A genetically linked pair of NLR immune receptors show contrasting patterns of evolution

Motoki Shimizu[1], Akiko Hirabuchi[1], Yu Sugihara[2], Akira Abe[1], Takumi Takeda[1], Michie Kobayashi[1,3], Yukie Hiraka[1], Eiko Kanzaki[1], Kaori Oikawa[1], Hiromasa Saitoh[1,4], Thorsten Langner[5], Mark J. Banfield[6], Sophien Kamoun[5] and Ryohei Terauchi[1, 2]

[1] Iwate Biotechnology Research Center, Kitakami, Iwate 024-0003, Japan

[2] Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University, Mozume, Muko, Kyoto 617-0001, Japan

[3] Institute of Agrobiological Sciences, National Agriculture and Food Research Organization (NARO), Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

[4] Department of Molecular Microbiology, Tokyo University of Agriculture, Setagaya-ku, Tokyo 156-8502, Japan

[5] The Sainsbury Laboratory, University of East Anglia, Norwich, NR4 7UH, UK

[6] Department of Biochemistry and Metabolism, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK

Correspondence to:
Motoki Shimizu (m-shimizu@ibrc.or.jp) and
Ryohei Terauchi (terauchi@ibrc.or.jp)

Abstract

Throughout their evolution, plant nucleotide-binding leucine-rich-repeat receptors (NLRs) have acquired widely divergent unconventional integrated domains that enhance their ability to detect pathogen effectors. However, the functional dynamics that drive the evolution of NLRs with integrated domains (NLR-IDs) remain poorly understood. Here, we reconstructed the evolutionary history of an NLR locus prone to unconventional domain integration and experimentally tested hypotheses about the evolution of NLR-IDs. We show that the rice (*Oryza sativa*) NLR Pias recognizes the effector AVR-Pias of the blast fungal pathogen *Magnaporthe oryzae*. Pias consists of a functionally specialized NLR pair, the helper Pias-1 and the sensor Pias-2, that is allelic to the previously characterized Pia pair of

NLRs: the helper RGA4 and the sensor RGA5. Remarkably, Pias-2 carries a C-terminal DUF761 domain at a similar position to the heavy metal–associated (HMA) domain of RGA5. Phylogenomic analysis showed that Pias-2/RGA5 sensor NLRs have undergone recurrent genomic recombination within the genus *Oryza*, resulting in up to six sequence-divergent domain integrations. Allelic NLRs with divergent functions have been maintained trans-species in different *Oryza* lineages to detect sequence-divergent pathogen effectors. By contrast, Pias-1 has retained its NLR helper activity throughout evolution and is capable of functioning together with the divergent sensor-NLR RGA5 to responds to AVR-Pia. These results suggest that opposite selective forces have driven the evolution of paired NLRs: highly dynamic domain integration events maintained by balancing selection for sensor NLRs, in sharp contrast to purifying selection and functional conservation of immune signaling for helper NLRs.

Keywords:
nucleotide-binding leucine-rich-repeat receptors (NLRs), paired NLR, integrated domains, evolution, rice, *Magnaporthe oryzae*

Significance statement

Plants have evolved sophisticated defense mechanisms to fend off pathogens. Plant nucleotide-binding leucine-rich repeat receptor (NLR) proteins play crucial roles in detecting pathogen molecules inside plant cells and mounting defense responses. Here, we identified the *Pias* gene from rice, which encodes the NLR pair Pias-1 "helper" and Pias-2 "sensor." These proteins function together to detect the pathogen molecule AVR-Pias of *Magnaporthe oryzae* and defend against rice blast disease. *Pias* is allelic to the previously reported *Pia* gene. A comparison of Pias/Pia alleles among *Oryza* species showed that Pias/Pia helper is evolutionarily and functionally conserved, whereas the Pias/Pia sensor shows highly dynamic evolution, with various host domains integrated into similar positions, allowing it to detect a wide variety of pathogen molecules.

Introduction

Plants are continually attacked by a multitude of microbial pathogens. Pathogens secrete effector molecules to enable the invasion of their hosts (Hogenhout et al. 2009). To counter this, plants have evolved a surveillance system that detects pathogen effectors inside the plant cell, leading to effector-triggered immunity (ETI; Jones and Dangl, 2006). Nucleotide-binding leucine-rich-repeat receptors (NLRs) play pivotal roles in ETI, which frequently leads to hypersensitive response (HR)–mediated cell death (Kourelis and van der Hoorn, 2018; Adachi et al. 2019a). NLR genes underwent lineage-specific expansions in most plant genomes (~150 in *Arabidopsis thaliana* and ~500 in rice (*Oryza sativa*)) and are among the most

variable genes in plants, pointing to strong selection pressure from pathogens (Clark et al. 2007; Van de Weyer et al. 2019). NLR proteins are characterized by a conserved nucleotide-binding (NB) domain and a leucine-rich-repeat (LRR) domain. NLRs are divided into two major groups depending on the type of N-terminal domain: NLRs with the N-terminal coiled-coil (CC) domain are called CNLs (CC-NLRs), and those with the N-terminal Toll-like domain are called TNLs (TIR-NLRs). CNLs are widespread in the plant kingdom. TIRs are grouped into the canonical TIR and TIR2 subclasses. Though TIR2-NB proteins are found in monocot plants, TNLs with canonical TIR domains have been detected in dicot but not in monocot plants (Pan et al. 2000; Sarris et al. 2016). ADP/ATP exchange at the NB domain (Bernoux et al. 2016; Wang et al. 2019a) and oligomerization of NLRs (Ma et al. 2020; Martin et al. 2020; Sharif et al. 2019; Tenthorey et al. 2017; Wang et al. 2019b; Zhang et al. 2015) trigger ETI signaling. The Arabidopsis CNL ZAR1 forms a pentamer "resistosome" complex after binding to a host protein and a pathogen effector. This leads to the protrusion of the N-terminal alpha helix, which perturbs the plasma membrane to trigger the HR (Wang et al. 2019ab), a process potentially mediated by Ca^{2+} influx (Bi et al. 2021). The Arabidopsis TNL RPP1 forms a tetramer "resistosome" upon binding to its cognate effector ATR1, followed by an increase in its NAD^{+}ase activity (Ma et al. 2020), which triggers cell death (Wan et al. 2019; Horsefield et al. 2019). These recent breakthroughs have started to reveal the biochemical links between NLR molecular structure and HR induction (Adachi et al. 2019b).

NLRs can function as singletons, in pairs or in networks (Adachi et al. 2019a). Arabidopsis ZAR1 (Wang et al. 2019a), RPP1 (Krasileva et al. 2010), and many other NLRs function as singleton NLRs and recognize AVRs directly or indirectly. Two NLR proteins encoded by genetically linked genes function together as paired NLRs. One of the paired NLRs frequently has a non-canonical domain called the integrated domain (ID). IDs are thought to have been derived from other host proteins (Kroj et al. 2016; Sarris et al. 2016). Examples of paired NLRs include RPS4/RRS1 in Arabidopsis (Williams et al. 2014; Le Roux et al. 2015) and RGA4/RGA5 (Okuyama et al. 2011; Cesari et al. 2013), Pik1/Pik2 (Ashikawa et al. 2008; Maqbool et al. 2015), and Pii1/Pii2 in rice (Takagi et al. 2013; Fujisaki et al. 2017). Multiple NLRs encoded by unlinked genes may function together; these network NLRs (Wu et al. 2017) include NRC2, NRC3, and NRC4 in Solanaceae plants. When NLRs function as a pair or network, one NLR is involved in the recognition of the AVR effector (the "sensor NLR"), whereas the other plays a role in signaling (the "helper NLR"). Recent studies have expanded our understanding of the genetic architecture of NLR pairs and networks; however, how they function together and how they evolved remain elusive.

A phylogenetic study of 4,184 NLRs in the genomes of seven Poaceae species including rice, wheat, and other grass species grouped the NLRs into 24 major clades, including three clades (MIC1, MIC2, and MIC3) containing the majority of NLRs with IDs (Bailey et al. 2018). NLRs in clade MIC1 (including RGA5) are characterized by a wide variety of IDs integrated into similar positions in the NLR after the LRR domain. Bailey et al. (2018) reported a conserved 43–amino acid motif between LRR and ID named the CID (conservation and association with IDs) motif. The authors proposed a possible evolutionary mechanism of ID generation and ID shuffling mediated by the CID motif. However, this study examined only a few genomes from each species, with a focus on inter-species diversity within the Poaceae, and provided limited

113 information about the recent diversification of NLRs within a single genus. In addition, experimental
114 validation of the hypotheses underpinning the functional diversity of NLRs with IDs remains limited.
115 Rice blast disease caused by blast fungus (*Magnaporthe oryzae* [syn. *Pyricularia oryzae*]) is a major
116 disease of rice that threatens world food security (Pennisi 2010). The best way to control this disease is to
117 deploy cultivars with resistance (*R*) genes. To date, 40 *R*-genes against blast have been reported, and 25 have
118 been cloned, most of which encode NLRs (Liu et al. 2010; Wu et al. 2015). However, the molecular
119 interactions between rice NLRs and blast AVRs are understood for only a small number of cases, including
120 Pita and AVR-Pita (Orbach et al. 2000; Jia et al. 2000), Pia and AVR-Pia/AVR1-CO39 (Cesari et al. 2013),
121 Pik and AVR-Pik (Kanzaki et al. 2012; Maqbool et al. 2015), Pii and AVR-Pii (Takagi et al. 2013; Fujisaki
122 et al. 2015; 2017), and Pizt and AVR-Pizt (Park et al. 2016, Wang et al. 2016).
123 The rice Pia pair, consisting of the NLR helper RGA4 and the NLR sensor RGA5, is one of the
124 most well-characterized NLR pairs (Okuyama et al. 2011; Cesari et al. 2013). Overexpression of *RGA4* in
125 *Nicotiana benthamiana* and rice protoplasts caused HR-like cell death in the respective plant species,
126 suggesting that the helper RGA4 is responsible for HR signaling. RGA4-mediated cell death was suppressed
127 by co-expression of RGA5, indicating that RGA5 negatively regulates RGA4-mediated defense signaling.
128 Finally, co-expression of RGA5, RGA4, and AVR-Pia triggered cell death via the direct binding of AVR-
129 Pia to HMA ID of the sensor RGA5 (Cesari et al. 2013). This study, together with a study of the Arabidopsis
130 RPS4/RRS1 pair (Williams et al. 2014; Le Roux et al. 2015), supported a negative regulation model for
131 paired NLRs in which helper NLRs function in HR signaling and sensor NLRs in the suppression of helper
132 NLR activity. Upon binding or modification of the sensor-NLR IDs by pathogen AVRs, this suppression is
133 released, allowing HR signaling to proceed. However, a recent study of rice Pikp paired NLRs suggests a
134 helper-sensor cooperation model. Expression of the helper NLR alone did not cause cell death in *N.*
135 *benthamiana* and co-expression of the helper and sensor was required for triggering effector-dependent HR-
136 like cell death (Zdrzałek et al. 2020).
137 Here, we describe the rice NLR pair Pias, which recognize the *M. oryzae* effector AVR-Pias. *Pias*
138 encodes the helper Pias-1 and the sensor Pias-2 and is allelic to the previously characterized *Pia* gene,
139 encoding the NLR pair helper RGA4 and sensor RGA5. Pias-2 carries a C-terminal DUF761 domain at a
140 similar position to the heavy metal–associated (HMA) domain of RGA5. We show that the Pias/Pia helper
141 NLR lineage is evolutionarily and functionally conserved, while its sensor NLR lineage shows highly
142 dynamic evolution with various host domains integrated into similar positions, possibly allowing it to detect
143 a wide variety of pathogen molecules.
144
145 **Results**
146
147 **Isolation of rice *Pias* NLR genes and the matching gene *AVR-Pias* from *Magnaporthe oryzae***
148 As part of a large genetic screen to identify novel rice blast resistance genes, we crossed *japonica*-type rice
149 cultivar Hitomebore to 20 different rice cultivars representing the worldwide genetic diversity of rice,
150 resulting in the generation of recombinant inbred lines (RILs) of the F_7-F_9 generations. We inoculated the

151 parental rice lines with a panel of *M. oryzae* isolates and recorded their resistance or susceptibility to each
152 isolate. Since Hitomebore was susceptible to *M. oryzae* isolate 2012-1 but *indica*-type accession WRC17
153 (cultivar Keiboba; Kojima et al. (2005)) was resistant to this isolate, we set out to isolate the resistance genes
154 in WRC17 rice against the 2012-1 pathogen. The 58 RILs derived from a cross between Hitomebore and
155 WRC17 segregated into 52 resistant and 6 susceptible lines (**Fig. 1A**), indicating that WRC17 likely contains
156 more than one locus conferring resistance against 2012-1.

157 To identify the resistance genes, we performed whole-genome sequencing and RNA-seq of WRC17
158 and conducted an association study using a bioinformatics pipeline we named "RaIDeN"
159 (https://github.com/YuSugihara/RaIDeN) using the RILs segregating for these phenotypes (**Fig. 1; Suppl.**
160 **Fig. 1**). We sequenced the genomes WRC17, Hitomebore, and six RILs showing susceptibility to the isolate
161 2012-1 on an Illumina DNA sequencer (**Suppl. Table 1**) and subjected the short reads of WRC17 to *de novo*
162 assembly with *DISCOVAR* (https://www.broadinstitute.org/software/discovar/blog/), resulting in the
163 WRC17 reference genome sequence (**Suppl. Table 2**). We also performed RNA-seq of WRC17 leaves that
164 had been inoculated with *M. oryzae* (2012-1). The RNA-seq reads were mapped to the WRC17 reference
165 genome, revealing 22,561 genes expressed from the WRC17 genome. The short reads obtained from
166 Hitomebore and six susceptible (S-) RILs were aligned to the genome sequences of the 22,561 expressed
167 genes. We reasoned that Hitomebore and the six S-RILs share the same DNA sequences in the candidate
168 genes responsible for their resistance that are different from the sequences in resistant WRC17 (**Suppl. Fig.**
169 **1**). Two types of DNA polymorphisms were considered: (1) presence/absence of the genes and (2) SNPs in
170 the genes. We identified 14 genes that were present in WRC17 but absent from Hitomebore and the six S-
171 RILs, in addition to 839 genes with shared SNPs among Hitomebore and the S-RILs that were different from
172 the sequences of WRC17.

173 From this group of 853 genes (**Suppl. Dataset 1**), we selected Resistance Gene Analogs (RGAs)
174 using "RGAugury" (Li et al. 2016), which predicts genes encoding putative NLRs, Receptor-Like Kinases
175 (RLKs), and Receptor Like Proteins (RLPs). This analysis identified 38 RGAs as the candidate genes (**Suppl.**
176 **Table 3**). Since most rice resistance genes against *M. oryzae* reported to date are *NLR*s, we focused on 18
177 *NLR* (11 *CNL* and 7 *NL*) genes as the candidate resistance genes of WRC17 against *M. oryzae* isolate 2012-
178 1. Among the 18 NLR genes, only one (*NL-04*) showed presence/absence polymorphisms, and the rest
179 contained SNPs associated with the phenotypes. We developed DNA markers in the candidate NLR genes
180 and studied their association with phenotypes using 58 RILs segregating for resistance and susceptibility to
181 the isolate 2012-1 (**Suppl. Fig. 2A**). This analysis showed that 16 NLRs were tightly linked to each other
182 and that the two other *NLR*s were linked to each other, suggesting that the two loci (designated *Pi-W17-1*
183 and *Pi-W17-2*) are involved in the resistance of WRC17 against 2012-1 (**Fig. 1B**). Genotyping of the 52
184 RILs using the markers located in the two loci, *Pi-W17-1* (*CNL-04*) and *Pi-W17-2* (*NL-07* and *W17.11718*),
185 suggested that the lines became resistant when either of the two loci contained the WRC17-type allele (**Fig.**
186 **1A; Suppl. Fig. 2A**).

187 The genomic position of *Pi-W17-1* corresponds to that of the previously reported *Pia* locus
188 (Okuyama et al. 2011), and the position of *Pi-W17-2* corresponds to that of the *Pita* (=*NL-07*) and *Ptr*

189 (=*W17.11718*) loci (Zhao et al. 2018; Orbach et al. 2000; Meng et al. 2020). Among the progeny showing
190 resistance against 2012-1, we selected 19 RILs containing *PiW17-1* but not *PiW17-2*. These lines shared 10
191 candidate NLR genes within the *PiW17-1* region. (**Fig. 1B; Suppl. Fig. 2B**). We performed RNAi-mediated
192 gene silencing of eight of these genes (encoding proteins over 900 amino acids long) (**Suppl. Fig. 2C**) using
193 the RIL HW-RIL7, which contains *Pi-W17-1* but lacks *Pi-W17-2*. When *CNL-04* or *CNL-05* was silenced,
194 its resistance against 2012-1 became compromised (**Suppl**. **Fig. 2CD**). We confirmed this result by
195 generating *CNL-04* and *CNL-05* knockout mutant lines by CRISPR/Cas9-mediated genome editing (**Fig.**
196 **1C; Suppl. Fig. 3**). These data suggest that *Pi-W17-1*-mediated resistance requires both the neighboring
197 NLRs *CNL-04* and *CNL-05*. Indeed, the position of *CNL-04* corresponds to that of *RGA4*, whereas the
198 position of *CNL-05* corresponds to that of *RGA5* of *Pia* (**Fig. 1D**; Okuyama et al. 2011). Interestingly, *CNL-*
199 *04* is similar to *RGA4* in terms of both structure and DNA sequence (96.6% DNA sequence identity), whereas
200 *CNL-05* has a distinct structure and a DNA sequence that diverged from *RGA5* (59.8% DNA sequence
201 identity) (**Suppl. Fig. 4**). *Pia RGA5* encodes a protein with a heavy metal–associated (HMA) domain in its
202 C terminus (Okuyama *et al*. 2011), whereas *CNL-05* encodes a protein with a 19–amino acid motif
203 corresponding to domain of unknown function 761 (DUF761) near its C terminus. The physical distance
204 between *CNL-04* and *CNL-05* is 8.7 kb, which is longer than that between *RGA4* and *RGA5* (3.7 kb) (**Fig.**
205 **1D**). In view of the substantial differences between *CNL-05* and *RGA5*, we decided to name this WRC17
206 allele *Pias*, *CNL-04* as *Pias-1*, and *CNL-05* as *Pias-2* (**Suppl. Fig. 5**).
207 To isolate the *AVR-Pias* avirulence gene cognate of rice NLR *Pias*, we performed an association
208 study of expressed genes encoding candidate effector proteins (See **Suppl. Fig. 6** and **Suppl. Table 4, 5** for
209 details). This analysis identified three genes (*G9141*, *G9435*, and *G9532*) as candidates of *AVR-Pias*. We
210 selected *M. oryzae* isolate Ao92-06-2, which is compatible with HW-RIL7, transformed it with each of the
211 candidate genes, and tested their interactions with HW-RIL7 (**Suppl. Fig. 7**). Transformation with one of
212 the candidate genes, *G9532*, rendered Ao92-06-2 incompatible with HW-RIL7 (**Suppl. Fig. 7A**), suggesting
213 that *G9532* is *AVR-Pias*. To validate this result, we generated a knockout mutant of *G9532* in the 2012-1
214 background, which became compatible with HW-RIL7 (**Fig. 1E**; **Suppl. Fig. 8**). These results indicate that
215 *G9532* is *AVR-Pias*, which is recognized by *Pias*. Both *Pias-1* and *Pias-2* are required for the recognition of
216 *AVR-Pias*, as knockout of either NLR gene abrogated resistance (**Suppl. Fig. 7B**). AVR-Pias is a 91–amino
217 acid protein with a secretion signal peptide (**Fig. 1F**). This protein contains a 12–amino acid Toxin18-like
218 motif (a feature of proteins belonging to the conotoxin O superfamily) in its C terminus.
219
220 **Various host domains are integrated into Pias/Pia sensor NLRs**
221 *Pias* and *Pia* are allelic to each other, and both are composed of a pair of NLRs. The helper NLRs (*Pias-1*
222 and *RGA4*) are conserved, whereas the sensor NLRs (*Pias-2* and *RGA5*) are divergent, with different
223 integrated domains. To explore the diversity and evolution of *Pias/Pia* NLRs in the entire *Oryza* genus, we
224 obtained the genomic DNA sequences of the *Pias/Pia* locus from 171 accessions representing 11 *Oryza*
225 species as well as four non-*Oryza* species of Poaceae, including *Setaria italica* (foxtail millet), *Panicum*
226 *hallii* (Hall's panicgrass), *Hordeum vulgare* (barley), and *Aegilops tauschii* (Tausch's goatgrass) (**Suppl.**

227 **Table 6**). To validate the gene structures, we used *RGA4/RGA5* and *Pias-1/Pias-2* genes as well as gene
228 models supported by RNA-seq for 10 *Oryza* accessions as queries to infer the gene models of 167 *Oryza*
229 accessions and four non-*Oryza* species using Exonerate software (http://www.ebi.ac.uk/~guy/exonerate)
230 (**Suppl. Table 7, 8**). The details of the gene model prediction pipeline and the results of RNA-seq alignment
231 to the Pias-2/RGA5 genes are given in **Suppl. Fig. 9**. Remarkably, Pias-2/RGA5 sensor proteins contain a
232 wide variety of IDs, with up to nine different domains, including HMA, DUF761, DUF677, Zinc_ribbon_12,
233 PKc_MAPKK (PKc_M), PKc_like, and WRKY, inserted into the identical position after the LRR domain
234 (**Fig. 2A**). Around the junction of the ID and LRR domains, we identified a conserved 145– to 146–amino
235 acid motif named the LRR-ID intervening (LII) region (**Fig. 2B**), which spans four LRR domains. The LII
236 region partially overlaps but is different from the previously reported CID motif (Bailey et al. 2018; **Suppl.**
237 **Fig. 10**).
238 We studied the frequency of different IDs in the 11 *Oryza* species used to reconstruct a phylogenetic
239 tree based on whole-genome sequences (**Fig. 2C; Suppl. Table 6, 8**). The IDs of the cultivated rice *O. sativa*
240 (44 samples) are shared by the HMA and DUF761 domains with similar frequencies. *O. rufipogon* (13
241 samples), the wild progenitor of *O. sativa*, contains DUF677 and PKc_M in addition to HMA and DUF761.
242 *Oryza* species belonging to the A-genome group (*O. sativa*, *O. rufipogon*, *O. barthii*, *O. glumaepatula*, and
243 *O. meridionalis*) contain a higher proportion of the DUF761 ID, suggesting its importance in their defense.
244 However, outside the A-genome species, the DUF761 ID is absent. Instead, PKc_M (*O. brachyantha* and *O.*
245 *punctata*), DUF677 (*O. officinalis*), and HMA (*O. australiensis*) are dominant, indicating that the integration
246 of the DUF761 domain into Pias-2/RGA5 likely occurred in the immediate ancestor of the A-genome species
247 that diverged from other *Oryza* lineages over 2.3 million years ago (TimeTree; http://www.timetree.org/).
248 A comparison of the genome sequences around *Pias/Pia* sensor-NLR genes between *O. punctata*
249 and *O. sativa* cultivar Nipponbare revealed an interesting conserved region (**Fig. 2DE**). Both of their coding
250 regions contain conserved LII sequences, whereas the IDs downstream of the LII are different: PKc_M for
251 *O. punctata* W1582 and DUF761 for *O. sativa* Nipponbare. However, in the 30.8-kb region downstream of
252 the *O. sativa* Nipponbare *Pias-2* gene, two regions (block **a** and **b**) share high DNA sequence similarity with
253 the LII and ID sequences of *O. punctata* W1582 Pias-2 homolog (**Fig. 2DE; Suppl. Fig. 11**): Block **a**
254 corresponds to region LII to exon-6, and block **b** corresponds to exon-7 to -10 of the *O. punctata* W1582
255 *Pias-2* homolog. Perhaps this conserved sequence downstream of the *O. sativa* Nipponbare *Pias* sensor is a
256 footprint of the replacement of the ID from PKc_M to DUF761 via homologous recombination at the LII
257 region (**Fig. 2F**). A survey of 51 A-genome *Oryza* accessions with the DUF761 ID revealed that the *O.*
258 *punctata* W1582 LII-ID-like sequence is widely conserved in *O. sativa*, *O. rufipogon*, *O. meridionalis*, and
259 an accession of *O. barthii* (Accession W1702), but not in the majority of *O. barthii* and *O. glumaepatula*
260 accessions (**Suppl. Table 9; Suppl. Fig. 12**). These results indicate that this possible recombination occurred
261 only once and that the footprint was probably lost in the latter two species. We did not detect similar
262 footprints in other sensor *Pias/Pia* NLRs with non-DUF761 IDs.
263
264 **Contrasting patterns of evolution of *Pias/Pia* sensors and helpers**

265 To explore the evolutionary patterns of the *Pias/Pia* NLR locus, we reconstructed phylogenetic trees of
266 *Pias*/*Pia* NLR pairs separately for the helper NLR (Pias-1 and RGA4) and the sensor NLR (Pias-2 and
267 RGA5) using the amino acid sequences of the helper NLRs (full-length protein) and sensor NLRs (full-
268 length protein except the ID domain) of 22 *Oryza* accessions and four accessions from other Poaceae genera.
269 Note that Pias/Pia helper and sensor NLRs belong to distantly related NLR clades (Bailey et al. 2018).
270 Overall, the helper NLR tree (**Fig. 3A**) is consistent with the species tree (**Fig. 2C**) with a few exceptions.
271 Species in the AA genome group (*O. sativa* subsp. *japonica*, *O. sativa* subsp. *indica*, *O. rufipogon*, *O. barthii*,
272 *O. glumaepatula*, and *O. meridionalis*) clustered together. Only the placement of *O. australiensis* (accession
273 W0008) is not congruent between the Pias-1/RGA4 helper NLR tree and the species tree. On the other hand,
274 the sensor-NLR tree is quite distinct from the helper NLR tree and is not congruent to the species tree.
275 Remarkably, the sensor NLRs show a higher level of divergence than the helper NLRs (**Fig. 3A**). The sensor
276 phylogeny includes two major clades (C1and C2) separated by deep branches. RGA5 of *O. sativa* cv.
277 Sasanishiki (Okuyama et al, 2011) belongs to the C1 clade, whereas Pias-2 of *O. sativa* WRC17 belongs to
278 the C2 clade. Four species of the AA genome group (*O. sativa* subsp. *japonica*, *O. sativa* subsp. *indica*, *O.*
279 *rufipogon*, *O. barthii*, and *O. meridionalis*) and *O. punctata* (BB genome) have sensor-NLR alleles from
280 both the C1 and C2 clades, whereas *O. australiensis* (EE genome), *O. granulata* (GG genome), and *O.*
281 *brachyantha* (FF genome) have C1 alleles, and *O. glumaepatula* (AA genome) and *O. officinalis* (CC
282 genome) have C2 alleles. Therefore, phylogenetic analysis of the sensor NLR of the *Pias/Pia* locus pointed
283 to trans-species polymorphism, which is reminiscent of the major histocompatibility (MHC) gene
284 polymorphism in vertebrates (Edwards and Hedrick, 1998).
285 To determine whether the paired NLRs have accumulated mutations at different rates, we calculated
286 the nucleotide diversity and Tajima's *D* (Tajima, 1989). The nucleotide diversities of the *Pias-2/RGA5*
287 sensor was markedly higher than that of the *Pias-1/RGA4* helper for each of the CC, NBS and LRR(-LII)
288 domains (**Fig. 3B**). Also, the three domains of *Pias-1/RGA4* helper and *Pias-2/RGA5* sensor had contrasting
289 negative and positive Tajima's *D*, respectively (**Fig. 3C**), indicating purifying selection in *Pias-1/RGA4* and
290 balancing selection in *Pias-2/RGA5*, especially in the LRR domain. These results suggest that the genetically
291 linked helper and sensor NLRs of the *Pias/Pia* locus in the genus *Oryza* have undergone contrasting modes
292 of evolution.
293 To determine the types of nucleotides substitutions in the *Pias-1/RGA4* helper and *Pias-2/RGA5* sensor,
294 we calculated d_N (non-synonymous mutations) and d_S (synonymous mutations) (**Fig. 3D**). Consistent with
295 higher nucleotide diversity and Tajima's *D*, we noted overall higher d_N and d_S in *Pias-2/RGA5*, notably in
296 NBS and LRR domains (**Fig. 3D**). We also observed that the pairwise d_N and d_S of *Pias-1/RGA4* was
297 constrained by the low genetic divergence in many of the examined pairs. Nonetheless, d_S was higher in
298 *Pias-1/RGA4* for CC domain than in *Pias-2/RGA5* (**Fig. 3D; Suppl. Fig. 13**). These results suggest that
299 different evolutionary patterns affect the different domains of *Pias-1/RGA4* and *Pias-2/RGA5*.
300
301 **Pias-1 helper functions with the RGA5 sensor to recognize AVR-Pia**

302 Pia comprises the helper NLR RGA4 and the sensor-NLR RGA5. The expression of RGA4 triggers HR cell
303 death in rice as well as *N. benthamiana*, and RGA5 suppresses RGA4-mediated cell death (Césari et al.
304 2014). Upon the binding of AVR-Pia to HMA ID of RGA5, this suppression is released, and HR-like cell
305 death is triggered (Césari et al. 2014). To address the functional conservation of helper/sensor NLRs of the
306 *Pias/Pia* locus, we examined the functions of helper NLRs by transiently overexpressing five Pias/Pia helper
307 NLRs in *N. benthamiana* using agroinfiltration (**Suppl. Fig. 14**). Only RGA4 supported strong cell death,
308 and other alleles including Pias-1 caused weaker cell death, indicating that the role of RGA4 as a strong HR
309 inducer is not typical among the tested helper NLRs (**Suppl. Fig. 14**). Next, we tested the effects of co-
310 expressing Pias-1 and Pias-2, which surprisingly resulted in stronger cell death than that caused by Pias-1
311 expression alone (**Suppl. Fig. 15**). Co-expression of *AVR-Pias* with *Pias-1* and *Pias-2* did not alter the level
312 of cell death (**Suppl. Fig. 15**). A co-immunoprecipitation analysis showed that Pias-1 and Pias-2 interact
313 (**Suppl. Fig. 16**), and a yeast two-hybrid assay showed that the CC-domains of Pias-1 and Pias-2 form homo-
314 and heterodimers (**Suppl. Fig.16**). These results suggest that Pias-1 and Pias-2 physically interact like
315 RGA4/RGA5, but their mode of action is different from that of the RGA4/RGA5 pair.

316 We next examined the effects of exchanging various helper and sensor Pias/Pia NLRs by performing
317 a transient expression assay in *N. benthamiana* leaves (**Fig. 4A, Suppl. Fig. 17**). Co-expression of the sensor
318 RGA5 with the two heterologous helpers (Pias-1 and RGA4-Ogr from *O. granulata*) suppressed HR-like
319 cell death, indicating that RGA5 can suppress the cell death induced by helper NLRs other than RGA4 (**Fig.
320 4A, Suppl. Fig. 17**). Furthermore, this RGA5-mediated suppression of the Pias-1-triggered HR was released
321 via co-expression of AVR-Pia, which resulted in cell death (**Fig. 4A; Suppl. Fig. 17**). A similar result was
322 obtained using RGA4-Ogr helper (**Fig. 4A; Suppl. Fig. 17**). These results indicate that the sensor RGA5
323 properly functions with helpers other than RGA4 to recognize AVR-Pia and mount the HR in *N.
324 benthamiana*.

325 Finally, we generated transgenic rice plants (HW-RIL7:35S-RGA5) expressing the *RGA5* transgene
326 driven by the 35S cauliflower mosaic virus (CaMV35S) promoter in the *Pias* (*Pias-1* and *Pias-2*)
327 background. We challenged these lines with two *M. oryzae* isolates: Ao92-62-2 harboring the *AVR-Pia*
328 transgene and Ao92-62-2 harboring the *AVR-Pias* transgene (**Fig. 4B; Suppl. Fig. 18**). Remarkably, HW-
329 RIL7:35S-RGA5 exhibited resistance against *M. oryzae* containing *AVR-Pia* (**Fig. 4B**). Although the HW-
330 RIL7:35S-RGA5 lines consistently showed resistance against *M. oryzae* containing *AVR-Pia*, their
331 resistance against *M. oryzae* containing *AVR-Pias* varied. We crossed an *rga4* mutant in the Sasanishiki
332 background (Sas1493; Okuyama et al. 2011) with HW-RIL7 and obtained F1-B plants. These plants,
333 harboring intact *RGA5* as well as *Pias-1* and *Pias-2*, now recognized and triggered resistance against *M.
334 oryzae* containing *AVR-Pia*, but not *AVR-Pias* (**Fig. 4B**). Similar results were obtained for F1-A plants
335 generated by crossing Sasanishiki (WT) with HW-RIL7 (**Fig. 4B**). These results suggest that *Pia* function is
336 dominant over *Pias* function in terms of the recognition of *AVR-Pia* and *AVR-Pias*.

337 These results suggest that Pias-1 helper functions together with RGA5 in rice to recognize and
338 mount resistance against *M. oryzae* containing *AVR-Pia*, suggesting that the helper function has been
339 conserved over the long history of Pias/Pia evolution.

340

341 **Discussion**

342

343 Our study investigated the evolution of a pair of genetically linked NLRs in the genus *Oryza* and provided
344 experimental evidence that the two paired NLRs have evolved in dramatically contrasting fashions. This
345 study points to the evolution of a modular architecture of paired NLRs. Division of roles between a conserved
346 helper NLR for signaling and a divergent sensor NLR with a cassette-like receptor domain for pathogen
347 sensing may have given plants the ability to efficiently fend off rapidly evolving microbe pathogens.

348 We identified and functionally characterized the rice R-gene *Pias*. This gene encodes the paired
349 NLRs Pias-1 helper and Pias-2 sensor, which recognizes the *M. oryzae* effector AVR-Pias. *Pias* is allelic to
350 the well-studied *R*-gene *Pia*, encoding the NLRs RGA4 helper and RGA5 sensor (Okuyama et al. 2011),
351 which recognizes the effectors AVR-Pia (Yoshida et al. 2009) and AVR1-CO39 (Cesari et al. 2013). The
352 allelic sensor-NLRs Pias-2 and RGA5 carry different domains at their C termini. In RGA5, the integrated
353 domain HMA directly binds to and recognizes two *M. oryzae* effectors, AVR-Pia and AVR1-CO39 (Cesari
354 et al. 2013). We have not yet detected direct binding between the DUF761 ID of Pias-2 and AVR-Pias
355 despite several attempts. Perhaps the recognition of AVR-Pias by Pias-2 requires other host components;
356 indeed, the recognition of AVR-Pii by Pii-2 requires the rice protein OsExo70-F3, which binds to both AVR-
357 Pii and Pii-2 (Fujisaki et al. 2015, 2017).

358 What is the origin of the integrated DUF761 domain of Pias-2? The rice sensor-NLRs RGA5 and
359 Pik-1 contain HMA domains as IDs. The ID HMA shares high amino acid sequence similarity with rice
360 small heavy metal–associated domain proteins (sHMAs: Oikawa et al. 2020; Białas et al. 2021). We revealed
361 that the *M. oryzae* effector AVR-Pik binds to and stabilizes sHMA proteins, likely to promote pathogen
362 infection (Oikawa et al. 2020; Maidment et al. 2021). To identify the proteins that provide the DUF761
363 domain to Pias-2 ID, we performed BLAST searches against the rice protein database using the short Pias-
364 2 DUF761 sequence (19 amino acids) as a query (**Suppl. Fig.19**). This identified 15 proteins with a similarity
365 threshold of E < 10. Most of the 15 proteins contained DUF761 at their C termini. Only a few functional
366 studies of DUF761-containing proteins have been performed. A study in cotton (*Gossypium hirsuta*) showed
367 that GhCFE1A, containing DUF761 and DUF4408 domains, binds to actin proteins and localizes to the
368 endoplasmic reticulum (ER) upon overexpression in *N. benthamiana* (Lv et al. 2015). Knockdown of
369 *GhCFE1A* did not cause any phenotypic changes, while its overexpression led to delayed cotton fiber cell
370 elongation. These results suggest that GhCFE1A is a linker protein that mediates the formation of the ER
371 network and actin cytoskeleton. Arabidopsis *A70*, encoding a DUF761-containing protein, is specifically
372 induced in the incompatible interaction with *Pseudomonas syringae*, but not in the compatible interaction
373 (Truman et al. 2007). Knockout of the Arabidopsis DUF761-containing protein gene *DUF761-1* did not
374 have any phenotypic effects, whereas overexpressing *DUF761-1* altered plant morphology and resulted in a
375 constitutive defense response, leading to enhanced resistance against *P. syringae* (Zhang et al. 2019). These
376 findings suggest that DUF761-containing proteins function in defense, presumably mediated by the actin-
377 ER network. Previous studies of the NLRome of Arabidopsis (Van de Weyer et al, 2019) and NLR-ID of

various members of the plant kingdom (Sarris et al. 2016) showed that DUF761 is one of the most common domains integrated into NLR as the ID. In Arabidopsis, DUF761 was integrated into TIR-NLR, and the NLRs with DUF761-ID are in almost all cases paired with helper NLRs (Van de Weyer et al, 2019). These findings suggest that the DUF761 domain is a major target of pathogen effectors and has been frequently integrated into NLR. Future studies should investigate how AVR-Pias is recognized by Pias and how AVR-Pias interferes with host cellular processes by its possible interaction with DUF761-containing proteins.

Divergent sensor NLRs in the Pias/Pia lineage

The Pias-2/RGA5 sensor-NLR lineage is extremely divergent among *Oryza* species, with up to six different ID motifs integrated at their C termini (**Fig. 2**). These hugely divergent IDs may mediate the detection of a diversity of effectors from the blast fungus and possibly other pathogens. We hypothesize that the diversity of the Pias-2/RGA5 lineage has been maintained by natural selection to maintain various IDs that recognize the invasion of pathogens by directly binding to effectors or guarding host factors that are modified by effectors.

Within *Oryza* species, there are two major clades, C1 and C2. Notably, alleles from both clades are maintained within the species *O. sativa*, *O. rufipogon*, *O. barthii*, *O. meridionalis*, and *O. punctata*. The observed trans-species allelic divergence and their roles in detecting pathogen molecules are similar to those of the major host incompatibility (MHC) locus of vertebrates (Edwards and Hedrick, 1998). Individuals with higher heterozygosity at the MHC locus might have higher fitness (overdominance) due to their ability to bind to a larger number of pathogen peptides (Takahata and Nei, 1990; Piertney and Oliver, 2006). Perhaps in the ancestral outcrossing *Oryza* species, heterozygous plants with a larger repertoire of NLRs with different IDs had a selective advantage against a multitude of pathogens. It is also possible that frequency-dependent selection helped maintain this polymorphism. When the frequency of a pathogen effector in a population increases, the frequency of an allele for a cognate sensor NLR will increase, resulting in a reduction in the frequency of pathogen alleles in the population. In turn, the frequency of another effector gene will increase in the pathogen, and the frequency of the cognate sensor NLR will increase. According to this Red Queen model, the allele frequencies of both effector and sensor NLRs oscillate and may be maintained for a long time by balancing selection (Takahata and Nei, 1990; Woolhouse et al. 2002; Terauchi and Yoshida, 2010). In summary, the highly divergent evolution of Pias/Pia sensor NLRs with variable IDs seems driven by the fitness gain obtained by an enhanced recognition capability of pathogen effectors.

Genetic mechanism of ID switching in Pias/Pia NLRs

The Pias/Pia sensor NLRs contain various ID sequences at the identical position downstream of the LRR domain. This suggests the presence of a mechanism for “cassette”-like exchange of IDs between sensor NLRs. Between the LRR and ID, we identified a highly conserved stretch of 145–amino acid sequences named LII. Bailey et al. (2018) reported a 43–amino acid CID motif conserved in the region between the LRR and ID of the MIC1 NLR clade by studying the NLRs of seven Poaceae species. The LII motif

416 identified encompasses the CID motif (**Suppl. Fig. 10**) but extends to the N-terminal direction by
417 approximately 100 amino acids. Bailey hypothesized that the CID could serve as a recombination point of
418 integration of genomic sequences. Despite the difference in the conserved motif, our data basically support
419 Bailey (2018)'s hypothesis that this region serves as the recombination point for the integration of
420 endogenous sequences matching various protein domains. In support of this idea, the downstream sequence
421 of a Pias sensor NLR with DUF761 contains a DNA sequence similar to the LII and ID regions of the *O.*
422 *punctata* Pias sensor homolog. Perhaps the PKc_M ID of the original *O. punctata* Pias was replaced by the
423 DUF761 ID, and this switch caused the translocation of PKc_M to the region downstream of the sensor NLR
424 gene (**Fig. 2F**). In view of the high sequence conservation between *O. punctata* PKc_M ID and the
425 downstream sequence, it is also possible that the downstream sequence was functional until the recent past,
426 the Pias sensor contained the dual IDs DUF761 and PKc_M in the same molecule, or Pias switched between
427 these two IDs, possibly via alternative splicing. Future studies should address the mechanism of LII-
428 mediated recombination.
429
430 **Function of *Pias/Pia* paired NLRs**
431
432 Pia has been extensively studied (Césari et al. 2014) and serves as a paradigm for paired NLRs together with
433 the paired NLRs RPS4/RRS1 (Williams et al. 2014). In these two cases, the helper NLR is regarded as a cell
434 death inducer and the sensor NLR as a suppressor that maintains the complex in an inactive state when the
435 pathogen is absent. Once the sensor has been modified by direct binding (RGA5) or modification (RRS1) of
436 the ID by pathogen effectors, the suppression of the helper is released and HR-like cell death occurs.
437 However, Pias NLRs do not function according to this model. In the *N. benthamiana* assay, Pias-1 functioned
438 as a weak cell death inducer and Pias-2 did not suppress Pias-1-mediated cell death. Indeed, Pias-1 and Pias-
439 2 together triggered stronger cell death than that caused by Pias-1 alone. A recent functional study of rice
440 Pikp, another paired NLR, also showed that the helper NLR alone does not cause cell death in the *N.*
441 *benthamiana* system and that the helper and sensor cooperate to trigger HR (Zdrzałek et al. 2020). Therefore,
442 the hypothesized functional roles of the helper as a cell death inducer and the sensor as a cell death suppressor
443 as well as a detector of effector molecules may not be as prevalent as assumed. The system of negative
444 regulation of a cell death inducer by a suppressor encoded by genetically separate factors carries tremendous
445 risks given that a loss-of-function mutation in the suppressor gene kills the carrier cells and incurs a genetic
446 load. Therefore, such an extreme negative regulation system is unlikely to be maintained over a long period
447 of evolution. We predict that cooperative NLRs in pairs or networks are more prevalent (Wu et al. 2017). It
448 is also possible that the Pias/Pia-paired NLR system is regulated by additional components in rice cells that
449 are absent from *N. benthamiana*. It would be interesting to experimentally determine what proportions of
450 paired NLRs function in negative regulation and in cooperation using *N. benthamiana* transient expression
451 assays. Further studies are needed to decipher the full regulatory network of Pias/Pia-paired NLR-mediated
452 resistance.
453

454 **Helper NLRs in the Pias/Pia lineage are functionally conserved**
455
456 Phylogenetic reconstruction of the Pias/Pia NLR locus revealed that the helper NLR Pias-1/RGA4 is
457 conserved, whereas the sensor NLR Pias-2/RGA5 is highly divergent (**Fig. 3**). An Arabidopsis NLRome
458 study showed that some NLR pairs coevolved, with the phylogenetic trees of helper and sensor NLRs
459 corresponding (Van de Weyer, 2019). Our findings for Pias/Pia NLRs do not align with these observations.
460 A functional study in *N. benthamiana* showed that Pias-1-mediated cell death was suppressed by RGA5 and
461 that Pias-1 together with RGA5 function in the recognition of AVR-Pia, leading to cell death (**Fig. 4A**).
462 Similarly, an RGA4 homolog of *O. granulata* that is phylogenetically distant from *O. sativ*a also functions
463 with RGA5 to recognize AVR-Pia (**Fig. 4A**). Moreover, a HW-RIL7 rice line harboring *Pias* as well as an
464 *RGA5* transgene recognized AVR-Pia (**Fig. 4B**). These results suggest that the function of the Pias-1/RGA4
465 helper lineage is conserved, which is in line with its conserved amino acid sequences. It is possible that the
466 separation of the roles of NLRs between the conserved helper and divergent sensor allowed for higher
467 flexibility of pathogen recognition compared to singleton NLRs. This flexibility would allow the plant to
468 cope with the rapid evolution of pathogens, which exhibit larger population sizes and shorter generation
469 times than the host plants. The functional understanding of the modular structure of paired NLRs revealed
470 here provides a basis for engineering NLRs to detect various effectors and to confer resistance to crops
471 against pathogens.
472
473
474 **Materials and Methods**
475
476 **Rice pathogenicity assays**
477 Rice leaf blade punch inoculation was performed using the *M. oryzae* isolates. A conidial suspension (3 ×
478 10^5 conidia mL^{-1}) was punch inoculated onto a rice leaf 1 month after seed sowing. The inoculated plants
479 were placed in a dew chamber at 27°C for 24 h in the dark and transferred to a growth chamber with a 16-
480 h-light/8-h-dark photoperiod. Disease lesions were scanned 10 days post-inoculation (dpi), and lesion size
481 was measured using Image J software (Schneider et al. 2012).
482
483 **RNA-seq of rice and barley leaves infected with *M. oryzae* 2012-1 isolate**
484 Total RNA was extracted from rice and barley infected leaves using an SV Total RNA Isolation System
485 (Promega, WI, USA). One microgram of total RNA was used to prepare each sequencing library with an
486 RNA Sample Prep Kit v. 2 (Illumina, CA, USA). The two types of libraries, created from infected barley
487 and rice leaves, were sequenced by paired-end (PE) and single-end (SE) sequencing using the NextSeq 500
488 platform.
489
490 **DNA-seq for the RaIDeN pipeline**

491 Genomic DNA was extracted from WRC17, Hitomebore, and S-RIL leaves using a NucleoSpin Plant II Kit
492 (Macherey Nagel Co, Düren, Germany). Libraries for PE short reads were constructed using an Illumina
493 TruSeq DNA LT Sample Prep Kit (Illumina, CA, USA). The PE library of WRC17 and other libraries were
494 sequenced on the Illumina MiSeq and HiSeq 4000 platforms.
495
496 **Generation of *PiW17-1* with knocked-down expression of candidate NLR genes in rice and**
497 **quantitative RT-PCR**
498 The eight types of gene knockdown (RNAi) constructs (pANDA-*PiW17-1* candidate NLRs) were generated
499 by PCR amplification of a specific fragment of each *PiW17-1* candidate NLR gene from WRC17 WT cDNA.
500 The sequences were cloned into the Gateway vector pENTR/D-TOPO (Invitrogen, CA, USA) and
501 transferred into recombination sites of the pANDA vector (Miki and Shimamoto, 2004) using LR Clonase
502 (Invitrogen, CA, USA). The resulting vectors with eight types of pANDA-*PiW17-1* candidate NLR genes
503 were introduced into *Agrobacterium tumefaciens* (strain EHA105) and used for *A. tumefaciens*–mediated
504 transformation of HW-RIL7 following the method of Okuyama et al. (2011). Total RNA was extracted from
505 leaves using an SV Total RNA Isolation System (Promega, WI, USA) and used for quantitative RT-PCR
506 (qRT-PCR). cDNA was synthesized from 500 ng total RNA using a Prime Script RT Reagent Kit (Takara
507 Bio, Otsu, Japan). qRT-PCR was performed using a StepOne Real-time PCR Instrument (Applied
508 Biosystems, CA, USA) with KAPA SYBR FAST PCR Master Mix (Kapa Biosystems, MA, USA). Melting
509 curve analysis (from 60 to 95°C) was included at the end of the cycles to ensure the consistency of the
510 amplified products. The comparative Ct (ΔΔCt) method was used to calculate the expression of *CNL-04*
511 (*CNL-05*) relative to the rice *ACTIN* gene (*LOC_Os03g50885*) as an internal control. The data presented are
512 the average and standard deviations from three experimental replications. The primers used to generate the
513 RNAi construct and for qRT-PCR are listed in **Suppl. Table 10**.
514
515 **Generation of rice mutants of *CNL-04* and *CNL-05* by CRISPR/Cas9-mediated genome editing**
516 Rice knockout mutants of *CNL-04* and *CNL-05* were generated using the CRISPR/Cas9 system developed
517 by Mikami et al. (2015). Sense and antisense target sequences were designed using the web-based service
518 CRISPR direct (http://crispr.dbcls.jp), annealed, and cloned into the pU6::ccdB::gRNA cloning vector
519 following digestion with *Bbs*I as the target sequence. The target sequence with the *OsU6* promoter was
520 cloned into the pZH::gYSA::MMCas9 vector following digestion with *Asc*I and *Pac*I. The resulting vectors
521 (pZH::gYSA::MMCas9-*CNL-04* and -*CNL-05*) were introduced into *A. tumefaciens* (strain EHA105) and
522 used for *A. tumefaciens*–mediated transformation of HW-RIL7 following the method of Okuyama et al.
523 (2011). The resulting regenerated T0 plants were sequenced, and the mutation type was confirmed using
524 primers listed in **Suppl. Table 10**.
525
526 **Genetic complementation of the candidate AVR-Pias**
527 Three candidate gene constructs (pCB1531-*AVR-Pias* candidate) were generated by PCR amplification of
528 the coding sequences of the *AVR-Pias* candidate genes from cDNA prepared from *M. oryzae* 2012-1-infected

529 barley leaves. The sequences were digested with *Xba*I and *BamH*I and cloned into pCB1531-pex22p-EGFP
530 (Yoshida et al. 2009) that had been linearized by digestion with *Xba*I and *Bam*HI. The resulting vectors were
531 used to transform Ao92-06-2 (lacking *AVR-Pias*) following the method of Sweigard et al. (1997). *M. oryzae*
532 isolate 2012-1 mutated in *G9532* was generated using the CRISPR/Cas9 system developed by Arazoe et al.
533 (2015). Sense and antisense target sequences were designed using the web-based service CRISPR direct
534 (http://crispr.dbcls.jp), annealed, and cloned into the pCRISPR/Cas-U6-1 cloning vector following the
535 method of Arazoe et al. (2015). To generate the targeting vector TV-*G9532*, the 5′ flanking region of *G9532*
536 was amplified and cloned into pCB1636 (Sweigard et al. 1997) containing a hygromycin resistance gene
537 that had been linearized by inverse PCR using primers pCB1636iv2fwd and pCB1636iv2rev as described
538 by Shimizu et al. (2019) using in-fusion cloning (Clontech, Madison, WI, USA). Subsequently, the 3′
539 flanking region was amplified and cloned into a plasmid containing the 5′ flanking region that had been
540 linearized by inverse PCR using primers pCB1636iv1fwd and pCB1636iv1rev as described by Shimizu et
541 al. (2019) using in-fusion cloning (Clontech, Madison, WI, USA). The resulting vectors were used to
542 transform the 2012-1 isolate (containing *AVR-Pias*) following the method of Sweigard et al. (1997). The
543 primers used for construct generation are listed in **Suppl. Table 10**.
544
545 **Genome sequences used for the study**
546 Genome sequences of 171 accessions of Poaceae plants were used, including 167 *Oryza* accessions as well
547 as one accession each from *Aegilops tauschii*, *Hordeum vulgare*, *Panicum hallii*, and *Setaria italica*. These
548 also included sequences of *O. sativa* 'WRC17' (this study) and *O. sativa* 'Sasanishiki' (Okuyama et al. 2011).
549 Genome sequences of 66 accessions were obtained from public databases: 52 accessions of *O. sativa* and *O.*
550 *rufipogon* (Zhao et al. 2018), *O. sativa* subsp. *indica* 'ShuHui498' (Du et al. 2017), *O. sativa* subsp. *japonica*
551 'Nipponbare' (IRGSP. 2004), *O. rufipogon* 'W1943' (National Center for Gene Research), 5 accessions
552 from *O. barthii*, *O. glumaepatula*, *O. meridionalis*, *O. punctata*, and *O. brachyantha* (The Oryza Map
553 Alignment Project (OMAP)), *O. officinalis* 'W0002' (National Institute of Genetics, Japan), *O. granulata*
554 'W0067B' (Wu et al. 2018), *Ae. tauschii* 'AL8/78' (Zimin et al. 2017), *H. vulgare* 'Morex' (Mascher et al.
555 2017), *P. hallii* 'FIL2' (DOE Joint Genome Institute), and *S. italica* 'Yugu18' (Bennetzen et al. 2012) (see
556 **Suppl. Table 6** for details). For 101 accessions of the wild *Oryza* species (*O. barthii*, *O. glumaepatula*, *O.*
557 *meridionalis*, *O. punctata*, *O. officinalis, O. brachyantha*, and *O. granulata*), NGS reads (fastq format) of
558 whole-genome sequences were retrieved from the NGS NCBI database (**Suppl. Table 6**) and used for *de*
559 *novo* assembly by MaSuRCA (Zimin et al. 2013). For the two wild *Oryza* accessions, *O. glumaepatula*
560 W2184 and *O. punctata* W1582, DNA sequencing was performed using Oxford Nanopore Technology
561 (ONT) using genomic DNA extracted from their leaves. Sequencing was performed using the PromethION
562 system with a FLO-PRO002 flow cell (ONT). Base calling of ONT reads was performed on FAST5 files
563 using Guppy (ONT). Subsequently, low-quality reads were filtered out, and *de novo* assembly was
564 performed using NECAT software (https://github.com/xiaochuanle/NECAT/). To further improve the
565 accuracy of the assembly, Racon software (https://github.com/lbcb-sci/racon) was applied twice, and
566 Medaka (https://github.com/nanoporetech/medaka) was used to correct mis-assembly. One round of

567 consensus correction was performed using BWA (Li and Durbin, 2010) and HyPo
568 (https://github.com/kensung-lab/hypo) on Illumina short reads for the accessions.
569
570 **Gene annotation and detection of Pias/Pia orthologs**
571 To infer the protein coding regions of *Pias-1/RGA4* and *Pias-2/RGA5* genes, and to obtain information about
572 their ID sequences, we used the pipeline shown in **Suppl. Fig. 9**. We retrieved *Pias-1/RGA4* and *Pias-2*
573 */RGA5* gene models from the genome sequences and RNA-seq data publicly available for seven *Oryza*
574 samples (*O. barthii* 'IRGC105608', *O. glumaepatula* 'GEN1233_2', *O. meridionalis* 'W2112', *O. punctata*
575 'IRGC105690', *O. australiensis* 'W0008', *O. brachyantha* 'IRGC101232', and *O. granulata* 'W0067B').
576 For *O. rufipogon* 'W1943' and *O. officinalis* 'W0002', genome sequences were publicly available, but we
577 performed new RNA-seq analyses to improve gene prediction. For the sample *O. punctata* 'W1582', we
578 performed genome sequencing and RNA-seq analyses. We also used the *RGA4/RGA5* gene model of *O.*
579 *sativa* cv. Sasanishiki (Okuyama et al. 2011) and the *Pias-1/Pias-2* gene model of *O. sativa* cv. Keiboba
580 (this study). The gene models of *Pias-1/RGA4* and *Pias-2/RGA5* of these 12 *Oryza* samples were used as
581 queries to annotate IDs in the genome assembly of 167 *Oryza* samples using Exonerate
582 (http://www.ebi.ac.uk/~guy/exonerate). However, 10 samples of *O. glumaepatula* and six samples of *O.*
583 *brachyantha* did not match to known domains. Therefore, we incorporated RNA-seq data for each sample
584 for the two species, resulting in the annotation of the Zinc_ribbon_12 (*O. glumaepatula*) and HMA (*O.*
585 *brachyantha*) IDs. In the next round, we used the gene models of 12 samples used in the first round of
586 Exonerate as well as two new samples (*O. glumaepatula* W2184 and *O. brachyantha* W0655) as queries to
587 infer IDs in the assembled genomes of 167 *Oryza* samples (**Suppl. Table 6, 7, 8**). This resulted in the
588 identification of five known domains and one unknown domain in the *Oryza* ID sequences.
589
590 **Amino acid sequences and accession numbers of Pias/Pia orthologs**
591 Amino acid sequences of Pias/Pia homologs used in this study were retrieved from the Eukaryotic Genome
592 Annotation of the NCBI database (Du et al. 2017; Mascher et al. 2017); accession numbers for Pias-1/RGA4
593 homologs and Pias-2/RGA5 homologs are as follows: XP_015617251.1 and XP_015617810.1 for *O. sativa*
594 subsp. *japonica* 'Nipponbare', OsR498G1119642600.01 and OsR498G1119642700.01 for *O. sativa* subsp.
595 *indica* 'ShuHui498', XP_020148260.1 and XP_020148256.1 for *Aegilops tauschii* 'AL8/78',
596 HORVU.MOREX.r2.4HG0288000.1 and HORVU.MOREX.r2.4HG0288010.1 for *Hordeum vulgare*
597 'Morex', XP_025826635.1 and XP_025827327.1 for *Panicum hallii* 'FIL2', and XP_004979045.1 and
598 XP_004979046.2 for *Setaria italica* 'Yugu18'.
599
600 **Phylogenetic analysis of Pias-1/RGA4 and Pias-2/RGA5 orthologs**
601 The protein sequences of the Pias-1/RGA4 and Pias-2/RGA5 orthologs were aligned using webPRANK
602 (https://www.ebi.ac.uk/goldman-srv/webprank/) (Löytynoja et al. 2010). We used the whole amino acid
603 sequences of the Pias-1/RGA4 orthologs but only the partial amino acid sequences (CC-NB-LRR-LII
604 domains) of the Pias-2/RGA5 orthologs due to the very low sequence similarity after the LII domains. A

605 maximum likelihood tree was constructed with IQ-TREE v2.0.3 (Minh et al. 2020) using 1,000 ultrafast
606 bootstrap replicates (Hoang et al., 2018). The models were automatically selected by ModelFinder
607 (Kalyaanamoorthy et al. 2017) in IQ-TREE (Minh et al. 2020). ModelFinder (Kalyaanamoorthy et al. 2017)
608 selected "JTT + G4" for Pias-1/RGA4 orthologs and "JTT + R2" for Pias-2/RGA5 orthologs. Finally, the
609 phylogenetic trees were drawn with FigTree v1.2.3 (http://tree.bio.ed.ac.uk/software/figtree/).
610
611 **Analysis of DNA polymorphisms, *dN*, and *dS***
612 The coding DNA sequences (CDSs) of the *Pias-1/RGA4* and *Pias-2/RGA5* orthologs were aligned using the
613 codon-based aligner MACSE v2.05 (Ranwez et al. 2018). We applied MACSE v2.05 not only to the entire
614 CDS of *Pias-1/RGA4* and CC-NBS-LRR-LII of *Pias-2/RGA5* but also to each domain (CC, NBS, and LRR[-
615 LII]) using default parameters. We evaluated DNA polymorphisms of the *Pias-1/RGA4* and *Pias-2/RGA5*
616 orthologs, calculating the nucleotide diversity (π) and Tajima's *D* (Tajima, 1989) using MEGAX v10.2.4
617 (Kumar et al. 2018). Then, for each alignment, the maximum likelihood trees were constructed using IQ-
618 TREE v2.0.3 with 1,000 ultrafast bootstrap replicates (Hoang et al., 2018) and ModelFinder
619 (Kalyaanamoorthy et al. 2017). Based on these alignments and trees, we calculated the pairwise *dN* and *dS*
620 using the YN00 program (Yang and Nielsen, 2000) in PAML v4.8 (Yang 1997).
621
622 **Expression constructs used in the cell death assay**
623 Expression constructs for five types of helper-NLRs, Pias-1 (*O. sativa* subsp. *Indica*, WRC17), RGA4 (*O.*
624 *sativa* subsp. *Japonica*, Sasanishiki), RGA4-Oru (*O. rufipogon* accession W1943), RGA4-Oau (*O.*
625 *australiensis* accession W0008), and RGA4-Ogr (*O. granulata* accession W0067B) (pCambia1300S-
626 "helper-NLR":HA), were generated by PCR amplification of the coding sequences from cDNA generated
627 from leaf material and cloned into the binary vector pCambia1300S (http://www.cambia.org) that had been
628 linearized by digestion with *Pst*I and *Spe*I by In-fusion cloning. Expression constructs for two types of sensor
629 NLRs, Pias-2 (*O. sativa* subsp. *Indica*, WRC17) and RGA5 (*O. sativa* subsp. *Japonica*, Sasanishiki)
630 (pCambia1300S-FLAG: "sensor NLR"), were generated by PCR amplification of the coding sequences from
631 cDNA generated from leaf material and cloned into the binary vector pCambia1300S
632 (http://www.cambia.org) that had been linearized by digestion with *Sal*I and *Pst*I by In-fusion cloning. The
633 resulting vectors were introduced into *A. tumefaciens* (strain GV3101). The primers used to generate the
634 expression constructs are listed in **Suppl. Table 10**.
635
636 **Immunoblot analysis**
637 Proteins were expressed in *Nicotiana benthamiana* leaves and extracted from leaf tissue (approximately 100
638 mg) in 200 µL of extraction buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, and 0.1 % NP-40 [v/v]). The
639 extracts were subjected to SDS-PAGE, followed by immunoblotting using anti-FLAG-HRP antibody, anti-
640 HA-HRP antibody, or anti-AVR-Pia antibody. Anti-mouse IgG-HRP was used as a secondary antibody
641 following incubation with anti-AVR-Pia antibody. Antibody-antigen complexes were detected using a

642 luminescent image analyzer (ImageQuant LAS-4000) (Cytiva, Tokyo, Japan). Equal loading of proteins in
643 the PAGE gel was confirmed by Coomassie blue staining.
644
645 **Cell death assay in *N. benthamiana***
646 Transient expression of Pias/Pia allelic NLR and AVR-Pia was performed by infiltrating 4- to 5-week-old
647 *N. benthamiana* plants with *A. tumefaciens* carrying the expression vector. *A. tumefaciens* suspensions in
648 infiltration buffer (10 mM MES, 10 mM $MgCl_2$, and 150 μM acetosyringone, pH 5.6) were adjusted to the
649 densities shown in **Suppl. Table 11**. The autofluorescence value under UV light was scored using a
650 luminescent image analyzer (ImageQuant LAS-4000) (Cytiva, Tokyo, Japan).
651
652 **Yeast two-hybrid assay**
653 To examine the protein–protein interactions between Pias-1:CC_{182} and Pias-2:CC_{177}, a yeast two-hybrid
654 assay was performed as described previously (Kanzaki et al. 2012). The CC domain of Pias-1 (Pias-1:CC_{182})
655 was amplified by PCR, digested with *Eco*RI and *Bam*HI, and cloned into pGADT7 (prey) or pGBKT7 (bait)
656 vector (Clontech, Madison, WI, USA) that had been linearized by digestion with *Eco*RI and *Bam*HI by in-
657 fusion cloning. The CC domain of Pias-2 (Pias-2:CC_{177}) was amplified by PCR, digested with *Sfi*I, and
658 cloned into pGADT7 or pGBKT7 vector that had been linearized by digestion with *Sfi*I. GFP (as negative
659 control) was amplified by PCR and cloned into pGADT7 or pGBKT7 vectors that had been linearized by
660 digestion with *Eco*RI and *Bam*HI by in-fusion cloning. The various combinations of bait and prey vectors
661 were transformed into yeast strain AH109 using the PEG/LiAc method. To detect protein–protein
662 interactions, a 10-fold dilution series (×1, $\times 10^{-1}$, $\times 10^{-2}$) of yeast cells (×1: OD_{600} = 1.0) was spotted onto
663 basal medium lacking Trp, Leu, Ade, and His (-HTLA) but containing 5-bromo-4-chloro-3-indolyl α-D-
664 galactopyranoside (Clontech, Madison, WI, USA). As a control, yeast growth on basal medium lacking Trp
665 and Leu (–TL) was also checked. To assess the protein accumulation in yeast cells, each transformant was
666 propagated in liquid basal medium lacking Trp and Leu with gentle shaking at 30°C overnight. Yeast cells
667 from 10 mL medium were collected, and 100 mg of yeast cells was treated with 400 μL of 0.3 N NaOH for
668 15 min at room temperature. The resulting yeast extracts were used for immunoblot analysis using anti-Myc
669 (MBL, Woburn, MA, United States) for bait proteins and anti-HA (Roche, Switzerland) for prey proteins.
670 The primers used to generate the constructs are listed in **Suppl. Table 10**.
671
672 **Co-immunoprecipitation assay**
673 Expression constructs for two types of helper-NLRs, Pias-1 (MHD) and RGA4 (MHD) (pCambia1300S-
674 "helper-NLR":HA), were generated by PCR and cloned into the binary vector pCambia1300S
675 (http://www.cambia.org) that had been linearized by digestion with *Pst*I and *Spe*I by in-fusion cloning. The
676 resulting vectors were introduced into *A. tumefaciens* (strain GV3101). Pias-1(MHD):HA and FLAG:Pias-
677 2 or FLAG:GFP were transiently co-expressed in *N. benthamiana.* Similarly, RGA4 (MHD):HA and
678 FLAG:RGA5 or FLAG:GFP were transiently co-expressed in *N. benthamiana*. The co-expressed proteins
679 were extracted with 50 mM sodium phosphate buffer containing 10 % (v/v) glycerol and 150 mM NaCl.

680 Extracted proteins were incubated with FLAG-agarose beads in the presence of protease inhibitor, 10 mM
681 DTT, and 2% (w/v) blocking reagent (Cytiva, Tokyo, Japan) at 4°C for 1 h. The beads were washed with
682 the same buffer four times, and then bound proteins were extracted with the same buffer containing 1% (w/v)
683 SDS by boiling. Bound fractions were analyzed by immunoblotting using anti-FLAG and anti-HA antibodies.
684 The primers used to generate constructs are listed in **Suppl. Table 10**.
685
686 **Data Availability**
687 The nucleotide sequences of *Pias-1*, *Pias-2*, and *AVR-Pias* were deposited in the DNA Databank of Japan
688 (*Pias-1*: LC672059, *Pias-2*: LC672060, and *AVR-Pias*: LC672061). The DNA-seq and RNA-seq data from
689 this study are listed in **Suppl. Dataset 2** and were deposited in the DNA Databank of Japan (DDBJ) under
690 the Bioproject accession numbers PRJDB9440, PRJDB12353, PRJDB12884, PRJDB12891, and
691 PRJDB12902. All data pertaining to this study are included in the article and/or supporting information.
692
693 **Acknowledgements**
694 This study was supported by JSPS KAKENHI 15H05779, 20H00421 and 16H06279 (PAGS) to RT and
695 21K14834 to MS, the Royal Society UK-Japan International exchange grants JPJSBP120215702 and
696 IEC\R3\203081 to RT, SK and MB, the Biotechnology and Biological Sciences Research Council (UKRI-
697 BBSRC, UK, grant BB/P012574), the European Research Council (ERC BLASTOFF project, 743165), the
698 John Innes Foundation and the Gatsby Charitable Foundation. Computations were partially performed on
699 the NIG supercomputer at ROIS National Institute of Genetics. We also thank the National Institute of
700 Genetics (National Bioresource Project), Japan, and the Genebank at the National Agriculture and Food
701 Research Organization, Japan, for providing the seeds of wild rice and WRC17, respectively.
702
703 **Competing interests**
704
705 S.K. receives funding from industry on NLR biology.
706
707 **References in the main text**
708
709 Adachi H, Derevnina L, Kamoun S. NLR singletons, pairs, and networks: evolution, assembly, and
710 regulation of the intracellular immunoreceptor circuitry of plants. Curr Opin Plant Biol. 2019:121-131. doi:
711 10.1016/j.pbi.2019.04.007.
712
713 Adachi H, Kamoun S, Maqbool A. A resistosome-activated 'death switch'. Nat Plants. 2019b 5:457-458. doi:
714 10.1038/s41477-019-0425-9.
715

716 Ashikawa I, Hayashi N, Yamane H, Kanamori H, Wu J, Matsumoto T, Ono K, Yano M. Two adjacent
717 nucleotide-binding site-leucine-rich repeat class genes are required to confer *Pikm*-specific rice blast
718 resistance. Genetics. 2008 180:2267-2276. doi: 10.1534/genetics.108.095034
719
720 Bailey PC, Schudoma C, Jackson W, Baggs E, Dagdas G, Haerty W, Moscou M, Krasileva KV. Dominant
721 integration locus drives continuous diversification of plant immune receptors with exogenous domain
722 fusions. Genome Biol. 2018 19:23. doi: 10.1186/s13059-018-1392-6.
723
724 Bernoux M, Burdett H, Williams SJ, Zhang X, Chen C, Newell K, Lawrence GJ, Kobe B, Ellis JG, Anderson
725 PA, Dodds PN. Comparative Analysis of the Flax Immune Receptors L6 and L7 Suggests an Equilibrium-
726 Based Switch Activation Model. Plant Cell. 2016 28:146-159. doi: 10.1105/tpc.15.00303.
727
728 Bi G, Su M, Li N, Liang Y, Dang S, Xu J, Hu M, Wang J, Zou M, Deng Y, Li Q, Huang S, Li J, Chai J, He
729 K, Chen YH, Zhou JM. The ZAR1 resistosome is a calcium-permeable channel triggering plant immune
730 signaling. Cell. 2021 S0092-867400600-0. doi: 10.1016/j.cell.2021.05.003.
731
732 Białas A, Langner T, Harant A, et al. Two NLR immune receptors acquired high-affinity binding to a fungal
733 effector through convergent evolution of their integrated domain. Elife. 2021 10:e66961.
734 doi:10.7554/eLife.66961.
735
736 Cesari S, Thilliez G, Ribot C, Chalvon V, Michel C, Jauneau A, Rivas S, Alaux L, Kanzaki H, Okuyama Y,
737 Morel JB, Fournier E, Tharreau D, Terauchi R, Kroj T. The rice resistance protein pair RGA4/RGA5
738 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell. 2013
739 25:1463-1481. doi: 10.1105/tpc.112.107201.
740
741 Césari S, Kanzaki H, Fujiwara T, Bernoux M, Chalvon V, Kawano Y, Shimamoto K, Dodds P, Terauchi R,
742 Kroj T. The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease
743 resistance. EMBO J. 2014 33:1941-1959. doi: 10.15252/embj.201487923.
744
745 Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds
746 DA, Chen H, Frazer KA, Huson DH, Schölkopf B, Nordborg M, Rätsch G, Ecker JR, Weigel D. Common
747 sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science. 2007 317:338-342.
748 doi: 10.1126/science.1138632.
749
750 Edwards SV, Hedrick PW. Evolution and ecology of MHC molecules: from genomics to sexual selection.
751 Trends Ecol Evol. 1998 13:305-311. doi: 10.1016/s0169-5347(98)01416-5.
752

753 Fujisaki K, Abe Y, Ito A, Saitoh H, Yoshida K, Kanzaki H, Kanzaki E, Utsushi H, Yamashita T, Kamoun
754 S, Terauchi R. Rice Exo70 interacts with a fungal effector, AVR-Pii, and is required for AVR-Pii-triggered
755 immunity. Plant J. 2015 83:875-887. doi: 10.1111/tpj.12934.
756
757 Fujisaki K, Abe Y, KanzakiE, Ito K, Utsushi H, Saitoh H, Białas A, Banfield MJ, Kamoun S, Terauchi R.
758 An unconventional NOI/RIN4 domain of a rice NLR protein binds host EXO70 protein to confer fungal
759 immunity. bioRxiv. 2017 doi:10.1101/239400.
760
761 Hogenhout SA, Van der Hoorn RA, Terauchi R, Kamoun S. Emerging concepts in effector biology of plant-
762 associated organisms. Mol Plant Microbe Interact. 2009 22:115-122. doi: 10.1094/MPMI-22-2-0115.
763
764 Horsefield S, Burdett H, Zhang X, Manik MK, Shi Y, Chen J, Qi T, Gilley J, Lai JS, Rank MX, Casey LW,
765 Gu W, Ericsson DJ, Foley G, Hughes RO, Bosanac T, von Itzstein M, Rathjen JP, Nanson JD, Boden M,
766 Dry IB, Williams SJ, Staskawicz BJ, Coleman MP, Ve T, Dodds PN, Kobe B. NAD^+ cleavage activity by
767 animal and plant TIR domains in cell death pathways. Science. 2019 365:793-799. doi:
768 10.1126/science.aax1911.
769
770 Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B. Direct interaction of resistance gene and avirulence
771 gene products confers rice blast resistance. EMBO J. 2000 19:4004-4014. doi: 10.1093/emboj/19.15.4004
772
773 Jones JD, Dangl JL. The plant immune system. Nature. 2006 444:323-329. doi: 10.1038/nature05286
774
775 Kanzaki H, Yoshida K, Saitoh H, Fujisaki K, Hirabuchi A, Alaux L, Fournier E, Tharreau D, Terauchi R.
776 Arms race co-evolution of *Magnaporthe oryzae AVR-Pik* and rice *Pik* genes driven by their physical
777 interactions. Plant J. 2012 72:894-907. doi: 10.1111/j.1365-313X.2012.05110.x.
778
779 Kojima Y, Ebana K, Fukuoka S, Nagamine T, Kawase M. Development of an RFLP-based rice diversity
780 research set of germplasm. Breeding Sci. 2005 55:431-440. doi: 10.1270/jsbbs.55.431
781
782 Kourelis J, van der Hoorn RAL. Defended to the Nines: 25 Years of Resistance Gene Cloning Identifies
783 Nine Mechanisms for R Protein Function. Plant Cell. 2018 30:285-299. doi: 10.1105/tpc.17.00579.
784
785 Krasileva KV, Dahlbeck D, Staskawicz BJ. Activation of an *Arabidopsis* resistance protein is specified by
786 the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. Plant Cell.
787 2010 22:2444-2458. doi: 10.1105/tpc.110.075358.
788

789 Kroj T, Chanclud E, Michel-Romiti C, Grand X, Morel JB. Integration of decoy domains derived from
790 protein targets of pathogen effectors into plant immune receptors is widespread. New Phytol. 2016 210:618-
791 626. doi: 10.1111/nph.13869.
792
793 Le Roux C, Huet G, Jauneau A, Camborde L, Trémousaygue D, Kraut A, Zhou B, Levaillant M, Adachi H,
794 Yoshioka H, Raffaele S, Berthomé R, Couté Y, Parker JE, Deslandes L. A receptor pair with an integrated
795 decoy converts pathogen disabling of transcription factors to immunity. Cell. 2015 161:1074-1088. doi:
796 10.1016/j.cell.2015.04.025.
797
798 Li P, Quan X, Jia G, Xiao J, Cloutier S, You FM. RGAugury: a pipeline for genome-wide prediction of
799 resistance gene analogs (RGAs) in plants. BMC Genomics. 2016 17:852. doi: 10.1186/s12864-016-3197-x.
800
801 Liu J, Wang X, Mitchell T, Hu Y, Liu X, Dai L, Wang GL. Recent progress and understanding of the
802 molecular mechanisms of the rice-*Magnaporthe oryzae* interaction. Mol Plant Pathol. 2010 11:419-427. doi:
803 10.1111/j.1364-3703.2009.00607.x.
804
805 Lv F, Wang H, Wang X, et al. GhCFE1A, a dynamic linker between the ER network and actin cytoskeleton,
806 plays an important role in cotton fibre cell initiation and elongation. J Exp Bot. 2015;66(7):1877-1889.
807 doi:10.1093/jxb/eru530
808
809 Ma S, Lapin D, Liu L, Sun Y, Song W, Zhang X, Logemann E, Yu D, Wang J, Jirschitzka J, Han Z, Schulze-
810 Lefert P, Parker JE, Chai J. Direct pathogen-induced assembly of an NLR immune receptor complex to form
811 a holoenzyme. Science. 2020 370:eabe3069. doi: 10.1126/science.abe3069.
812
813 Maidment JHR, Franceschetti M, Maqbool A, Saitoh H, Jantasuriyarat C, Kamoun S, Terauchi R, Banfield
814 MJ. Multiple variants of the fungal effector AVR-Pik bind the HMA domain of the rice protein OsHIPP19,
815 providing foundation to engineer plant defense. J Biol Chem 2021 196:100371. doi: 10.1016/j.jbc.2021.
816 100371.
817
818 Maqbool A, Saitoh H, Franceschetti M, Stevenson CE, Uemura A, Kanzaki H, Kamoun S, Terauchi R,
819 Banfield MJ. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune
820 receptor. Elife. 2015 4:e08709. doi: 10.7554/eLife.08709.
821
822 Martin R, Qi T, Zhang H, Liu F, King M, Toth C, Nogales E, Staskawicz BJ. Structure of the activated Roq1
823 resistosome directly recognizing the pathogen effector XopQ. bioRxiv. 2020 doi:
824 https://doi.org/10.1101/2020.08.13.246413.
825

826 Meng X, Xiao G, Telebanco-Yanoria MJ, Siazon PM, Padilla J, Opulencia R, Bigirimana J, Habarugira G,
827 Wu J, Li M, Wang B, Lu GD, Zhou B. The broad-spectrum rice blast resistance (*R*) gene *Pita2* encodes a
828 novel R protein unique from *Pita*. Rice (N Y). 2020 13:19. doi: 10.1186/s12284-020-00377-5.
829
830 Oikawa K, Fujisaki K, Shimizu M, Takeda T, Saitoh H, Hirabuchi A, Hiraka Y, Białas A, Langner T, Kellner
831 R, Bozkurt T, Cesari S, Kroj T, Maidment J, Banfield M, Kamoun S, Terauchi R. The blast pathogen effector
832 AVR-Pik binds and stabilizes rice heavy metal-associated (HMA) proteins to co-opt their function in
833 immunity. bioRxiv. 2020 doi:10.1101/2020.12.01.406389.
834
835 Okuyama Y, Kanzaki H, Abe A, Yoshida K, Tamiru M, Saitoh H, Fujibe T, Matsumura H, Shenton M,
836 Galam DC, Undan J, Ito A, Sone T, Terauchi R. A multifaceted genomics approach allows the isolation of
837 the rice *Pia*-blast resistance gene consisting of two adjacent NBS-LRR protein genes. Plant J. 2011 66:467-
838 79. doi: 10.1111/j.1365-313X.2011.04502.x.
839
840 Orbach MJ, Farrall L, Sweigard JA, Chumley FG, Valent B. A telomeric avirulence gene determines efficacy
841 for the rice blast resistance gene Pi-ta. Plant Cell. 2000 12:2019-32. doi: 10.1105/tpc.12.11.2019.
842
843 Pan Q, Wendel J, Fluhr R. Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and
844 cereal genomes. J Mol Evol. 2000 50:203-13. doi: 10.1007/s002399910023.
845
846 Park CH, Shirsekar G, Bellizzi M, Chen S, Songkumarn P, Xie X, Shi X, Ning Y, Zhou B, Suttiviriya P,
847 Wang M, Umemura K, Wang GL. The E3 Ligase APIP10 Connects the Effector AvrPiz-t to the NLR
848 Receptor Piz-t in Rice. PLoS Pathog. 2016 12:e1005529. doi: 10.1371/journal.ppat.1005529.
849
850 Pennisi E. Armed and dangerous. Science. 2010 327:804-805. doi: 10.1126/science.327.5967.804.
851
852 Piertney SB, Oliver MK. The evolutionary ecology of the major histocompatibility complex. Heredity
853 (Edinb). 2006 96:7-21. doi: 10.1038/sj.hdy.6800724.
854
855 Sarris PF, Cevik V, Dagdas G, Jones JD, Krasileva KV. Comparative analysis of plant immune receptor
856 architectures uncovers host proteins likely targeted by pathogens. BMC Biol. 2016 14:8. doi:
857 10.1186/s12915-016-0228-7.
858
859 Sharif H, Wang L, Wang WL, Magupalli VG, Andreeva L, Qiao Q, Hauenstein AV, Wu Z, Núñez G, Mao
860 Y, Wu H. Structural mechanism for NEK7-licensed activation of NLRP3 inflammasome. Nature. 2019
861 570:338-343. doi: 10.1038/s41586-019-1295-z.
862

863 Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics.
864 1989 123:585-595. doi: 10.1093/genetics/123.3.585.
865
866 Takagi H, Uemura A, Yaegashi H, Tamiru M, Abe A, Mitsuoka C, Utsushi H, Natsume S, Kanzaki H,
867 Matsumura H, Saitoh H, Yoshida K, Cano LM, Kamoun S, Terauchi R. MutMap-Gap: whole-genome
868 resequencing of mutant F2 progeny bulk combined with de novo assembly of gap regions identifies the rice
869 blast resistance gene *Pii*. New Phytol. 2013 200:276-283. doi: 10.1111/nph.12369.
870
871 Takahata N, Nei M. Allelic genealogy under overdominant and frequency-dependent selection and
872 polymorphism of major histocompatibility complex loci. Genetics. 1990 124:967-78.
873
874 Tenthorey JL, Haloupek N, López-Blanco JR, Grob P, Adamson E, Hartenian E, Lind NA, Bourgeois NM,
875 Chacón P, Nogales E, Vance RE. The structural basis of flagellin detection by NAIP5: A strategy to limit
876 pathogen immune evasion. Science. 2017 358:888-893. doi: 10.1126/science.aao1140.
877
878 Terauchi R, Yoshida K. Towards population genomics of effector-effector target interactions. New Phytol.
879 2010 187:929-939. doi: 10.1111/j.1469-8137.2010.03408.x.
880
881 Truman W, Bennett MH, Kubigsteltig I, Turnbull C, Grant M. Arabidopsis systemic immunity uses
882 conserved defense signaling pathways and is mediated by jasmonates. Proc Natl Acad Sci U S A. 2007
883 104:1075-1080. doi: 10.1073/pnas.0605423104.
884
885 Van de Weyer AL, Monteiro F, Furzer OJ, Nishimura MT, Cevik V, Witek K, Jones JDG, Dangl JL, Weigel
886 D, Bemm F. A Species-Wide Inventory of NLR Genes and Alleles in *Arabidopsis thaliana*. Cell. 2019
887 178:1260-1272.e14. doi: 10.1016/j.cell.2019.07.038.
888
889 Wan L, Essuman K, Anderson RG, Sasaki Y, Monteiro F, Chung EH, Osborne Nishimura E, DiAntonio A,
890 Milbrandt J, Dangl JL, Nishimura MT. TIR domains of plant immune receptors are NAD^+-cleaving enzymes
891 that promote cell death. Science. 2019 365:799-803. doi: 10.1126/science.aax1771.
892
893 Wang J, Wang J, Hu M, Wu S, Qi J, Wang G, Han Z, Qi Y, Gao N, Wang HW, Zhou JM, Chai J. Ligand-
894 triggered allosteric ADP release primes a plant NLR complex. Science. 2019a 364:eaav5868. doi:
895 10.1126/science.aav5868.
896
897 Wang J, Hu M, Wang J, Qi J, Han Z, Wang G, Qi Y, Wang HW, Zhou JM, Chai J. Reconstitution and
898 structure of a plant NLR resistosome conferring immunity. Science. 2019b 364:eaav5870. doi:
899 10.1126/science.aav5870.
900

901 Wang R, Ning Y, Shi X, He F, Zhang C, Fan J, Jiang N, Zhang Y, Zhang T, Hu Y, Bellizzi M, Wang GL.
902 Immunity to Rice Blast Disease by Suppression of Effector-Triggered Necrosis. Curr Biol. 2016 26:2399-
903 2411. doi: 10.1016/j.cub.2016.06.072.
904
905 Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, Ve T, Ma Y, Saucet SB, Ericsson DJ,
906 Casey LW, Lonhienne T, Winzor DJ, Zhang X, Coerdt A, Parker JE, Dodds PN, Kobe B, Jones JD. Structural
907 basis for assembly and function of a heterodimeric plant immune receptor. Science. 2014 344:299-303. doi:
908 10.1126/science.1247357.
909
910 Woolhouse ME, Webster JP, Domingo E, Charlesworth B, Levin BR. Biological and biomedical
911 implications of the co-evolution of pathogens and their hosts. Nat Genet. 2002 32:569-577. doi:
912 10.1038/ng1202-569.
913
914 Wu CH, Abd-El-Haliem A, Bozkurt TO, Belhaj K, Terauchi R, Vossen JH, Kamoun S. NLR network
915 mediates immunity to diverse plant pathogens. Proc Natl Acad Sci U S A. 2017 114:8113-8118. doi:
916 10.1073/pnas.1702041114.
917
918 Wu Y, Xiao N, Yu L, Pan C, Li Y, Zhang X, Liu G, Dai Z, Pan X, Li A. Combination Patterns of Major R
919 Genes Determine the Level of Resistance to the *M. oryzae* in Rice (*Oryza* sativa L.). PLoS One. 2015
920 10:e0126130. doi: 10.1371/journal.pone.0126130.
921
922 Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Yoshida K, Tosa Y, Chuma I, Takano Y, Win
923 J, Kamoun S, Terauchi R. Association genetics reveals three novel avirulence genes from the rice blast
924 fungal pathogen *Magnaporthe oryzae*. Plant Cell. 2009 21:1573-1591. doi: 10.1105/tpc.109.066324.
925
926 Zdrzałek R, Kamoun S, Terauchi R, Saitoh H, Banfield MJ. The rice NLR pair Pikp-1/Pikp-2 initiates cell
927 death through receptor cooperation rather than negative regulation. PLoS One. 2020 15:e0238616. doi:
928 10.1371/journal.pone.0238616.
929
930 Zhao H, Wang X, Jia Y, Minkenberg B, Wheatley M, Fan J, Jia MH, Famoso A, Edwards JD, Wamishe Y,
931 Valent B, Wang GL, Yang Y. The rice blast resistance gene *Ptr* encodes an atypical protein required for
932 broad-spectrum disease resistance. Nat Commun. 2018 9:2039. doi: 10.1038/s41467-018-04369-4.
933
934 Zhang L, Chen S, Ruan J, Wu J, Tong AB, Yin Q, Li Y, David L, Lu A, Wang WL, Marks C, Ouyang Q,
935 Zhang X, Mao Y, Wu H. Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals
936 nucleated polymerization. Science. 2015 350:404-9. doi: 10.1126/science.aac5789.
937

Zhang Y, Zhang F, Huang X. Characterization of an *Arabidopsis thaliana* DUF761-containingprotein with a potential role in development and defense responses. Theor. Exp. Plant Physiol. 2019 31:303-316.

References in the MM and the supplemental data (Figure and Table)

Arazoe T, Miyoshi K, Yamato T, Ogawa T, Ohsato S, Arie T, Kuwata S. Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015 112:2543-2549. doi: 10.1002/bit.25662.

Bennetzen JL, Schmutz J, Wang H, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012 30:555-561. doi:10.1038/nbt.2196

Du H, Yu Y, Ma Y, et al. Sequencing and *de novo* assembly of a near complete indica rice genome. Nat Commun. 2017 8:15324. doi:10.1038/ncomms15324

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol. 2018 35:518-522. doi: 10.1093/molbev/msx281.

Hyon GS, Naga NTT, Chuma I, Inoue Y, Asano H, Murata N, Tosa Y. Characterization of interactions between barley and various host-specific subgroups of *Magnaporthe oryzae* and *M. grisea*. J. Gen. Plant Pathol. 2012 78: 237-246. doi: https://doi.org/10.1007/s10327-012-0386-6.

International Rice Genome Sequencing Project. The map-based sequence of the rice genome. Nature. 2005 436:793-800. doi:10.1038/nature03895

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017 14:587-589. doi: 10.1038/nmeth.4285.

Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015 12:357-360. doi:10.1038/nmeth.3317

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018 35:547. doi: 10.1093/molbev/msy096.

Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010 26:589-595. doi:10.1093/bioinformatics/btp698

Löytynoja A, Goldman N. webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. BMC Bioinformatics. 2010 11:579. doi: 10.1186/1471-2105-11-579

Martin EC, Sukarta OCA, Spiridon L, Grigore LG, Constantinescu V, Tacutu R, Goverse A, Petrescu AJ. LRRpredictor-A New LRR Motif Detection Method for Irregular Motifs of Plant NLR Proteins Using an Ensemble of Classifiers. Genes (Basel). 2020 11:286. doi: 10.3390/genes11030286.

Mascher M, Gundlach H, Himmelbach A, et al. A chromosome conformation capture ordered sequence of the barley genome. Nature. 2017;544:427-433. doi:10.1038/nature22043

Mikami M, Toki S, Endo M. Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. Plant Mol Biol. 2015 88:561-72. doi: 10.1007/s11103-015-0342-x.

Miki D, Shimamoto K. Simple RNAi vectors for stable and transient suppression of gene function in rice. Plant Cell Physiol. 2004 45:490-495. doi: 10.1093/pcp/pch048.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020 37:1530-1534. doi: 10.1093/molbev/msaa015.

Okuyama Y, Kanzaki H, Abe A, Yoshida K, Tamiru M, Saitoh H, Fujibe T, Matsumura H, Shenton M, Galam DC, Undan J, Ito A, Sone T, Terauchi R. A multifaceted genomics approach allows the isolation of the rice *Pia*-blast resistance gene consisting of two adjacent NBS-LRR protein genes. Plant J. 2011 66:467-479. doi: 10.1111/j.1365-313X.2011.04502.x.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015 33:290-295. doi:10.1038/nbt.3122

Ranwez V, Douzery EJ, Cambon C, Chantret N, Delsuc F. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Mol Biol Evol. 2018 35:2582-2584. doi: 10.1093/molbev/msy159

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011 29:24-26. doi: 10.1038/nbt.1754.

1013 Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods.
1014 2012 9:671-675. doi: 10.1038/nmeth.2089.
1015
1016 Shimizu M, Nakano Y, Hirabuchi A, Yoshino K, Kobayashi M, Yamamoto K, Terauchi R, Saitoh H. RNA-
1017 Seq of in planta-expressed *Magnaporthe oryzae* genes identifies MoSVP as a highly expressed gene required
1018 for pathogenicity at the initial stage of infection. Mol Plant Pathol. 2019 20:1682-1695. doi:
1019 10.1111/mpp.12869.
1020
1021 Sweigard J, Chumley F, Carroll A, Farrall L, Valent B. A series of vectors for fungal transformation. Fungal
1022 Genet. Newsl. 1997 44:52-55. doi: https://doi.org/10.4148/1941-4765.1287.
1023
1024 Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics.
1025 1989 123:585–595. doi: 10.1093/genetics/123.3.585.
1026
1027 Wu Z, Fang D, Yang R, et al. *De novo* genome assembly of *Oryza granulata* reveals rapid genome expansion
1028 and adaptive evolution. Commun Biol. 2018 1:84. doi:10.1038/s42003-018-0089-4.
1029
1030 Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci.
1031 1997 13:555-556. doi: 10.1093/bioinformatics/13.5.555.
1032
1033 Yang Z, Nielsen R. 2000. estimating synonymous and nonsynonymous substitution rates under realistic
1034 evolutionary models. Mol Biol Evol 17:32-43. doi:10.1093/oxfordjournals.molbev.a026236
1035
1036 Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Yoshida K, Tosa Y, Chuma I, Takano Y, Win
1037 J, Kamoun S, Terauchi R. Association genetics reveals three novel avirulence genes from the rice blast
1038 fungal pathogen *Magnaporthe oryzae*. Plant Cell. 2009 21:1573-1591. doi: 10.1105/tpc.109.066324.
1039
1040 Zhao Q, Feng Q, Lu H, et al. Pan-genome analysis highlights the extent of genomic variation in cultivated
1041 and wild rice. Nat Genet. 2018 50:278-284. doi:10.1038/s41588-018-0041-z
1042
1043 Zimin AV, Puiu D, Luo MC, et al. Hybrid assembly of the large and highly repetitive genome of Aegilops
1044 tauschii, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. Genome Res. 2017 27:787-
1045 792. doi:10.1101/gr.213405.116
1046
1047 Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. The MaSuRCA genome assembler.
1048 Bioinformatics. 2013 29:2669-2677. doi:10.1093/bioinformatics/btt476
1049
1050

1051 **Figure legends**

1052

1053 **Figure 1. The *Pias* gene of rice line WRC17 (cultivar Keiboba) encodes a CC-NLR protein and is**
1054 **allelic to *Pia*.**

1055 (A) Segregation of the resistance and susceptibility traits among the 58 RILs derived from a cross between
1056 WRC17 (cultivar Keiboba) and Hitomebore. Disease symptoms of WRC17, Hitomebore, and six RILs
1057 showing a susceptible phenotype after punch inoculation of *Magnaporthe oryzae* isolate 2012-1 (leaf
1058 photographs) and the frequency distribution of disease lesion area of the 58 RILs (bar graphs). (B) Linkage
1059 maps of candidate NLR genes at the *PiW17-1* and *PiW17-2* loci. (C) Both *CNL-04* and *CNL-05* are required
1060 for *PiW17-1*-mediated resistance against *M. oryzae* 2012-1. HW-RIL7 contains only *PiW17-1* and is
1061 resistant to *M. oryzae* 2012-1. Knockout of *CNL-04* (*cnl-04*) and *CNL-05* (*cnl-05*) in HW-RIL7 rendered
1062 plants susceptible to 2012-1. (D) Gene structures of *Pia* consisting of *RGA4* and *RGA5* and *Pias* consisting
1063 of *Pias-1* and *Pias-2*. The positions of protein domains (CC, NB-ARC, LRR, HMA, and DUF761) encoded
1064 by the NLRs are indicated. (E) The *M. oryzae* 2012-1 *AVR-Pias* knockout mutant became virulent to HW-
1065 RIL7. (F) Amino acid sequence of G9532 protein (AVR-Pias). The secretion signal is indicated by red letters
1066 and the Toxin18-like motif by blue letters. The Toxin18-like motif was annotated by Pfam
1067 (https://pfam.xfam.org/).

1068

1069 **Figure 2. Recurrent integration of extraneous domains in Pias/Pia sensor NLRs.**

1070 (A) A simplified scheme of the structures of the Pias/Pia NLR pairs. Pias-1/RGA4 helper NLRs are shown
1071 in green, and Pias-2/RGA5 sensor NLRs are shown in white. The conserved LRR-ID intervening motif (LII)
1072 is indicated by black ellipses. Integrated domains are shown by different-colored hexagons. (B) A sequence
1073 logo showing conserved amino acids of the LRR-ID Intervening (LII) motif. The red lines indicate the LRR
1074 in LII motif. (C) Distribution of ID motifs among *Oryza* species. The pie charts show the frequencies of
1075 different ID motifs in a given species. The colors correspond to the ID colors in (A). The numbers below the
1076 pie charts indicate the sample numbers. A cladogram showing the phylogenetic relationships of 11 *Oryza*
1077 species and four other Poaceae species (*Setaria italica*, *Panicum hallii*, *Hordeum vulgare*, and *Aegilops*
1078 *tauschii*) based on Time Tree, the time scale of life web database (http://www.timetree.org/). The numbers
1079 on the branches indicate the estimated time of the splitting of lineages (MYA: million years ago). (D) DNA
1080 sequence similarity between *O. punctata Pias-2/RGA5 sensor NLR* and the downstream sequence of *O.*
1081 *sativa* (Nipponbare) *Pias-2/RGA5.* LII: LRR-ID intervening motif. (E) Dot-plot analysis of the *O. punctata*
1082 (W1582) *Pias-2/RGA5 sensor NLR* and *O. sativa* (Nipponbare) *Pias-2/RGA5 NLR* downstream sequences
1083 using the Dotmatcher tool (http://emboss.bioinformatics.nl/cgi-bin/emboss/dotmatcher/). (F) Possible
1084 evolutionary process of ID replacement that might have occurred between the *O. punctata* and the *O. sativa*
1085 *Pias-2*/*RGA5* lineages. We still do not know the mode of interaction between the AVR-Pias effector and
1086 DUF761-containing protein, so it is indicated by "?".

1087

1088 **Figure 3. Contrasting evolutionary patterns of the helper and sensor NLRs of the *Pias/Pia* locus.**

1089 (A) Phylogenetic tree of the *Pias-1/RGA4* helper NLR gene (left) and *Pias-2/RGA5* sensor NLR (right) gene
1090 based on the full-length amino acid sequence of Pias-1/RGA4 and the sequence in the region CC to LII for
1091 Pias-2/RGA5. *Pias-2/RGA5* sensor NLRs form two major clades (C1 and C2). The numbers indicate
1092 bootstrap values. (B) Nucleotide diversity (π) of the CC, NBS and LRR(-LII) domains of *Pias-1/RGA4*
1093 helper NLR gene and *Pias-2/RGA5* sensor NLR gene in 22 *Oryza* samples. (C) Tajima's *D* of the CC, NBS
1094 and LRR(-LII) domains of *Pias-1/RGA4* helper NLR gene and *Pias-2/RGA5* sensor NLR gene in 22 *Oryza*
1095 samples. (D) Pairwise d_N and d_S values of CC-NBS-LRR(-LII), CC, NBS and LRR domains of the *Pias-*
1096 *1/RGA4* helper and *Pias-2/RGA5* sensor NLR gene in 22 *Oryza* samples.
1097
1098 **Figure 4. The NLR helper Pias-1 is functionally conserved.**
1099 (A) Representative images of *N. benthamiana* leaves after agroinfiltration with Pias-1:HA, Pias-
1100 1:HA/FLAG:RGA5, FLAG:RGA5/AVR-Pia, and Pias-1:HA/FLAG:RGA5/AVR-Pia (left) and RGA4-
1101 Ogr:HA derived from *O. granulata*, RGA4-Ogr:HA/FLAG:RGA5, FLAG:RGA5/AVR-Pia, and RGA4-
1102 Ogr:HA/FLAG:RGA5/AVR-Pia (right). Autofluorescence under UV light is shown. (B) Pias-1 cooperates
1103 with RGA5 to recognize AVR-Pia and induces resistance in rice. The rice line HW-RIL7 with *Pias* (*Pias-1*
1104 and *Pias-2*) recognizes the Ao-92-06-2 strain with *AVR-Pias* (Ao92-06-2+*pex22p*:*AVR-Pias*) and induces
1105 resistance. However, HW-RIL7 cannot recognize the Ao02-06-2 strain with *AVR-Pia* (Ao92-06-
1106 2+*pex22p*:*AVR-Pia*). Two lines (Line #1 and #2) contain the *35S-RGA5* transgene in the HW-RIL7
1107 background. F1-A is a progeny derived from a cross between Sasanishiki with *Pia* (*RGA4* and *RGA5*) and
1108 HW-RIL7. F1-B is a progeny derived from a cross between a Sasanishiki mutant (Sas1493) with *pia* (*rga4*
1109 and *RGA5*) and HW-RIL7.
1110
1111 **Legends to Supplementary Figures**
1112
1113 **Suppl. Fig. 1. The RaIDeN method used to identify candidate genes responsible for the resistance of**
1114 **rice line WRC17 to *M. oryzae* 2012-1.**
1115 (A) Segregation of the resistance and susceptibility traits among the 58 RILs derived from a cross between
1116 WRC17 (cultivar Keiboba) and Hitomebore. (B) The whole-genome sequence of the resistant parent WRC17
1117 was obtained by Illumina sequencing. RNA-seq of WRC17 leaves inoculated with *M. oryzae* was also
1118 performed. Next, short reads of whole-genome sequences of the susceptible parent Hitomebore and the six
1119 susceptible RILs were obtained by Illumina sequencing. (C) *De novo* assembly of R-parental line WRC17
1120 was performed using *DISCOVAR* software. (D) Expressed genes of WRC17 were identified by mapping
1121 RNA-seq reads to the WRC17 reference genome. (E) Mapping of short reads of the susceptible parent
1122 (Hitomebore) and susceptible RILs to the predicted genes identified DNA polymorphisms shared by the
1123 susceptible lines using the RaIDeN pipeline (the script of the RaIDeN pipeline and details about the pipeline
1124 are available at https://github.com/YuSugihara/RaIDeN).
1125

1126 **Suppl. Fig. 2. Mapping of *NLR*s in the *PiW17-1* and *PiW17-2* loci that confer resistance against *M.***
1127 ***oryzae* 2012-1 to WRC17.** (A) Association between resistance/susceptible phenotypes and genotypes
1128 among the RILs (the primers used for genotyping are shown in Suppl. Table 9). *CNL-04* and *CNL-05*, *CNL-*
1129 *06* and *CNL-07*, and *CNL-08*, *NL-2*, and *CNL-09* are located on the same contigs. (B) Fine maps of candidate
1130 NLR genes in *PiW17-1* and *PiW17-2*. Blue and yellow indicate WRC17- and Hitomebore-type genotypes,
1131 respectively. The 10 NLRs shown in the red rectangle were considered to be the candidates of *Pi-W17-1.*
1132
1133 **Suppl. Fig. 2 (continued). Results of RNAi-mediated silencing of candidate NLR genes.**
1134 (C) RNAi-mediated silencing of eight candidate genes encoding proteins over 900 amino acids long was
1135 performed. Gene silencing of *CNL-04* and *CNL-05* made plants susceptible to *M. oryzae* 2012-1. (D) Gene
1136 expression levels of *CNL-04* and *CNL-05* in plants transformed with the RNAi constructs. Control (Con) is
1137 WRC17 transformed with empty vector.
1138
1139 **Suppl. Fig. 3**. **CRISPR/Cas9-mediated knockout of *CNL-04* abolishes *PiW17-1*-mediated resistance**
1140 **against *M. oryzae* isolate 2012-1.**
1141 (A) The position of guide DNA targeted to *CNL-04*. The PAM is marked with blue letters, and the sgRNA
1142 sequence is marked with red letters. (B) Sequences of the sgRNA positions in the *CNL-04* knockout lines.
1143 (C) Resistance/susceptible phenotypes of the selfed progeny of the *cnl-04_1.14.2* line heterozygous for wild-
1144 type (*a*) and mutated (*b*: 4-bp deletion) alleles. The *b/b* homozygous plants became susceptible.
1145
1146 **Suppl. Fig. 3 (continued). CRISPR/Cas9-mediated knockout of *CNL-05* abolishes *PiW17-1* mediated**
1147 **resistance against *M. oryzae* isolate 2012-1.**
1148 (D) The positions of guide DNAs targeted to *CNL-05*. The PAM is marked with blue letters, and the sgRNA
1149 sequence is marked with red letters. (E) Sequences of the sgRNA positions in the *CNL-05* knockout lines.
1150 (F) Resistance/susceptible phenotypes of the selfed progeny of *cnl-05_3.7.2* line heterozygous for wild-type
1151 (*a*) and mutated (*b*: 1-bp insertion) alleles. The *b/b* homozygous plants became susceptible.
1152
1153 **Suppl. Fig. 4. Comparison of the gene structures of *Pia* (*RGA4* and *RGA5*) and *Pias* (*Pias-1* and *Pias-***
1154 ***2*).**
1155 DNA sequence similarities obtained by ClustalW are shown.
1156
1157 **Suppl. Fig. 5. Gene structures and predicted amino acid sequences of *Pias-1* and *Pias-2*.**
1158 (A) Gene structures of *Pias-1* and *Pias-2*. (B) RNA-seq read alignment across the genome sequence of the
1159 *Pias* region as visualized by IGV (Robinson et al. 2011).
1160
1161 **Suppl. Fig. 5 (continued).**
1162 (C) Amino acid sequences and putative domains of Pias-1 and Pias-2. The RX-CC and NB-ARC domains
1163 of Pias-1 and Pias-2 were annotated by a CD-search in NCBI

1164 (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), and the LRR was predicted using LRR predictor
1165 (Martin et al 2020).
1166
1167 **Suppl. Fig. 6. Isolation of *AVR-Pias* from *M. oryzae* isolate 2012-1**.
1168 A simplified scheme showing the procedure used to isolate *AVR-Pias*. First, the genome of isolate 2012-1
1169 was sequenced on the Illumina platform, and the short reads were used for *de novo* assembly with
1170 *DISCOVAR*, resulting in the 2012-1 genome reference (**Suppl. Table 4, 5**). Next, the isolate 2012-1 was
1171 used to inoculate barley (*Hordeum vulgare*) cultivar 'Nigrate', which is highly susceptible to *M. oryzae*
1172 (Hyon et al. 2012), and the infected leaves were subjected to RNA-seq (Shimizu et al. 2019), resulting in the
1173 identification of 10,991 genes in 2012-1 that were expressed during host infection. Next, eight *M. oryzae*
1174 isolates (r1–r8) were selected and used for inoculation of rice line HW-RIL7 with *Pias* but without *Pi-W17-*
1175 *2*. The eight isolates were compatible with HW-RIL7 (result shown in the bottom), suggesting that they lack
1176 *AVR-Pias*. The eight isolates were then subjected to genome resequencing on the Illumina platform, and the
1177 RaIDeN method was used to identify only presence/absence polymorphisms (**Suppl. Table 4**). The short
1178 reads of eight isolates were aligned to the expressed genes of the 2012-1 isolate, resulting in the identification
1179 of 87 expressed genes that were specific to isolate 2012-1. These transcripts were further filtered based on
1180 four criteria: (1) genes showing a higher level of expression 24 h after inoculation (FPKM > 200) and genes
1181 encoding (2) a putative secreted protein, (3) a non-transmembrane protein, and (4) a small protein (<150
1182 amino acids). This analysis identified three transcripts (*G9141*, *G9435*, and *G9532*) as the candidates of
1183 *AVR-Pias*.
1184
1185 **Suppl. Fig. 7. The expressed gene *G9532* of *M. oryzae* isolate 2012-1 is *AVR-Pias*.**
1186 (A) Results of inoculation of rice line HW-RIL7 with *Pias* or Hitomebore without *Pias* with *M. oryzae*
1187 isolate Ao92-06-2 wild type (WT) or Ao92-06-2 with the *G9141*, *G9435*, or *G9532* transgene, all driven by
1188 the *pex22* promoter. When Ao92-06-2 contained *G9532*, the interaction became incompatible, indicating
1189 that *G9532* is *AVR-Pias*. (B) Introduction of *pex22p*:*G9532* into *M. oryzae* isolate Ao92-06-2 confers
1190 avirulence to the pathogen against HW-RIL7 with *Pias-1* and *Pias-2*. Knockout of the host gene *Pias-1* or
1191 *Pias-2* compromises the avirulence caused by *G9532*.
1192
1193 **Suppl. Fig. 8. Deletion of *AVR-Pias* from *M. oryzae* isolate 2012-1 causes a loss of avirulence against**
1194 **rice line HW-RIL7 with *Pias*.**
1195 (A) Schematic overview of the target gene replacement (TGR) strategy at the *AVR-Pias* locus using RNA-
1196 guided nuclease. The PAM is marked with blue letters, and the sgRNA sequence is marked with red letters.
1197 The primers used for PCR analysis are indicated by horizontal arrows. (B) Inoculation assay of rice line
1198 HW-RIL7 with *Pias* wild type (WT) and TGR transformants. (C) PCR analysis of TGR events at the *AVR-*
1199 *Pias* locus. Upper and lower images show PCR results using *AVR-Pias*- and *Hygromycin* (*Hyg*)-specific
1200 primers, respectively. pCB1636: Replacement vector containing the hygromycin resistance gene and
1201 genomic regions neighboring *AVR-Pias*. The isolates corresponding to lanes 2–5 show compatibility with

1202 rice line HW-RIL7 with *Pias*, whereas the wild-type isolate (2012-1) in lane 6 is incompatible with HW-
1203 RIL7.
1204
1205 **Suppl. Fig. 9. Prediction of ID sequences in the Pias/Pia sensor NLRs of the *Oryza* species.**
1206 (A) Overview of gene prediction methods used in this study. (B) Gene models of *Pias-1/RGA4* and *Pias-2/*
1207 *RGA5* of 12 *Oryza* samples supported by RNA-seq data were used as queries to annotate IDs in the genome
1208 assemblies of 167 *Oryza* samples using Exonerate (http://www.ebi.ac.uk/~guy/exonerate). However, 10
1209 samples of *O. glumaepatula* and six samples of *O. brachyantha* did not match known domains. Therefore,
1210 we incorporated RNA-seq data of each sample from the two species, resulting in the annotation of the
1211 Zinc_ribbon_12 (*O. glumaepatula*) and HMA (*O. brachyantha*) IDs. In the second round, we used the 12
1212 samples used in the first round of Exonerate as well as two new samples (*O. glumaepatula* W2184 and *O.*
1213 *brachyantha* W0655) as queries to infer IDs in the assembled genomes of the 167 *Oryza* samples.
1214
1215 **Suppl. Fig. 9 (continued). Prediction of ID sequences in the Pias/Pia sensor NLRs of the *Oryza* species.**
1216 (C) RNA-seq read alignment across the genome sequence of the *Pias/Pia* homologous region in the AA
1217 genome *Oryza* species as visualized by IGV (Robinson et al. 2011).
1218
1219 **Suppl. Fig. 9 (continued). Prediction of ID sequences in the Pias/Pia sensor NLRs of the *Oryza* species.**
1220 (D) RNA-seq read alignment across the genome sequence of the *Pias/Pia* homologous region in the non-
1221 AA genome *Oryza* species as visualized by IGV (Robinson et al. 2011).
1222
1223 **Suppl. Fig. 10. Positions of the LII motif (this study) and the CID motif (Bailey et al. 2018).**
1224 Amino acid sequence alignment of RGA5, Pias-2, and LOC_Os12g18360 and the positions of the LII motif
1225 (this study) and the CID motif as described by Bailey et al. (2018). The aqua and orange boxes indicate the
1226 LII and CID motif regions, respectively. The red lines indicate LRR in the LII motif.
1227
1228 **Suppl. Fig. 11. The *O. punctata* W1582 LII-ID sequence is conserved in the downstream sequence of**
1229 ***Pias* in *O. sativa* cv. Nipponbare.**
1230 Alignment of the DNA sequences of the *O. punctata* W1582 LII-ID and the W1582 LII-ID-like sequence in
1231 *O. sativa* Nipponbare. The blue boxes and green lines indicate the exon sequence of *O. punctata* W1582
1232 Pias-2 homolog and the LII region, respectively.
1233
1234 **Suppl. Fig. 12. The LII + ID sequence of the *O. punctata* (W1582) *Pias-2* homolog is conserved in the**
1235 **downstream sequence of *O. meridionalis* (W2112) Pias-2.**
1236 Dot-plot analysis of *O. punctata* (W1582) *Pias-2* homolog and the genome sequence of *O. meridionalis*
1237 (W2112). The purple line corresponds to block (a) in Figure 2. The block (b) sequence in Figure 2 is deleted
1238 in *O. meridionalis* (W2112).
1239

1240 **Supple. Fig. 13 (A) Comparison of ω, d_N, and d_S between *Pias-1/RGA4* and *Pias-2/RGA5*.** The 'a'
1241 indicates that *Pias-2/RGA5* is significantly larger than *Pias-1/RGA4* by two-sided Welch's t-test ($p < 0.0001$).
1242 The 'b' indicates that *Pias-2/RGA5* is significantly smaller than *Pias-1/RGA4* by two-sided Welch's t-test
1243 ($p < 0.0001$). **(B) Comparison between d_N, and d_S in *Pias-1/RGA4* and *Pias-2/RGA5*.**
1244
1245 **Suppl. Fig. 14 HR-like cell death after overexpression of helper NLRs of the RGA4/Pias-1 lineage in**
1246 ***N. benthamiana*.**
1247 (A) Boxplots of autofluorescence values after transient expression of RGA4/Pias-1 homologs. GUS vector
1248 was included as a control. The number of spots inoculated with *A. tumefaciens* are indicated at the top of the
1249 boxplot. Box bounds represent the 25th and 75th percentile, center line represents the median, and whiskers
1250 indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). (B)
1251 Representative image of Pias-1/RGA4-meditated HR in *N. benthamiana*. (C) Immunoblot analysis of Pias-
1252 1:HA, RGA4:HA, RGA4-Oau:HA (cloned from *O. australiensis* accession W0008), RGA4-Oru:HA (cloned
1253 from *O. rufipogon* accession W1943), and RGA4-Ogr:HA (cloned from *O. granulata* accession W0067B)
1254 proteins detected by anti-HA (α-HA) antibody. Coomassie blue staining of Rubisco small subunit shows
1255 equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The
1256 positions of molecular size marker are indicated on the left (kDa)
1257
1258 **Suppl. Fig. 15 HR-like cell death after overexpression of Pias-1:HA helper and Pias-1/RGA5 sensors**
1259 **in *N. benthamiana*.**
1260 (A) Boxplots of autofluorescence values after transient expression of Pias-1:HA, FLAG:Pias-2, and
1261 FLAG:RGA5 separately or in combination in *N. benthamiana*. The number of spots inoculated with *A.*
1262 *tumefaciens* are indicated at the top of the boxplot. Box bounds represent the 25th and 75th percentile, center
1263 line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x
1264 interquartile range (IQR). (B) Representative image of HR after transient expression of Pias-1:HA,
1265 FLAG:Pias-2, and FLAG:RGA5 separately or in combination in *N. benthamiana*. (C) Immunoblot analysis
1266 of Pias-1:HA protein detected by anti-HA (α-HA) antibody and FLAG:Pias-2 and FLAG:RGA5 proteins
1267 detected by anti-FLAG antibody. Coomassie blue staining of Rubisco small subunit shows equal protein
1268 loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of
1269 molecular size marker are indicated on the left (kDa).
1270
1271 **Suppl. Fig. 15 (continued). HR-like cell death after overexpression of Pias-1 helper, Pias-2 sensor, and**
1272 **AVR-Pias in *N. benthamiana*.**
1273 (D) The rice line HW-RIL7 with *Pias* recognizes the *M. oryzae* Ao-92-06-2 isolate with *AVR-Pias:FLAG*
1274 (Ao92-06-2+*pex22p*:*AVR-Pias:FLAG*) and shows resistance to this isolate. (E) Boxplots of autofluorescence
1275 values after transient expression of Pias-1:HA, FLAG:Pias-2, and AVR-Pias:FLAG separately or in
1276 combination in *N. benthamiana*. The number of spots inoculated with *A. tumefaciens* are indicated above
1277 the boxplot. Box bounds represent the 25th and 75th percentile, center line represents the median, and

1278 whiskers indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). (F)
1279 Representative image of HR after transient expression of Pias-1:HA, FLAG:Pias-2, and FLAG:RGA5
1280 separately or in combination in *N. benthamiana*. (G) Immunoblot analysis of Pias-1:HA protein detected by
1281 anti-HA (*α*-HA) antibody and FLAG:Pias-2 and AVR-Pias:FLAG proteins detected by anti-FLAG antibody.
1282 Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands expressed
1283 from the constructs are marked by red asterisks. The positions of the molecular size marker are indicated on
1284 the left (kDa).
1285
1286 **Suppl. Fig. 16. Co-immunoprecipitation shows that Pias-1 interacts with Pias-2.**
1287 (A) Co-immunoprecipitation (Co-IP) of Pias-1 (MHD mutant):HA with FLAG:Pias-2 or FLAG:GFP
1288 (negative control), as well as Co-IP of RGA4 (MHD mutant):HA with FLAG:RGA5 or FLAG:GFP
1289 (negative control) were performed. Pias-1 (MHD mutant):HA and FLAG:Pias-2 or FLAG:GFP were
1290 transiently co-expressed in *N. benthamiana.* Similarly, RGA4 (MHD mutant):HA and FLAG:RGA5 or
1291 FLAG:GFP were transiently co-expressed in *N. benthamiana*. Instead of the wild-type Pias-1 and RGA4,
1292 their MHD mutants (TYG to MHD in ARC2 subdomain as described in Cesari et al. 2014) were used to
1293 avoid HR-like cell death that reduces protein accumulation. We judge the weak bands detected by anti-HA
1294 antibody after Co-IP in Pias-1(MHD):HA+FLAG:GFP (lane 2) and RGA4(MHD):HA+FLAG:GFP (lane 4)
1295 are non-specific. Bound fractions were analyzed by immunoblotting using anti-FLAG and anti-HA
1296 antibodies. Coomassie blue staining of the Rubisco small subunit shows equal protein loading. The protein
1297 bands expressed from the constructs are marked by red asterisks. The positions of the molecular size marker
1298 are indicated on the left (kDa). We obtained similar results in three independent experiments.
1299
1300 **Suppl. Fig. 16 (continued). Yeast two-hybrid assays indicate that the Pias-1 CC domain and Pias-2 CC**
1301 **domain homo- and heterodimerize.**
1302 (B) The Pias-1 and Pias-2 CC domains form homo- and heterocomplexes. A dilution series of yeast cells
1303 expressing a GAL4-AD and GAL4-BD fusion of Pias-1 1-182 (Pias-1:CC_{1-182}) and/or Pias-2 1-177 (Pias-
1304 2:CC_{1-177}) on selective media lacking Trp, Leu, Ade, and His (–HTLA) and non-selective media lacking Trp
1305 and Leu (–TL) with 5-bromo-4-chloro-3-indolyl α-D-galactopyranoside (X-a-gal). GFP was used as the
1306 negative control. Photographs were taken 5 days after growth. (C). Immunoblot analysis confirms the protein
1307 production in the Y2H assay shown in (B). The bait protein was tagged with the Myc epitope and the prey
1308 protein with the HA epitope. The protein bands expressed from each vector are marked by red asterisks. The
1309 positions of the molecular size marker are indicated on the right (kDa). We obtained similar results in three
1310 independent experiments. We attempted Y2H assay using the full-length Pias-1 and Pias-2 constructs.
1311 However, we could not detect interactions between the full-length Pias-1 and Pias-2.
1312
1313 **Suppl. Fig. 17. HR-like cell death caused by Pias1 and RGA4-Ogr expression is suppressed by RGA5**
1314 **expression, and additional AVR-Pia expression induces cell death.**

1315 (A) Boxplots of autofluorescence values after transient expression of Pias-1:HA, FLAG:RGA5, and AVR-
1316 Pia separately or in combination in *N. benthamiana* leaves. The number of spots inoculated with *A.*
1317 *tumefaciens* are indicated at the top of the boxplot. Box bounds represent the 25th and 75th percentile, center
1318 line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x
1319 interquartile range (IQR). (B) Immunoblot analysis of Pias-1:HA protein detected by anti-HA (α-HA)
1320 antibody, FLAG:RGA5 protein detected by anti-FLAG antibody, and AVR-Pia detected by anti-AVR-Pia
1321 antibody. Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands
1322 expressed from the constructs are marked by red asterisks. The positions of molecular size markers are
1323 indicated on the left (kDa). (C) Boxplots of autofluorescence values after transient expression of RGA4-
1324 Ogr:HA, FLAG:RGA5, and AVR-Pia separately or in combination in *N. benthamiana* leaves.The number
1325 of spots inoculated with *A. tumefaciens* are indicated at the top of the boxplot. Box bounds represent the
1326 25th and 75th percentile, center line represents the median, and whiskers indicate the range of the maximum
1327 or minimum data within a 1.5 x interquartile range (IQR). (D) Immunoblot analysis of RGA4-Ogr:HA
1328 protein detected by anti-HA (α-HA) antibody, FLAG:RGA5 protein detected by anti-FLAG (α-FLAG)
1329 antibody, and AVR-Pia detected by anti-AVR-Pia (α-AVR-Pia) antibody. Coomassie blue staining of
1330 Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are
1331 marked by red asterisks. The positions of molecular size markers are indicated on the left (kDa).
1332
1333 **Suppl. Fig. 18. Results of inoculation of HW-RIL7, Sasanishiki, and Sas1493 rice plants with *M. oryzae***
1334 **isolates Ao92-06-2, Ao92-06-2 +*pex22p:AVR-Pia*, and Ao92-06-2 +*pex22p:AVR-Pias*.**
1335
1336 **Suppl. Fig. 19. The DUF761-containing gene family in rice.**
1337 A BLASTP search was performed using the amino acid sequence of the DUF761 domain of Pias-2 as a
1338 query. Fifteen DUF761 domain-containing genes were retrieved using a cutoff e-value < 10. The middle
1339 panel shows the amino acid sequence alignment, and the right panel shows the domain structure of each
1340 protein.
1341

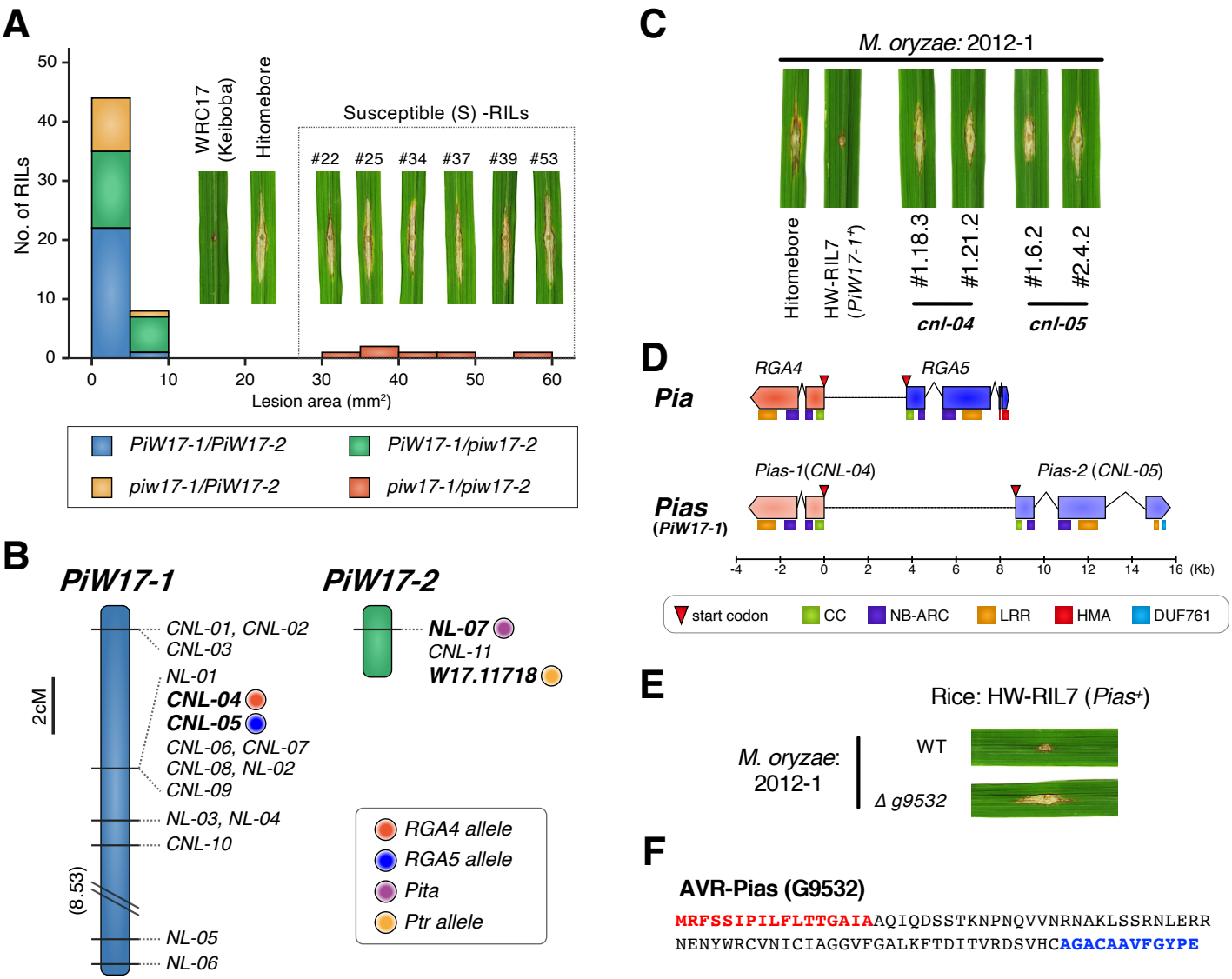
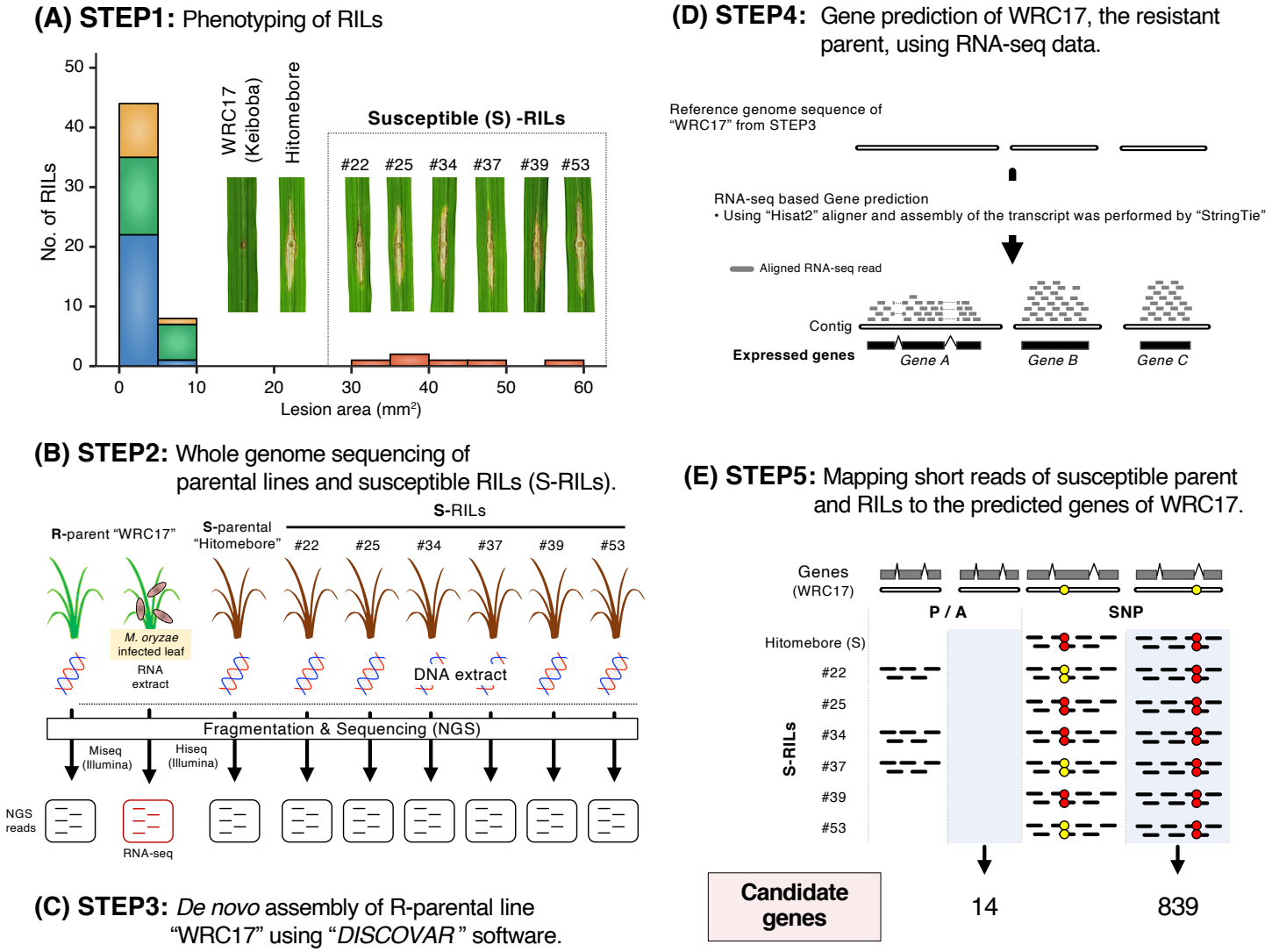


Figure 1. The *Pias* gene of rice line WRC17 (cultivar Keiboba) encodes a CC-NLR protein and is allelic to *Pia*.

(A) Segregation of the resistance and susceptibility traits among the 58 RILs derived from a cross between WRC17 (cultivar Keiboba) and Hitomebore. Disease symptoms of WRC17, Hitomebore and six RILs showing a susceptible phenotype after punch inoculation of *Magnaporthe oryzae* isolate 2012-1 (leaf photographs) and the frequency distribution of disease lesion area of the 58 RILs (bar graphs). (B) Linkage maps of candidate NLR genes at the *PiW17-1* and *PiW17-2* loci. (C) Both *CNL-04* and *CNL-05* are required for *PiW17-1*-mediated resistance against *M. oryzae* 2012-1. HW-RIL7 contains only *PiW17-1* and is resistant to *M. oryzae* 2012-1. Knockout of *CNL-04* (*cnl-04*) and *CNL-05* (*cnl-05*) in HW-RIL7 rendered plants susceptible to 2012-1. (D) Gene structures of *Pia* consisting of *RGA4* and *RGA5* and *Pias* consisting of *Pias-1* and *Pias-2*. The positions of protein domains (CC, NB-ARC, LRR, HMA, and DUF761) encoded by the NLRs are indicated. (E) The *M. oryzae* 2012-1 *AVR-Pias* knockout mutant became virulent to HW-RIL7. (F) Amino acid sequence of G9532 protein (AVR-Pias). The secretion signal is indicated by red letters and the Toxin18-like motif by blue letters. The Toxin18-like motif was annotated by Pfam (https://pfam.xfam.org/).



Suppl. Fig. 1. The RalDeN method used to identify candidate genes responsible for the resistance of rice line WRC17 to *M. oryzae* 2012-1.

(A) Segregation of the resistance and susceptibility traits among the 58 RILs derived from a cross between WRC17 (cultivar Keiboba) and Hitomebore. (B) The whole-genome sequence of the resistant parent WRC17 was obtained by Illumina sequencing. RNA-seq of WRC17 leaves inoculated with *M. oryzae* was also performed. Next, short reads of whole-genome sequences of the susceptible parent Hitomebore and the six susceptible RILs were obtained by Illumina sequencing. (C) *De novo* assembly of R-parental line WRC17 was performed using *DISCOVAR* software. (D) Expressed genes of WRC17 were identified by mapping RNA-seq reads to the WRC17 reference genome. (E) Mapping of short reads of the susceptible parent (Hitomebore) and susceptible RILs to the predicted genes identified DNA polymorphisms shared by the susceptible lines using the RalDeN pipeline (the script of the RalDeN pipeline and details about the pipeline are available at https://github.com/YuSugihara/RalDeN).

Suppl. Table 1. Summary of Illumina sequencing results used for RaIDeN methods

Line	Sequencing platform	Trim adapters and quality filter	Total no. of paired-end (PE) reads after filtering	Max length of short reads	Purpose
Whole genome sequencing					
R-parental line					
WRC17	Miseq	No	36,047,742	250	for *de novo* assemble
S-parental line					
Hitomebore	Hiseq4000	Yes	32,016,561	126	for RaIDeN
S-RILs					for RaIDeN
#22	Hiseq4000	Yes	30,830,739	126	for RaIDeN
#25	Hiseq4000	Yes	39,330,775	126	for RaIDeN
#34	Hiseq4000	Yes	36,109,189	126	for RaIDeN
#37	Hiseq4000	Yes	34,054,341	126	for RaIDeN
#39	Hiseq4000	Yes	42,282,097	126	for RaIDeN
#53	Hiseq4000	Yes	34,013,397	126	for RaIDeN
RNAseq					
M. oryzae infected leaf of WRC17	Nextseq500	Yes	39,744,107	75	for RaIDeN

Suppl. Table 2. Summary metrics of genome assembly of the resistance rice line WRC17. Genome assembly was performed by *DISCOVAR* software.

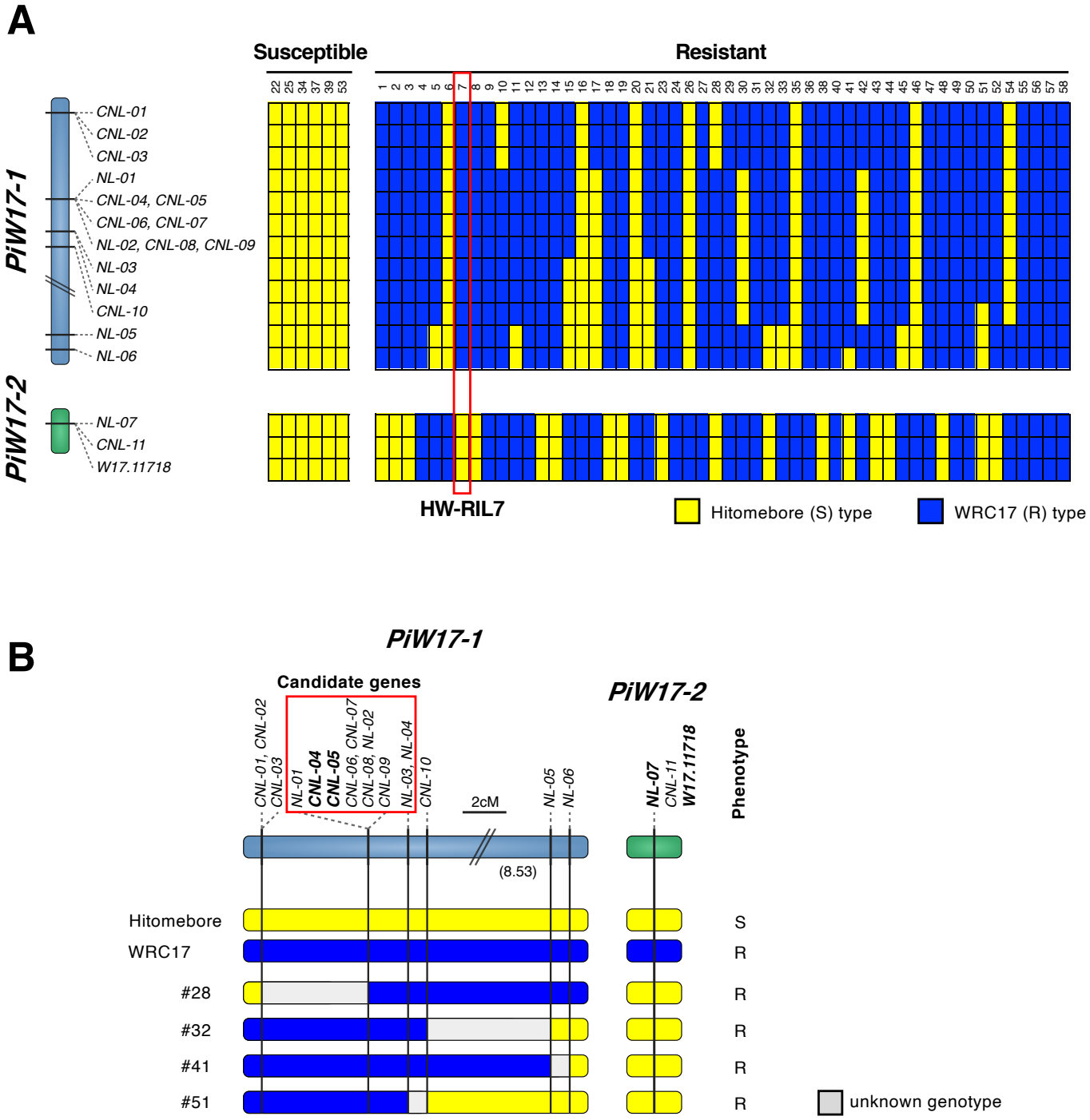
No. of contigs (>1,000bp)	25,278
Contig N50 (kb)	55.1
Largest contig (kb)	459.5
Total size of the assembled genome (Mb)	356.2

Suppl. Table 3. Summary of 38 candidate resistance gene analogs (RGAs) expressed in WRC17 leaves and possibly responsible for WRC17 resistance against 2012-1 as predicted by RGAugury (Li et al. 2016). These were selected from the 853 candidate genes as identified by the RaIDeN method. From the 38 RGAs, we selected 18 genes belonging to the categories of CNL (Coiled-Coil, NBS and Leucine rich repeat protein genes: 11 genes) and NL (NBS and Leucine rich repeat protein genes: 7 genes) as the candidate NLRs conferring resistance to WRC17 against *M. oryzae* 2012-1.

NBS encoding						TX	Other	RLP	RLK	TM-CC
NBS	CNL	TNL	CN	TN	NL					
3	11	0	1	0	7	0	0	1	12	3

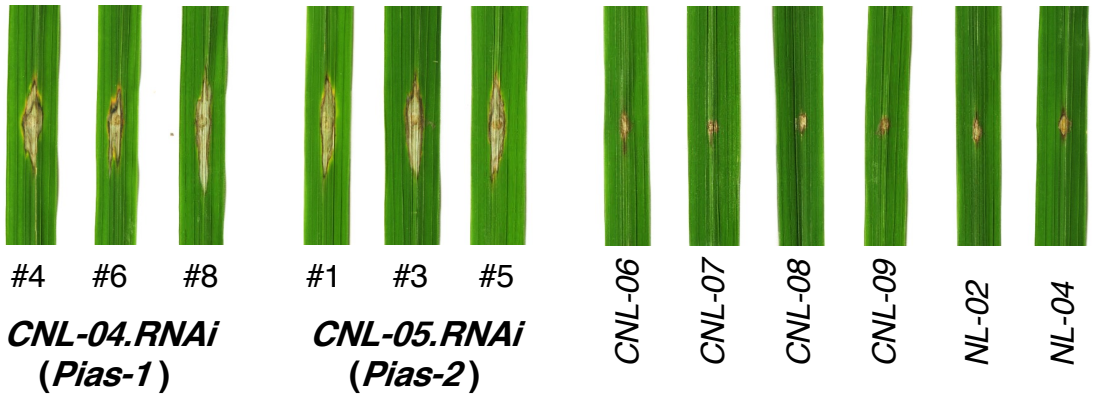
CC, C: coiled-coil
T: Toll/Interleukin-1 receptor
NBS, N: nucleotide-binding site
L: leucine rich repeat

TX: TIR-unknown domain
RLK: receptor like kinase
RLP: receptor like protein
TM: transmembrane

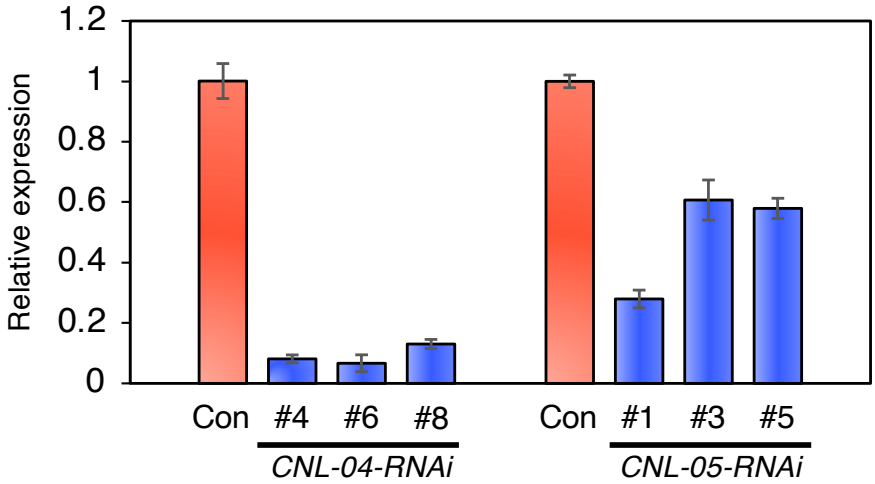


Suppl. Fig. 2. Mapping of *NLR*s in the *PiW17-1* and *PiW17-2* loci that confer resistance against *M. oryzae* 2012-1 to WRC17. (A) Association between resistance/susceptible phenotypes and genotypes among the RILs (the primers used for genotyping are shown in Suppl. Table 9). *CNL-04* and *CNL-05*, *CNL-06* and *CNL-07*, and *CNL-08*, *NL-2* and *CNL-09* are located on the same contigs. (B) Fine maps of candidate NLR genes in *PiW17-1* and *PiW17-2*. Blue and yellow indicate WRC17- and Hitomebore-type genotypes, respectively. The 10 NLRs shown in the red rectangle were considered to be the candidates of *Pi-W17-1.*

C

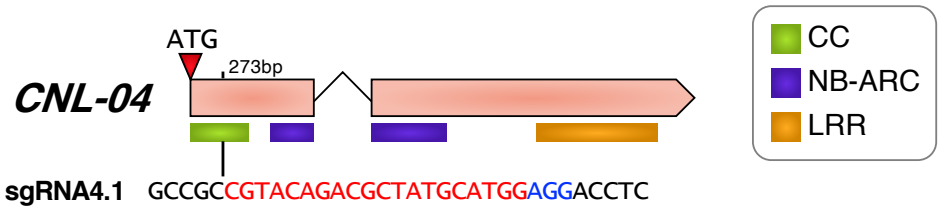


D



Suppl. Fig. 2 (continued). Results of RNAi-mediated silencing of candidate *NLR* genes.
(C) RNAi-mediated silencing of eight candidate genes encoding proteins over 900 amino acids long was performed. Gene silencing of *CNL-04* and *CNL-05* made plants susceptible to *M. oryzae* 2012-1. (D) Gene expression levels of *CNL-04* and *CNL-05* in plants transformed with the RNAi constructs. Control (Con) is WRC17 transformed with empty vector.

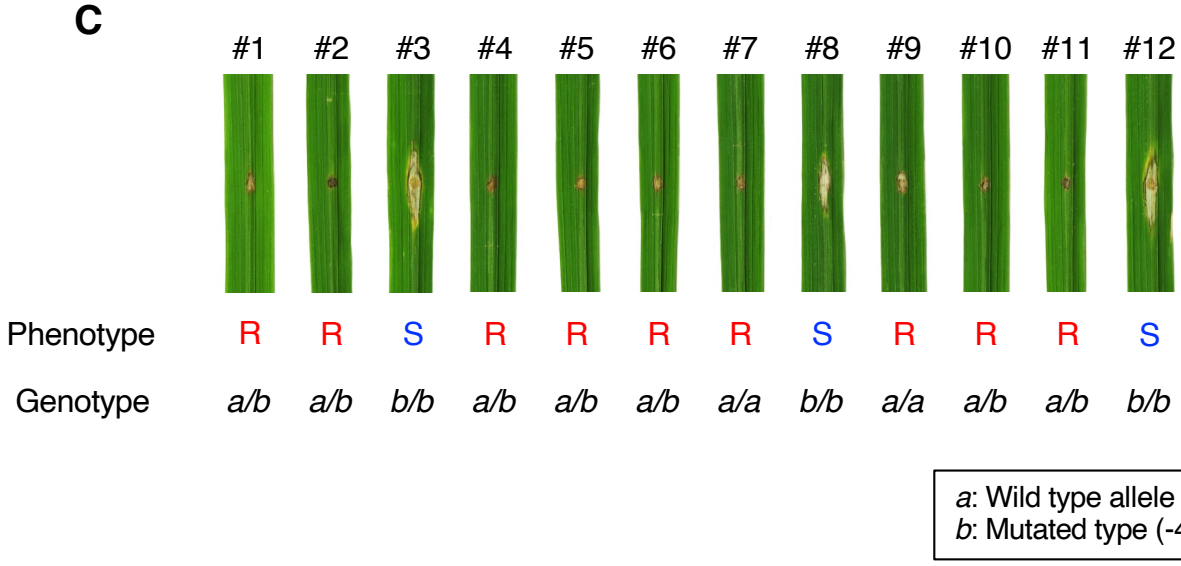
A



B

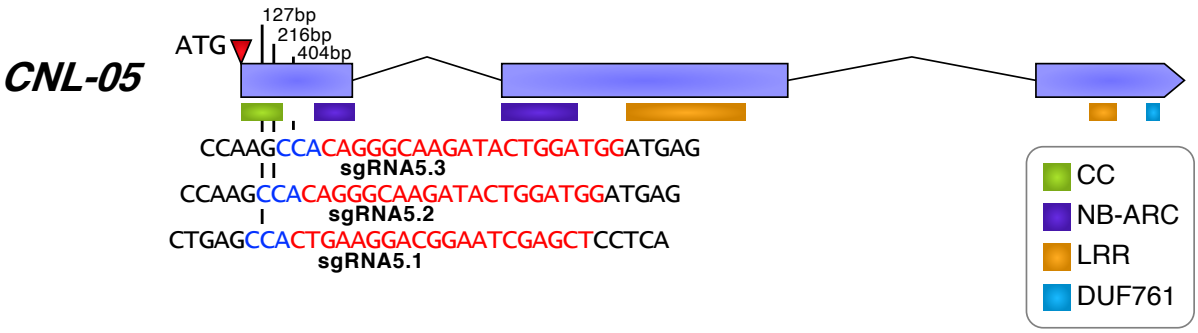
	Sequence (sgRNA4.1)	Modification
WT	GCCGCCGTACAGACGCTATGCA-TGGAGGACCTC	n/a
cnl-04_1.18.3	GCCGCCGTACAGACGCTATGC--TGGAGGACCTC	1bp deletion, premature stop 127aa
cnl-04_1.14.2	GCCGCCGTACAGACGCTATGCA-TGGAGGACCTC	n/a
	GCCGCCGTACAGACGCTA-----TGGAGGACCTC	4bp deletion, premature stop 126aa
cnl-04_1.21.2	GCCGCCGTACAGACGCTATGCACTGGAGGACCTC	1bp insertion, premature stop 162aa

C



Suppl. Fig. 3. CRISPR/Cas9-mediated knockout of *CNL-04* abolishes *PiW17-1*-mediated resistance against *M. oryzae* isolate 2012-1.
(A) The position of guide DNA targeted to *CNL-04*. The PAM is marked with blue letters, and the sgRNA sequence is marked with red letters. (B) Sequences of the sgRNA positions in the *CNL-04* knockout lines. (C) Resistance/susceptible phenotypes of the selfed progeny of the *cnl-04_1.14.2* line heterozygous for wild-type (*a*) and mutated (*b*: 4-bp deletion) alleles. The *b/b* homozygous plants became susceptible.

D



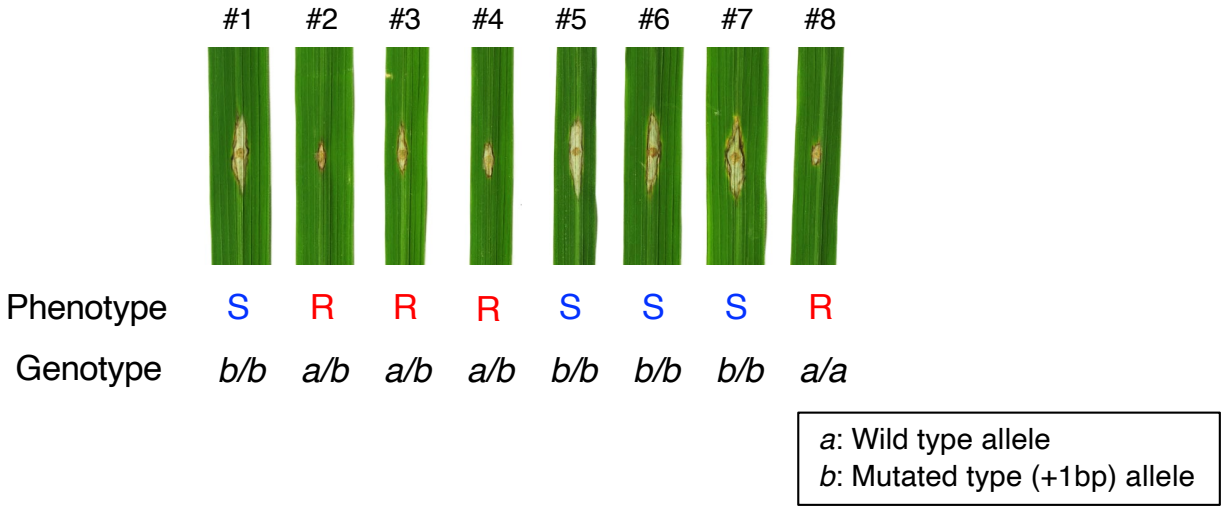
E

	Sequence (sgRNA5.1)	Modification
WT	CTGAGCCACTGAAGGACGGAATCGAGCTCCTCA	n/a
cnl-05_1.6.2	CTGAGCCACTG-AGGACGGAATCGAGCTCCTCA	1bp deletion, premature stop 45aa

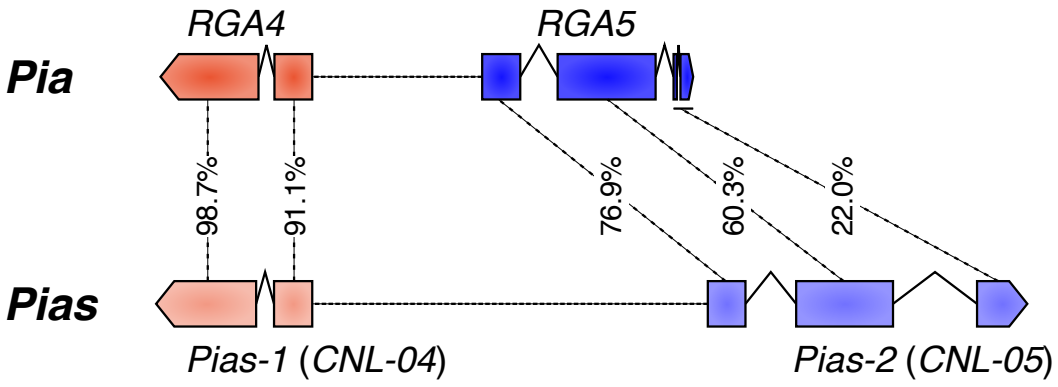
	Sequence (sgRNA5.2)	Modification
WT	CCAAGCCACAGGGCAAGATACTGGATGGATGAG	n/a
cnl-05_2.4.2	CCAAGCCACA-GGCAAGATACTGGATGGATGAG	1bp deletion, premature stop 109aa

	Sequence (sgRNA5.3)	Modification
WT	GCTCGCCAGGA-TCGCGGAGCTCAGGGCTCTGGT	n/a
cnl-05_3.7.2	GCTCGCCAGGA-TCGCGGAGCTCAGGGCTCTGGT	n/a
	GCTCGCCAGGACTCGCGGAGCTCAGGGCTCTGGT	1bp insertion, premature stop 148aa

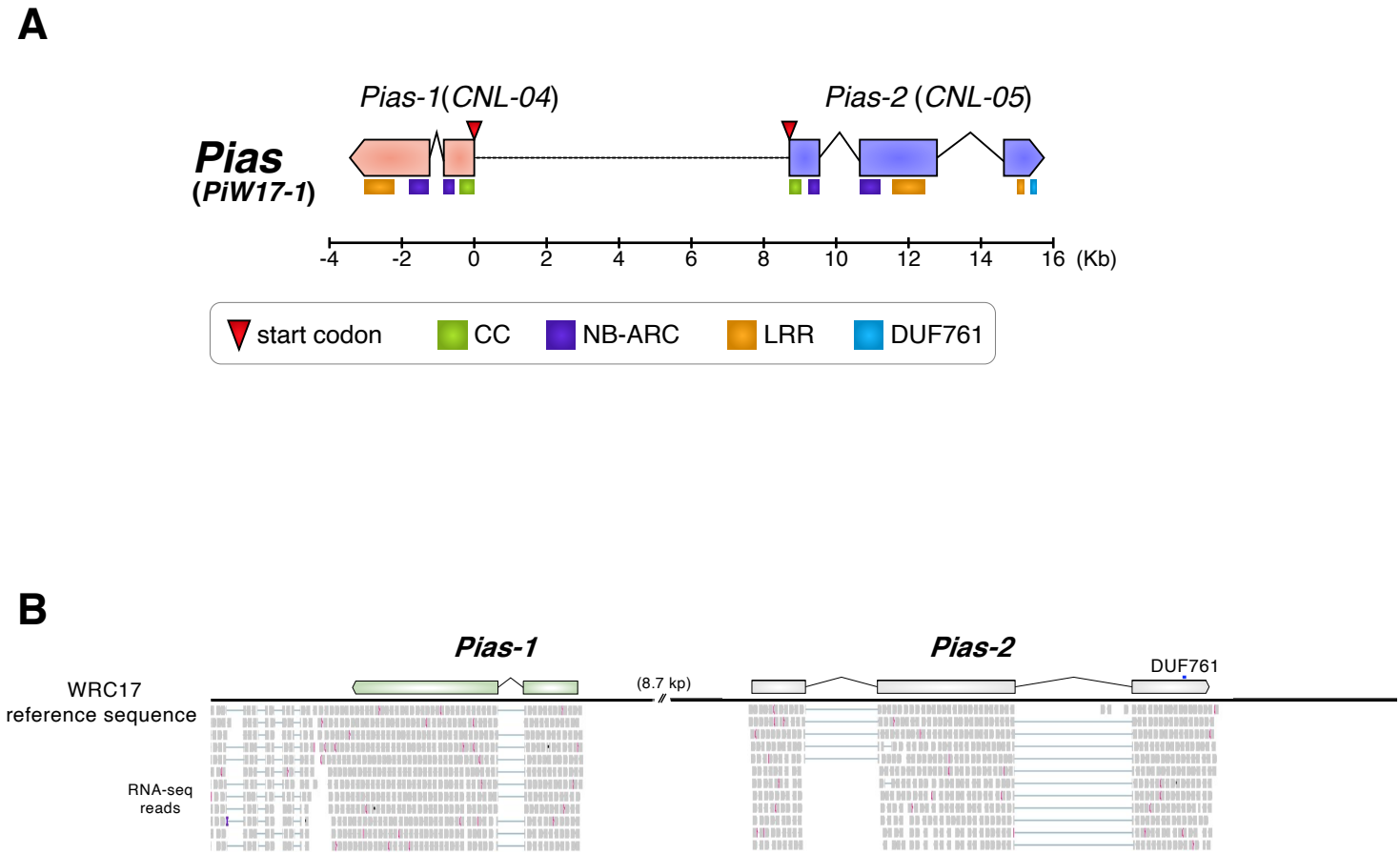
F



Suppl. Fig. 3 (continued). CRISPR/Cas9-mediated knockout of *CNL-05* abolishes *PiW17-1* mediated resistance against *M. oryzae* isolate 2012-1.
(D) The positions of guide DNAs targeted to *CNL-05*. The PAM is marked with blue letters, and the sgRNA sequence is marked with red letters. (E) Sequences of the sgRNA positions in the *CNL-05* knockout lines. (F) Resistance/susceptible phenotypes of the selfed progeny of *cnl-05_3.7.2* line heterozygous for wild-type (*a*) and mutated (*b*: 1-bp insertion) alleles. The *b/b* homozygous plants became susceptible.



Suppl. Fig. 4. Comparison of the gene structures of *Pia* (*RGA4* and *RGA5*) and *Pias* (*Pias-1* and *Pias-2*). DNA sequence similarities obtained by ClustalW are shown.



Suppl. Fig. 5. Gene structures and predicted amino acid sequences of *Pias-1* and *Pias-2*.
(A) Gene structures of *Pias-1* and *Pias-2*. (B) RNA-seq read alignment across the genome sequence of the *Pias* region as visualized by IGV (Robinson et al. 2011).

C

Pias-1

RX-CC

```
MEAALLSGFIKAILPRLFSLVNDKLNLHKGVKGDIDFLIKELRMIVGAIDDDLSVEHGAAAAAAAV
QTLCMEDLRELAHGIEDCIDGVLYRAAREQRRSSSLLPRTVRATKKLLQTNQHLAQELQRLKRMVE
E

ANQRKQRYTAAAPGQHGQVYSSAAAQVDEPWPSCSSASDPRIHEADLVG
```

NB-ARC

```
VDADRAELLEQLAERQPEQLKVIAIVGFCGLGKTALAAEAYNRETRGGRFERHAWVCAAHRSAREV
LGELLRRIDAACHGDSDAGQLCVDIRQQLEKKRYFIVIDDIQTEDQWKSIKSAFPTDKDIGSRIVV
TTTIQSVANACCSANGYLHKMSRLDKNCSKQLLSKKACPERYSHYKQPDSAAILKKCDGQPLALVT
IGEFLQANGWPTGPNCEDLCNRLHYHLENDKTLERMRRVLVRNYTSLPGHALKACLLYFGMFPSDH
PIRRKSLLRRWLAEGFVEPV

SSSSNLDSTAAFDVLMDRNIIEPINVSNNDKVKTCQTYGMMREFISHMSISQNFVTFFCDDKFLPK
Y
```

LRR

```
VRRLSLHGDTVVNGDNFNGIDLSL
VRSLAVFGEAGTTVLDFSKYQL
LRVLDLEKCDDLNDDHLKEICNLVL
LKYLSLGGNISKLPKDIAKLKD
LEALDVRRSKVKIMPVEVFGLPC
LIHLLGKFKLSDKVKQKTEVQEFLLKGKSN
LQTLAGFASNGSEGFLHLMRYMNK
LRKLKIWCTSSAGSTDWTDLREAIQQFILDEKEANIG
TRSLSLHFTGCSEDAINSLKEPCY
LSSLKLHGNFPQLPQFVTSLRG
LKELCLSSTKFTTGLLEALSNLSY
LQYLKLVADELEKFIIKVQGFPR
LLRLCIVLQCPTFPVIEEGALPF
LVTLQLLCKDLHGLSDIQIECFKH
LQEVTLHSGVTPATRQEWVKAAKEHPN
RPKVLLLKSVD

TAESEHTDVDSVMEAVKSETTEYYIAPEGPEQVIDMNNKMQLDHGLESSSVLNKQNNFADQSSSKD
QLHYSFNNMGLSDVSPAVSELPNGMVPSCT
```

Pias-2

RX-CC

```
MDAPVSFSLGAMGPLLRKLDSLPVAPEIRLPEPLKDGIELLKEDLEEIGAAVVEQSTVDSPSHRAR
YWMDEVRDLSYHIEDCIDTMFSMRCGGDDGKPRSVRR

HKVGRVKVDGFSKTQKPCTRLARIAELRALVREASERHERYQLGDGRASSSSSSSHRVFTAHGQVP
APCRNLVGM
```

NB-ARC

```
DEPKTKLTNMLTDEAELHMKVVCILGSAGIGKTTLAEQVYRKLRWQFDCHAFVRASRKPDMRRLLG
AILSQVQPRIRISDTSTVQSLIDNLWEYLQKKRYFIVIDELYETATWDIITSAFPEDNNCSRIMTT
AGIEGVALECCSYHSVNIFKMKPLGLDDSAKLFFNRVFGSEQQCPYELSEVSYRITAKCGGLPLAV
IIIAGLLASLPCKTELWYNIDGCLCSSVTTDIDLDEILKEIISLGYDNLPHYLKTCLLYLSLYSEG
FIIWTADLLKQWISEGFIAVIDG

EDIEEVAESYFYNLVNRGMIQSVKTKYNNQVLCTVHHTVFDLIIHKSKEEKFISAIDYSQTMPGNS
LEARRLSFHFSNTRYATEVAGITLSQVRSFAFLGLLKCMPSIMEFKL
```

LRR

```
LRVLILEFWGDNHGCMSFNVARICRLFQ
LRYLKISSQIIIELPAQIRGLKY
LETLEIDARVTAVPSDIIHLRS
LLHLYFQDGIVLPDGIGCIRSLRT
LKYFDLGSNSEENIRSLGQLTN
LRDLHLTCSAPKSNQQAKRNLVILASYTGKLGN
LKSVKFSPGDSGMDISFLFYGIGISVDRSRTASSLPVS
VRTLELPSICIFARLPDWIGQLRK
LHTLNLAVRELIENDIDSLAGLPD
LIVLSMHIMKAPMERIVFNRKAFPV
LKYFKFICGTLRMAFQAGAMAN
LHRLKLGFNAH

KGEKYDNILVGIEHLLNLKKIAVRIGGAAEAKESDRMAAEAALKEAIRKHLMFLDDLDIARVECVK
EEYKCIKKKHKIKIEDSISEKNGDSKKQHSVEKKAVWGKTMKNIADSGVFPEDYTMSREQRVAEGF
VVGIEKCRAEDAAERIIRNVPVDYDGLGQVSTSKIQDHLPELAPRAVQNEKFGSSNDLSIMIQINK
YARLPSYEWRDTDISKLNFRL
```

LRR

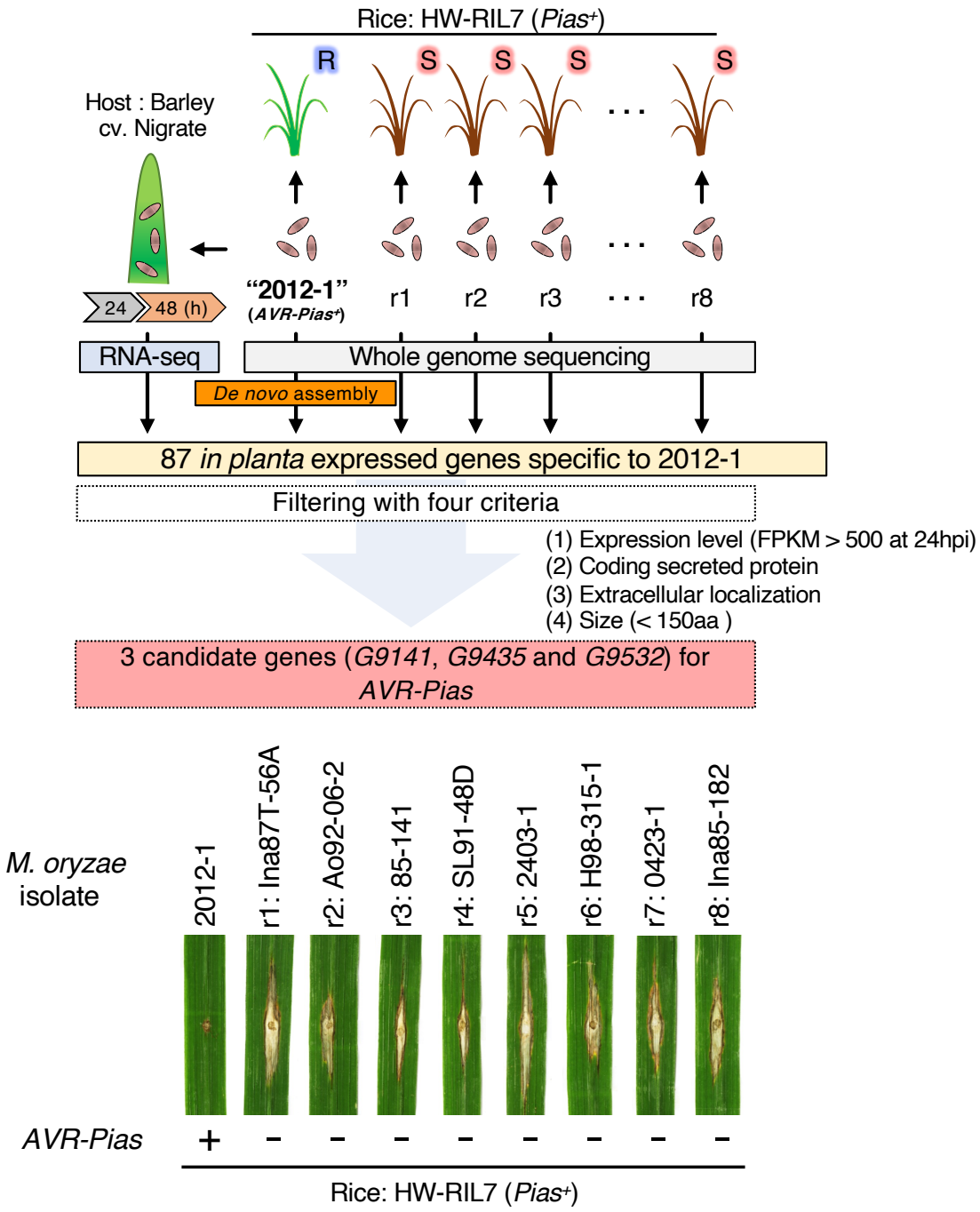
```
LRAPMLLEAVTARCHLLDLILIGSNN
ITVLDLGRPTITKLPASIECLPN
LRYLRLQGTQLKSLSEVIVKMPT
IRGLDIKNTKTEELPQGILRMKK
LSHLSMGEKQKNIQVFMEKMQT
LAETVQDSDDLSDETEGIADDDEGEFSTRANASTPKVD
```

DUF761

```
EDEVDRRANNFIARFRKQI

TIRNSGFPKRRAASTNDCGHEISMSANSPREVDDFKKKFDEIIARNRHTWKPIESPRSVKHGKYFV
RCPPSLITSAVFSFFFLSIFLVWYSLRFRILSGKP
```

Suppl. Fig. 5 (continued).
(C) Amino acid sequences and putative domains of Pias-1 and Pias-2. The RX-CC and NB-ARC domains of Pias-1 and Pias-2 were annotated by a CD-search in NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), and the LRR was predicted using LRR predictor (Martin et al 2020).



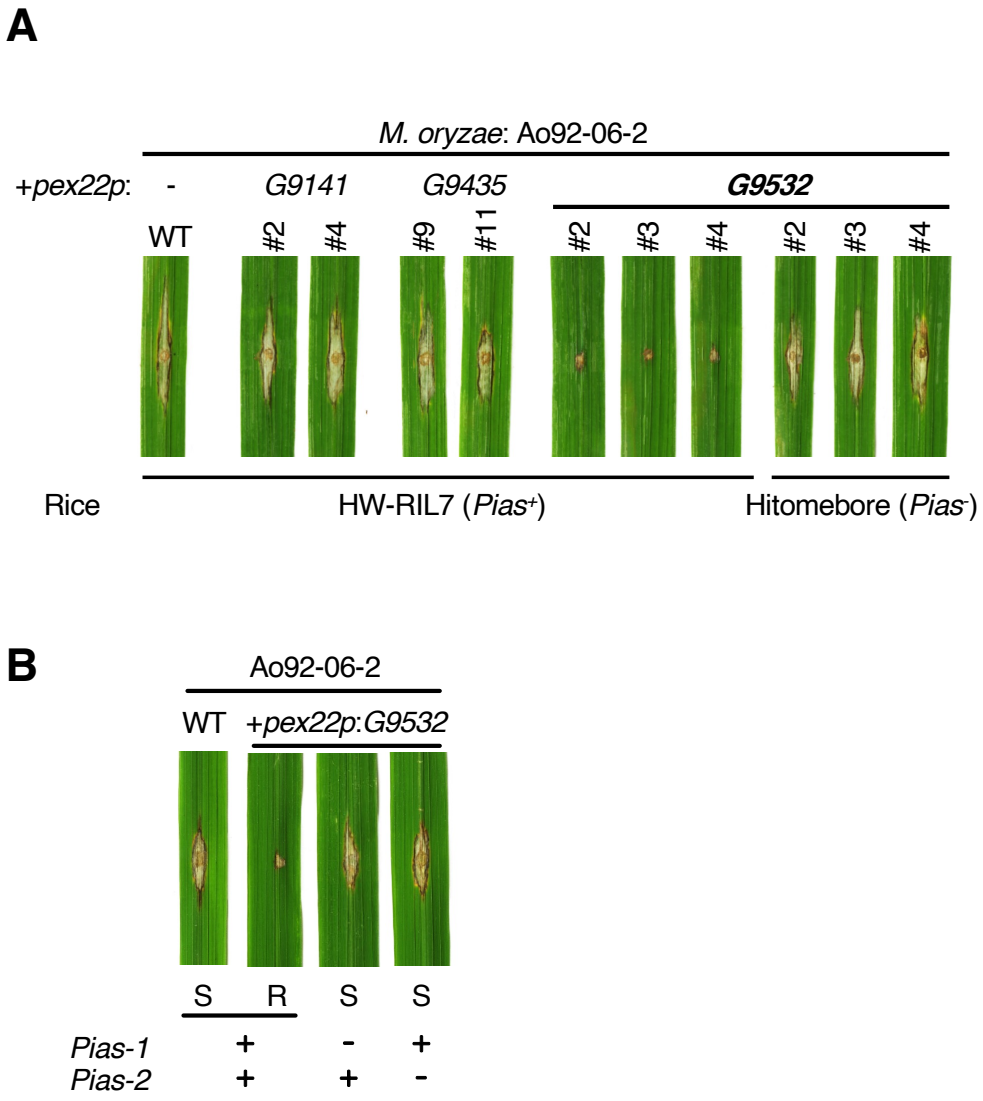
Suppl. Fig. 6. Isolation of *AVR-Pias* from *M. oryzae* isolate 2012-1.
A simplified scheme showing the procedure used to isolate *AVR-Pias*. First, the genome of isolate 2012-1 was sequenced on the Illumina platform, and the short reads were used for *de novo* assembly with *DISCOVAR*, resulting in the 2012-1 genome reference (**Suppl. Table 4, 5**). Next, the isolate 2012-1 was used to inoculate barley (*Hordeum vulgare*) cultivar 'Nigrate', which is highly susceptible to *M. oryzae* (Hyon et al. 2012), and the infected leaves were subjected to RNA-seq (Shimizu et al. 2019), resulting in the identification of 10,991 genes in 2012-1 that were expressed during host infection. Next, eight *M. oryzae* isolates (r1–r8) were selected and used for inoculation of rice line HW-RIL7 with *Pias* but without *Pi-W17-2*. The eight isolates were compatible with HW-RIL7 (result shown in the bottom), suggesting that they lack *AVR-Pias*. The eight isolates were then subjected to genome resequencing on the Illumina platform, and the RaIDeN method was used to identify only presence/absence polymorphisms (**Suppl. Table 4**). The short reads of eight isolates were aligned to the expressed genes of the 2012-1 isolate, resulting in the identification of 87 expressed genes that were specific to isolate 2012-1. These transcripts were further filtered based on four criteria: (1) genes showing a higher level of expression 24 h after inoculation (FPKM > 200) and genes encoding (2) a putative secreted protein, (3) a non-transmembrane protein and (4) a small protein (<150 amino acids). This analysis identified three transcripts (*G9141*, *G9435*, and *G9532*) as the candidates of *AVR-Pias*.

Suppl. Table 4. Summary of whole genome sequencing of the nine *M. oryzae* isolates and RNA sequencing of *M. oryzae* isolate 2012 for identification of *AVR-Pias*

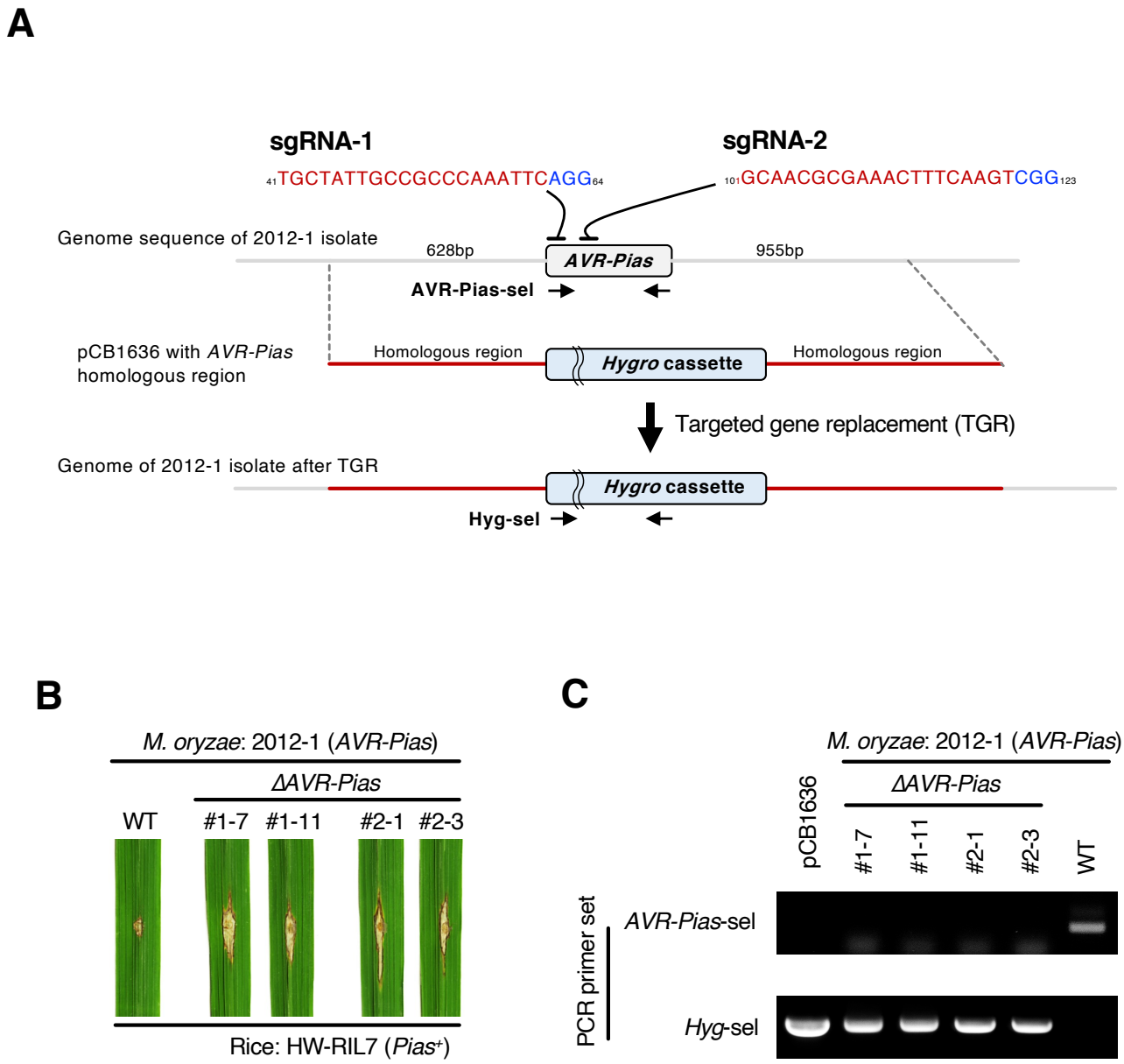
Line	Race code	Sequencing platform	Trim adapters and quality filter	Total no. of paired-end (PE) reads after filtering	Max length of short reads	Purpose
Whole genome sequencing						
2012-1		Miseq	No	3,332,000	250	for *de novo* assemble
S-pathogen for HW-RIL7						
Ina87T-56A	**r1**	Miseq	Yes	1,776,902	125	for RaIDeN
Ao92-06-2	**r2**	Miseq	Yes	2,691,934	125	for RaIDeN
85-141	**r3**	Miseq	Yes	3,034,450	125	for RaIDeN
SL91-48D	**r4**	Miseq	Yes	3,108,538	125	for RaIDeN
2403-1	**r5**	Miseq	Yes	3,388,670	125	for RaIDeN
H98-315-1	**r6**	Miseq	Yes	2,933,673	125	for RaIDeN
0423-1	**r7**	Miseq	Yes	2,996,542	125	for RaIDeN
Ina85-182	**r8**	Miseq	Yes	3,148,657	125	for RaIDeN
RNAseq						
M. oryzae isolate 2012-1 inoculated to the leaf of *Hordeum vulgare* var. Nigrate				**Total no. of reads (SE) after filtering**		
24H after inoculation		Nextseq500	Yes	15,658,123	75	for RaIDeN and expression analysis
48H after inoculation		Nextseq500	Yes	15,770,613	75	for RaIDeN and expression analysis

Suppl. Table 5. Summary metrics of genome assembly of *M. oryzae* 2012-1 isolate. Genome assembly was performed by *DISCOVAR* software.

No. of contigs (>500bp)	2,414
Contig N50 (kb)	154.7
Largest contig (Mb)	1.0
Total size of the assembled genome (Mb)	41.0



Suppl. Fig. 7. The expressed gene *G9532* of *M. oryzae* isolate 2012-1 is *AVR-Pias*.
(A) Results of inoculation of rice line HW-RIL7 with *Pias* or Hitomebore without *Pias* with *M. oryzae* isolate Ao92-06-2 wild type (WT) or Ao92-06-2 with the *G9141*, *G9435* or *G9532* transgene, all driven by the *pex22* promoter. When Ao92-06-2 contained *G9532*, the interaction became incompatible, indicating that *G9532* is *AVR-Pias*. (B) Introduction of *pex22p*:*G9532* into *M. oryzae* isolate Ao92-06-2 confers avirulence to the pathogen against HW-RIL7 with *Pias-1* and *Pias-2*. Knockout of the host gene *Pias-1* or *Pias-2* compromises the avirulence caused by *G9532*.



Suppl. Fig. 8. Deletion of *AVR-Pias* from *M. oryzae* isolate 2012-1 causes a loss of avirulence against rice line HW-RIL7 with *Pias*.

(A) Schematic overview of the target gene replacement (TGR) strategy at the *AVR-Pias* locus using RNA-guided nuclease. The PAM is marked with blue letters, and the sgRNA sequence is marked with red letters. The primers used for PCR analysis are indicated by horizontal arrows. (B) Inoculation assay of rice line HW-RIL7 with *M. oryzae* 2012-1 wild type (WT) and TGR transformants. (C) PCR analysis of TGR events at the *AVR-Pias* locus. Upper and lower images show PCR results using *AVR-Pias*- and *Hygromycin* (*Hyg*)-specific primers, respectively. pCB1636: Replacement vector containing the hygromycin resistance gene and genomic regions neighboring *AVR-Pias*. The isolates corresponding to lanes 2–5 show compatibility with rice line HW-RIL7 with *Pias*, whereas the wild-type isolate (2012-1) in lane 6 is incompatible with HW-RIL7.

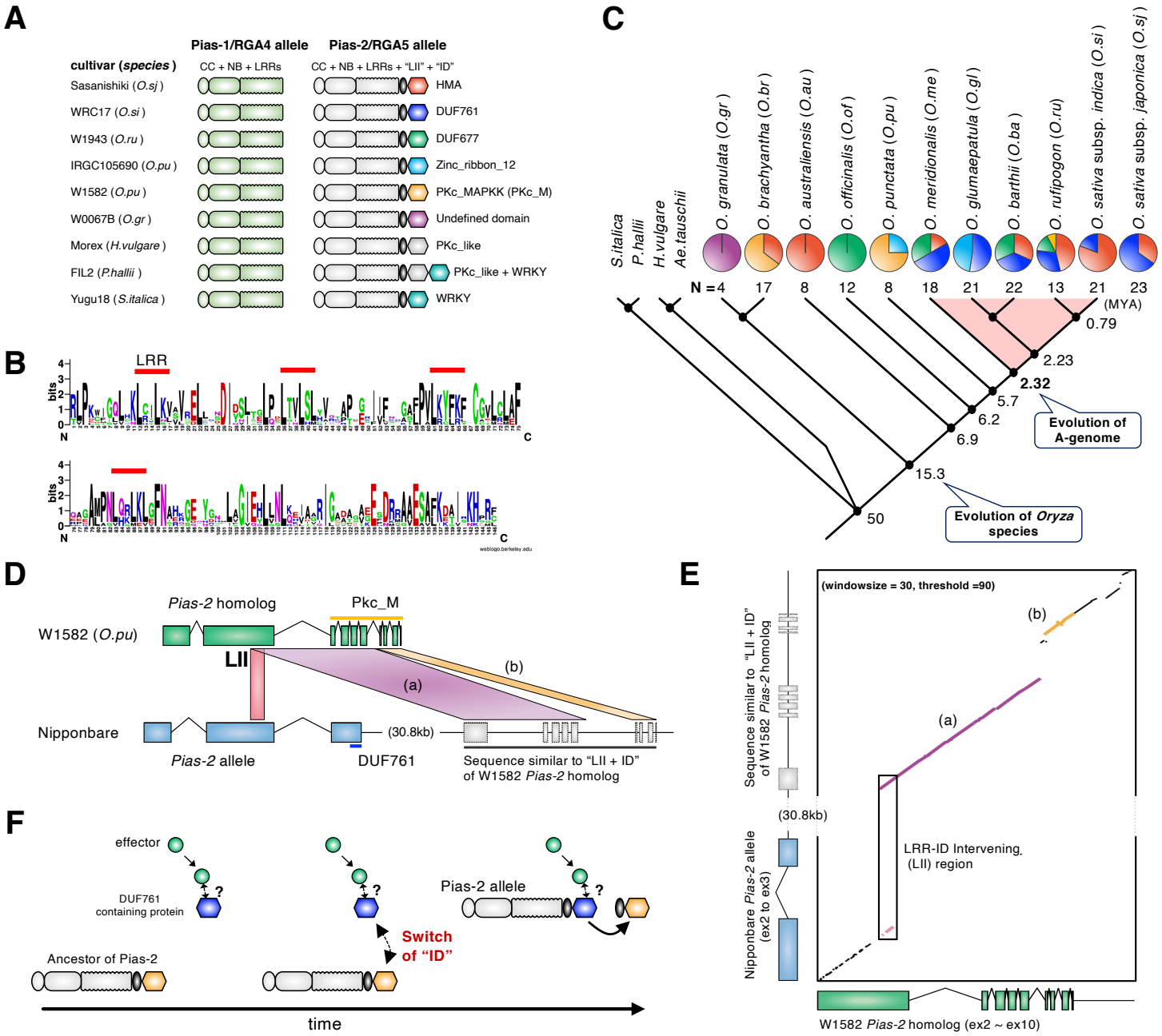
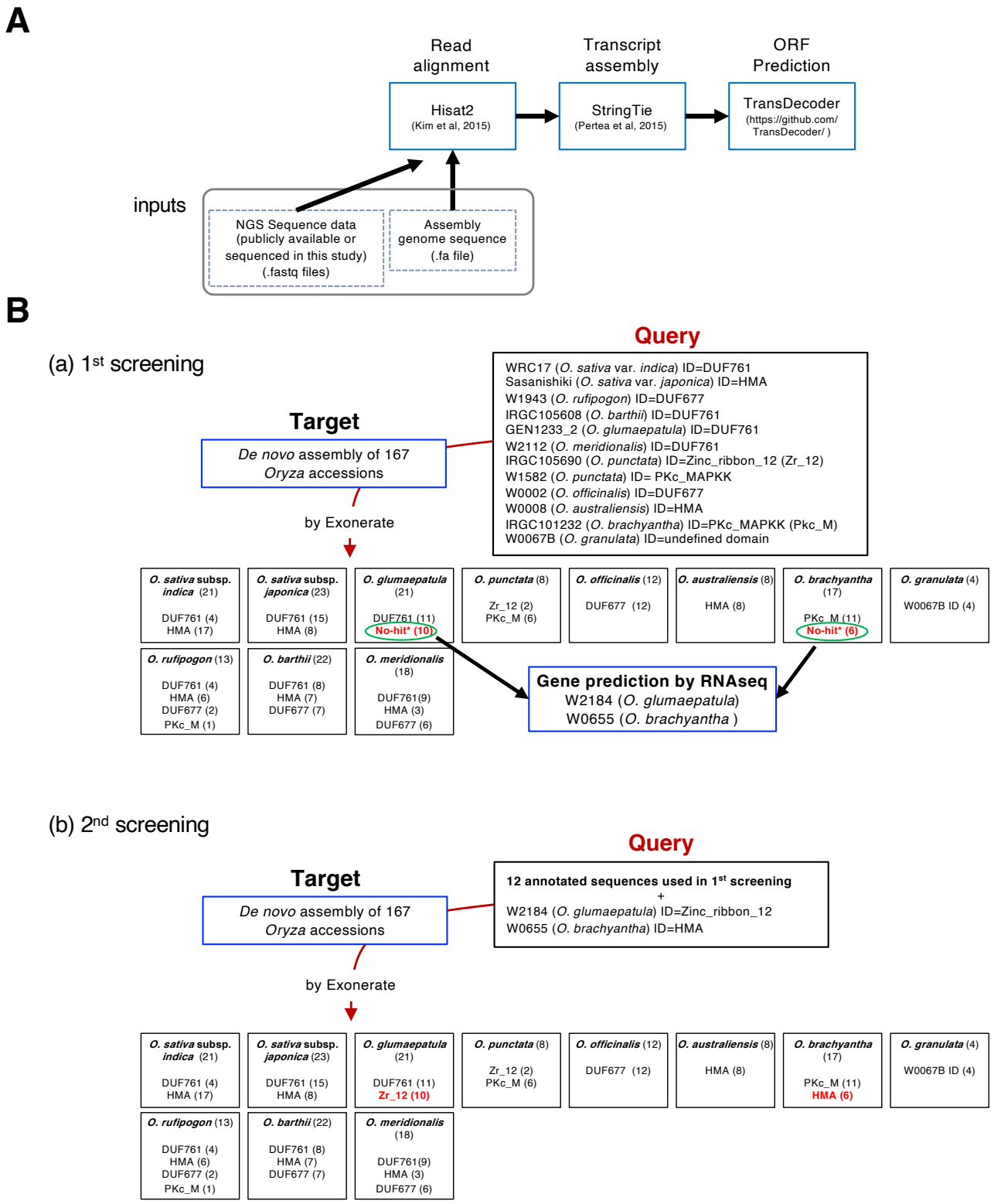
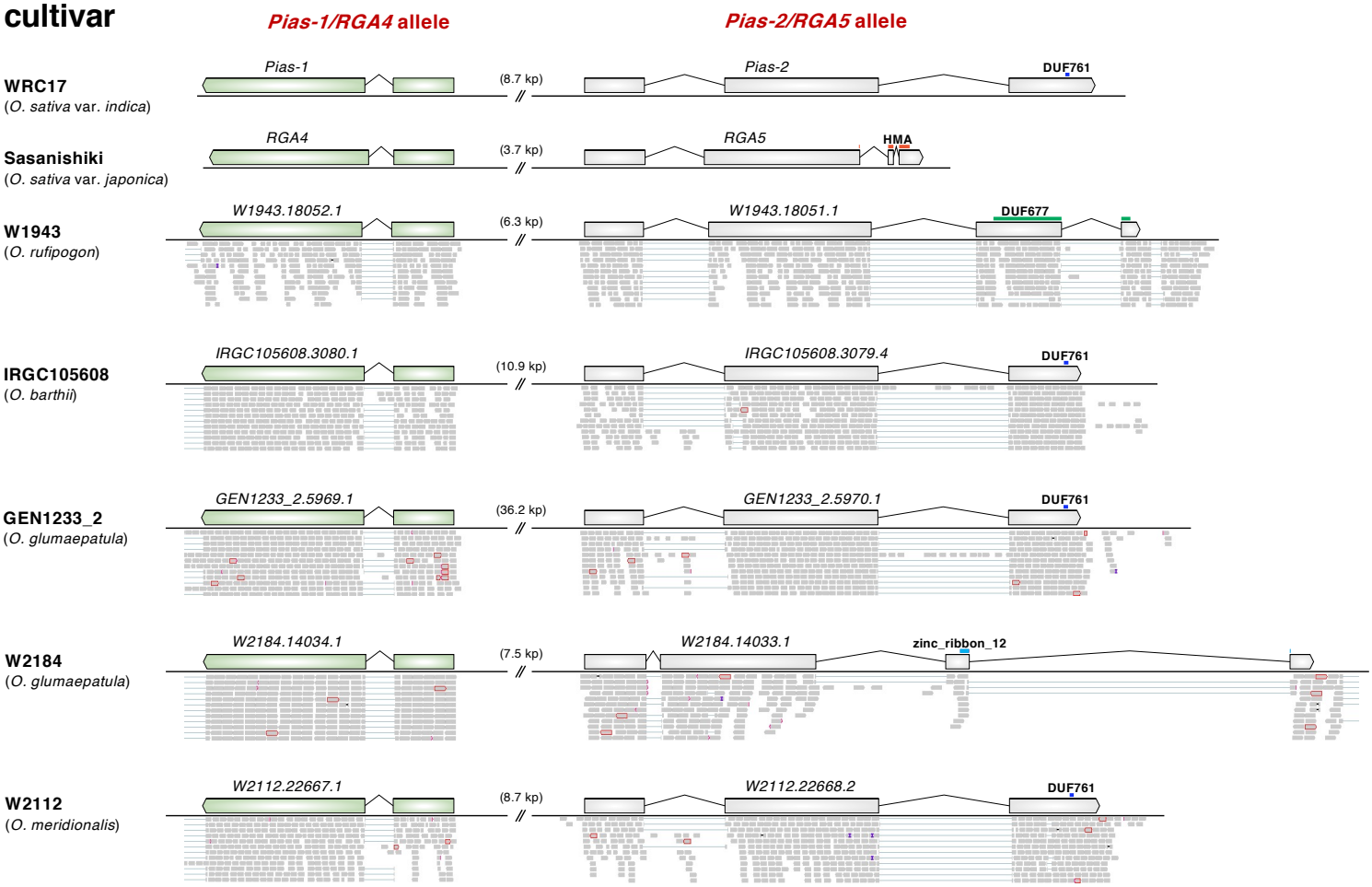


Figure 2. Recurrent integration of extraneous domains in Pias/Pia sensor NLRs.
(A) A simplified scheme of the structures of the Pias/Pia NLR pairs. Pias-1/RGA4 helper NLRs are shown in green, and Pias-2/RGA5 sensor NLRs are shown in white. The conserved LRR-ID intervening motif (LII) is indicated by black ellipses. Integrated domains are shown by different-colored hexagons. (B) A sequence logo showing conserved amino acids of the LRR-ID Intervening (LII) motif. The red lines indicate the LRR in LII motif. (C) Distribution of ID motifs among *Oryza* species. The pie charts show the frequencies of different ID motifs in a given species. The colors correspond to the ID colors in (A). The numbers below the pie charts indicate the sample numbers. A cladogram showing the phylogenetic relationships of 11 *Oryza* species and four other Poaceae species (*Setaria italica*, *Panicum hallii*, *Hordeum vulgare* and *Aegilops tauschii*) based on Time Tree - the time scale of life web-database (http://www.timetree.org/). The numbers on the branches indicate the estimated time of the splitting of lineages (MYA: million years ago). (D) DNA sequence similarity between *O. punctata Pias-2/RGA5 sensor NLR* and the downstream sequence of *O. sativa* (Nipponbare) *Pias-2/RGA5.* LII: LRR-ID intervening motif. (E) Dot-plot analysis of the *O. punctata Pias-2/RGA5 sensor NLR* and *O. sativa* (Nipponbare) *Pias-2/RGA5 NLR* downstream sequences using the Dotmatcher tool (http://emboss.bioinformatics.nl/cgi-bin/emboss/dotmatcher/). (F) Possible evolutionary process of ID replacement that might have occurred between the *O. punctata* lineage and the *O. sativa* of *Pias-2/RGA5* lineage. We still do not know the mode of interaction between the AVR-Pias effector and DUF761-containing protein, so it is indicated by "?".

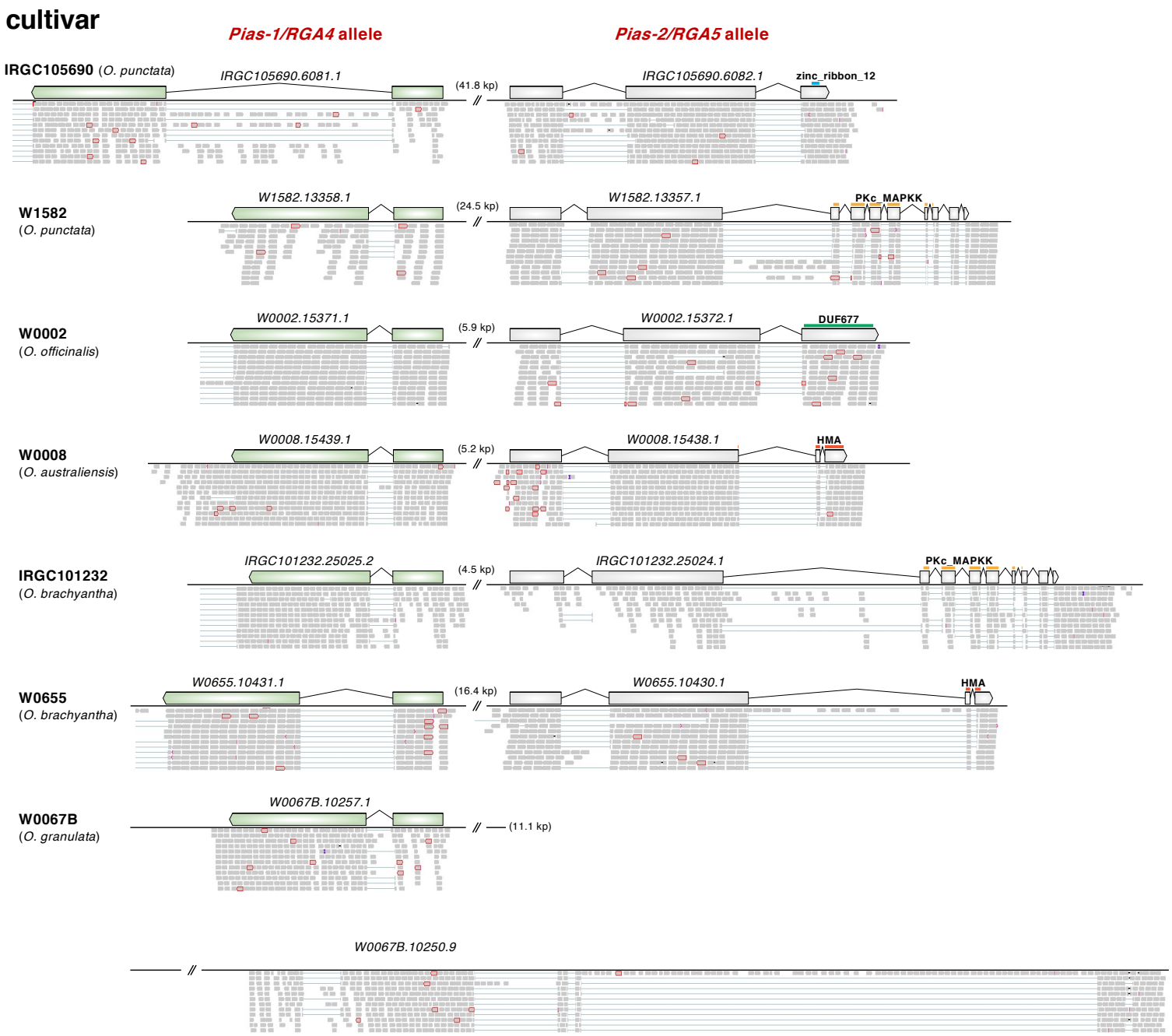


Suppl. Fig. 9. Prediction of ID sequences in the Pias/Pia sensor NLRs of the *Oryza* species.
(A) Overview of gene prediction methods used in this study. (B) Gene models of *Pias-1/RGA4* and *Pias-2/RGA5* of 12 *Oryza* samples supported by RNA-seq data were used as queries to annotate IDs in the genome assemblies of 167 *Oryza* samples using Exonerate ((http://www.ebi.ac.uk/~guy/exonerate)). However, 10 samples of *O. glumaepatula* and six samples of *O. brachyantha* did not match known domains. Therefore, we incorporated RNA-seq data of each sample from the two species, resulting in the annotation of the Zinc_ribbon_12 (*O. glumaepatula*) and HMA (*O. brachyantha*) IDs . In the second round, we used the 12 samples used in the first round of Exonerate as well as two new samples (*O. glumaepatula* W2184 and *O. brachyantha* W0655) as queries to infer IDs in the assembled genomes of the 167 *Oryza* samples.

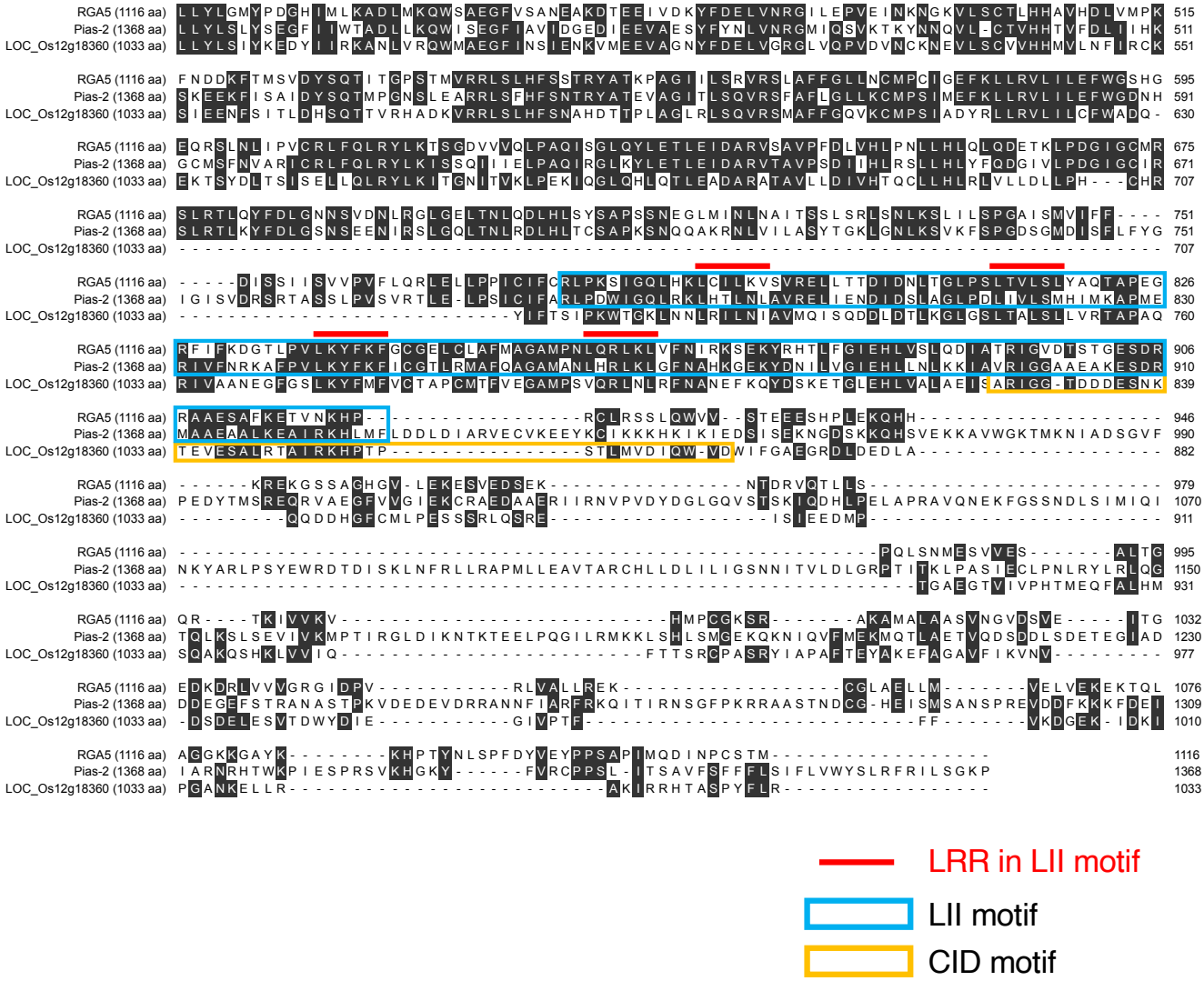
C



Suppl. Fig. 9 (continued). Prediction of ID sequences in the Pias/Pia sensor NLRs of the *Oryza* species. (C) RNA-seq read alignment across the genome sequence of the *Pias/Pia* homologous region in the AA genome *Oryza* species as visualized by IGV (Robinson et al. 2011).



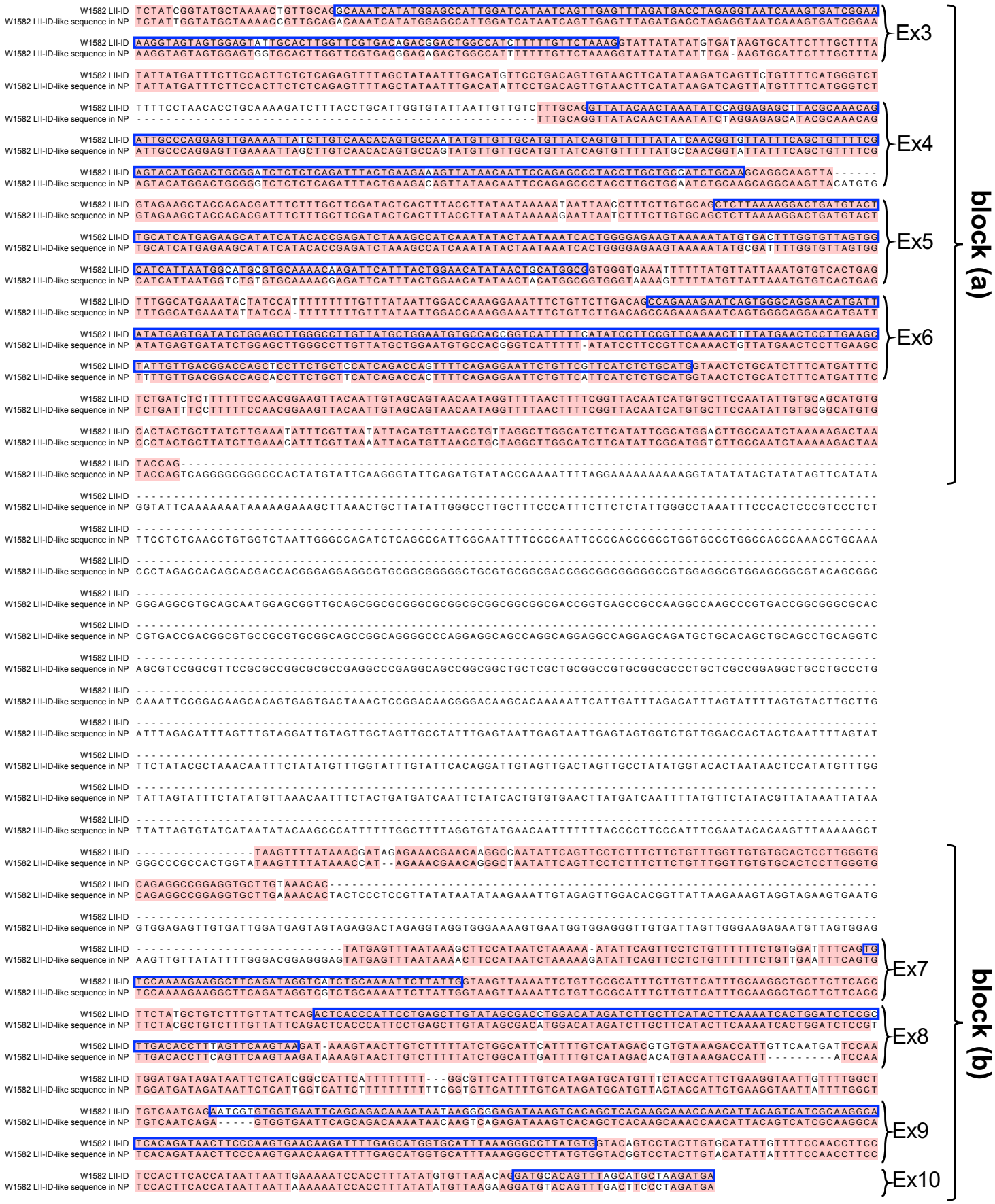
Suppl. Fig. 9 (continued). Prediction of ID sequences in the Pias/Pia sensor NLRs of the *Oryza* species. (D) RNA-seq read alignment across the genome sequence of the *Pias/Pia* homologous region in the non-AA genome *Oryza* species as visualized by IGV (Robinson et al. 2011).



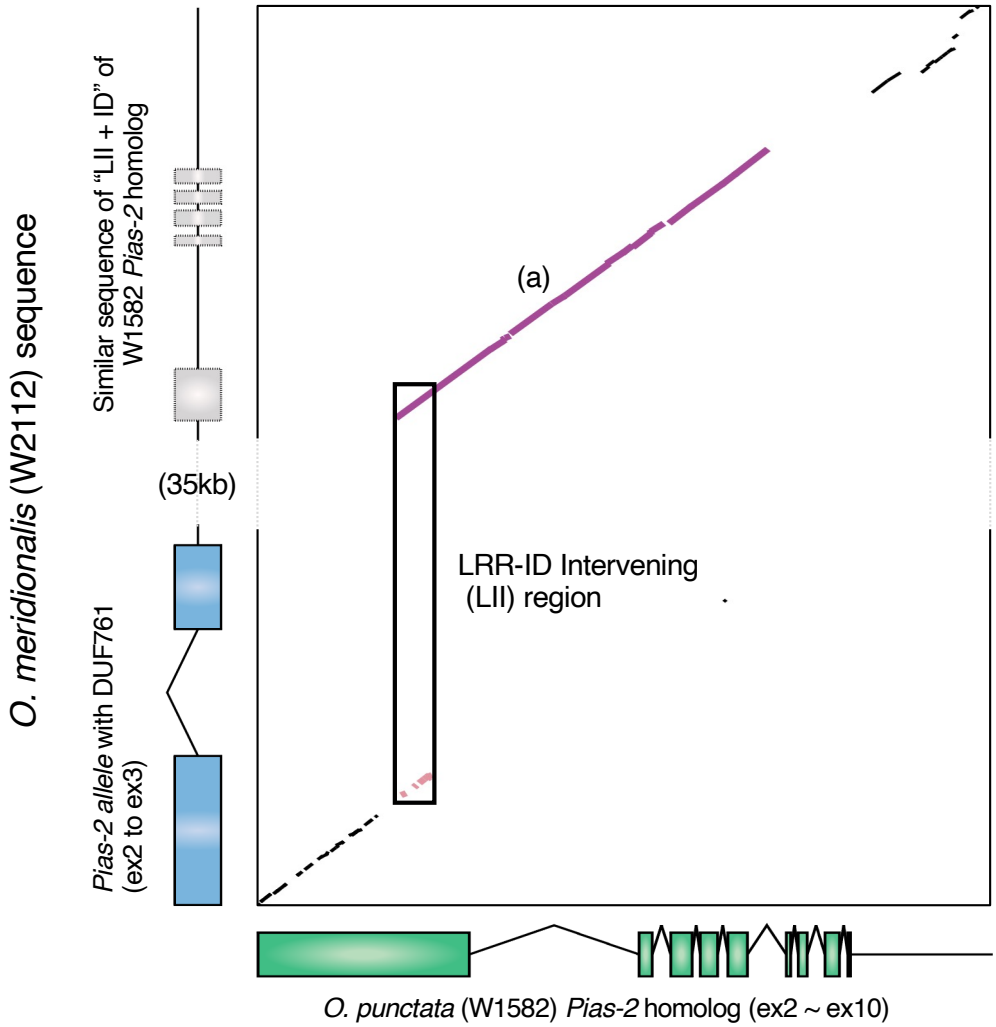
Suppl. Fig. 10. Positions of the LII motif (this study) and the CID motif (Bailey et al. 2018).
Amino acid sequence alignment of RGA5, Pias-2 and LOC_Os12g18360 and the positions of the LII motif (this study) and the CID motif as described by Bailey et al. (2018). The aqua and orange boxes indicate the LII and CID motif regions, respectively. The red lines indicate LRR in the LII motif.

Suppl. Fig. 11. The *O. punctata* W1582 LII-ID sequence is conserved in the downstream sequence of *Pias* in *O. sativa* cv. Nipponbare.

Alignment of the DNA sequences of the *O. punctata* W1582 LII-ID and the W1582 LII-ID-like sequence in *O. sativa* Nipponbare. The blue boxes and green lines indicate the exon sequence of *O. punctata* W1582 Pias-2 homolog and the LII region, respectively.



Suppl. Fig. 11. Continued.



Suppl. Fig. 12. The LII + ID sequence of the *O. punctata* (W1582) *Pias-2* homolog is conserved in the downstream sequence of *O. meridionalis* (W2112) *Pias-2*.
Dot-plot analysis of *O. punctata* (W1582) *Pias-2* homolog and the genome sequence of *O. meridionalis* (W2112). The purple line corresponds to block (a) in Figure 2. The block (b) sequence in Figure 2 is deleted in *O. meridionalis* (W2112).

Suppl. Table 9. A survey of 51 A-genome *Oryza* accessions with DUF761 ID revealed that the *O. punctata* LII-ID-like downstream sequence is widely conserved.

BLASTN search was performed on each genome sequence using *O. punctata* W1582 LII-ID-like sequence of Nipponbare cv. as a query. Block (a) and block (b) region correspond to those in Figure 2.DE.

Species	with DUF761 as ID	*O. punctata* W1582 LII-ID-like sequence conserved	
		block (a) and (b)	block (a) only
O. sativa	19	19	0
subsp. *indica*	4	4	0
subsp. *japonica*	15	15	0
O. rufipogon	4	4	0
O. barthii	8	1	0
O. glumipatula	11	0	0
O. meridionalis	9	1	7

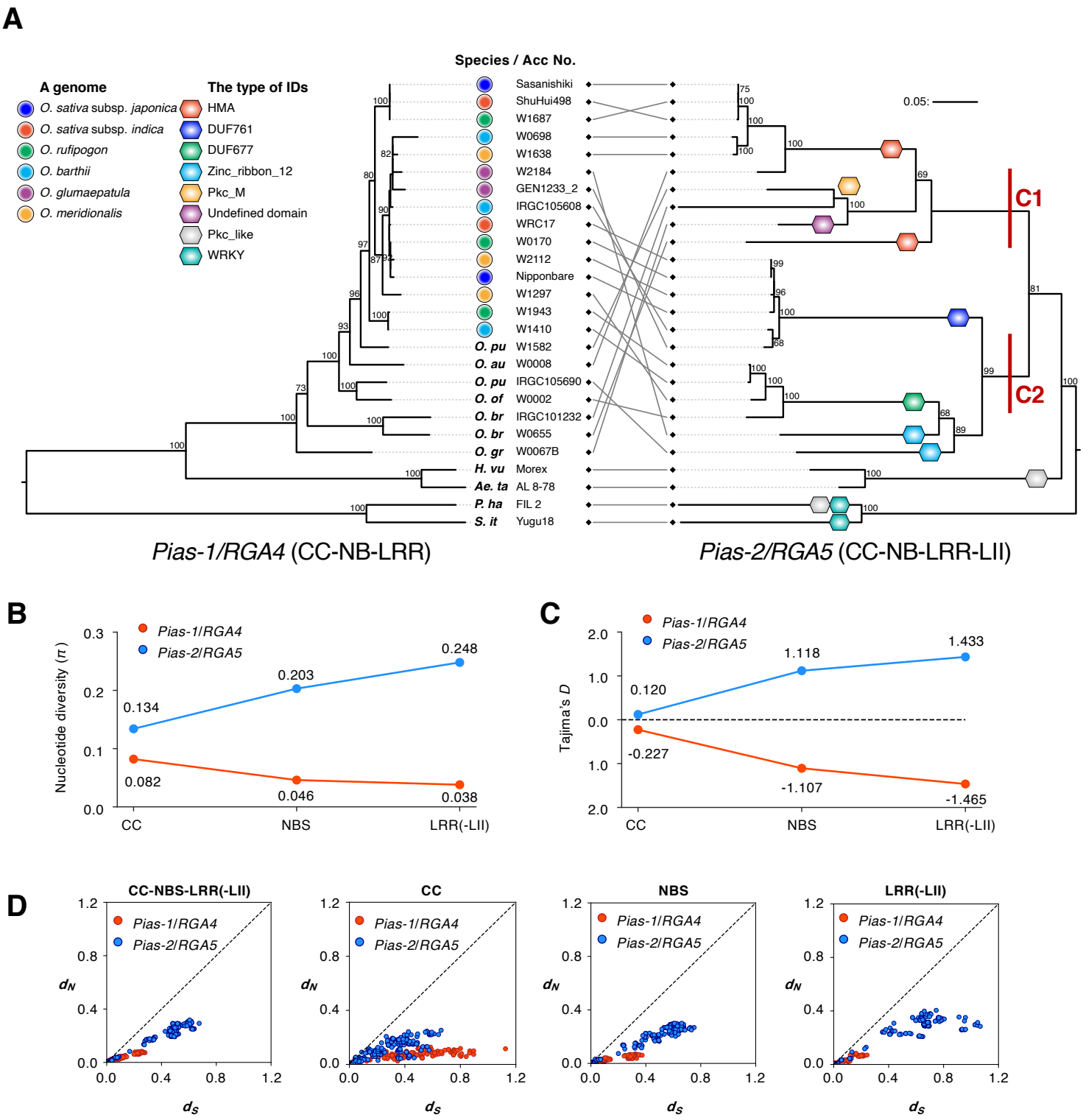
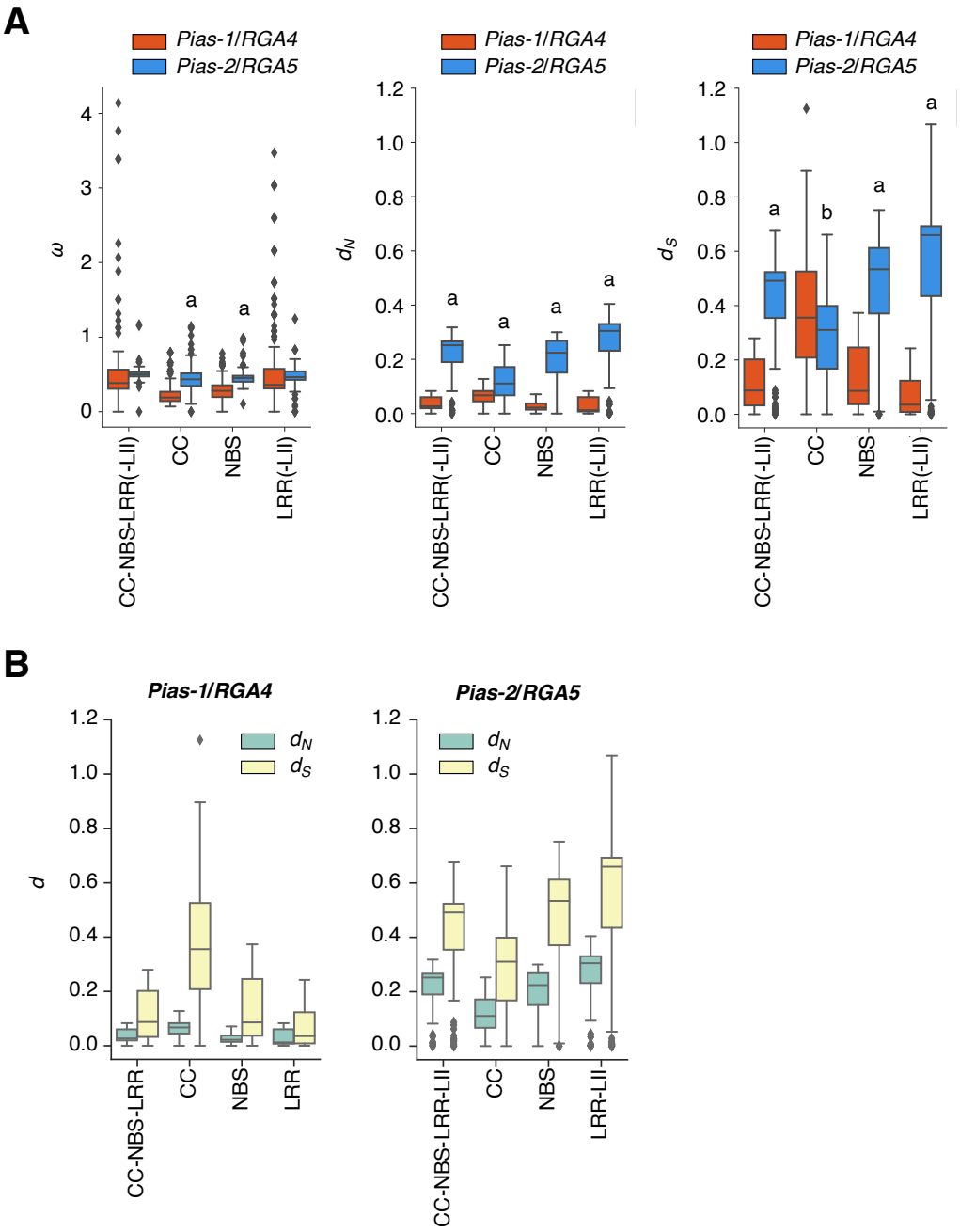
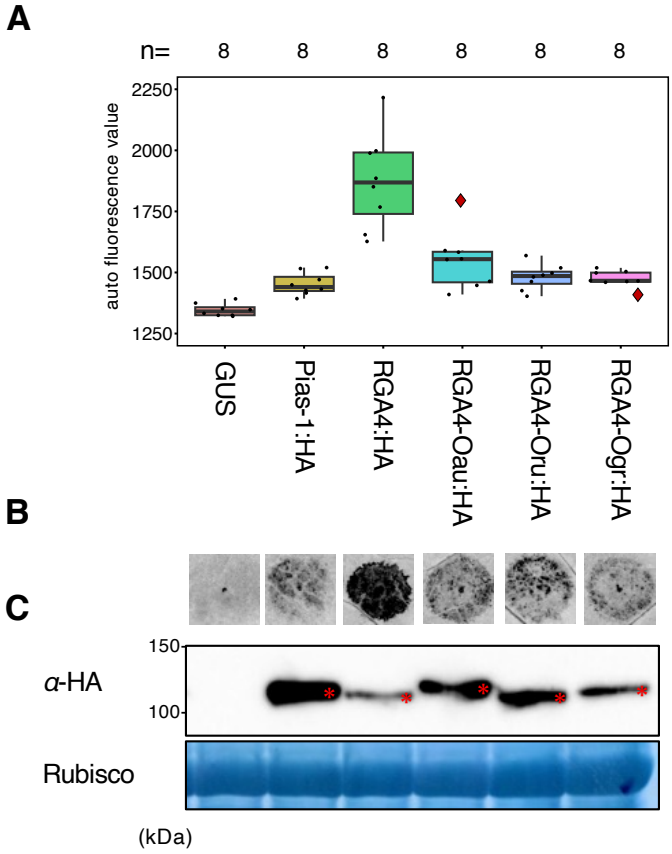


Figure 3. Contrasting evolutionary patterns of the helper and sensor NLRs of the *Pias/Pia* locus. (A) Phylogenetic tree of the *Pias-1/RGA4* helper NLR gene (left) and *Pias-2/RGA5* sensor NLR (right) gene based on the full-length amino acid sequence of Pias-1/RGA4 and the sequence in the region CC to LII for Pias-2/RGA5. *Pias-2/RGA5* sensor NLRs form two major clades (C1 and C2). The numbers indicate bootstrap values. (B) Nucleotide diversity (π) of the CC, NBS and LRR(-LII) domains of *Pias-1/RGA4* helper NLR gene and *Pias-2/RGA5* sensor NLR gene in 22 *Oryza* samples. (C) Tajima's *D* of the CC, NBS and LRR(-LII) domains of *Pias-1/RGA4* helper NLR gene and *Pias-2/RGA5* sensor NLR gene in 22 *Oryza* samples. (D) Pairwise d_N and d_S values of CC-NBS-LRR(-LII), CC, NBS and LRR domains of the *Pias-1/RGA4* helper and *Pias-2/RGA5* sensor NLR gene in 22 *Oryza* samples.

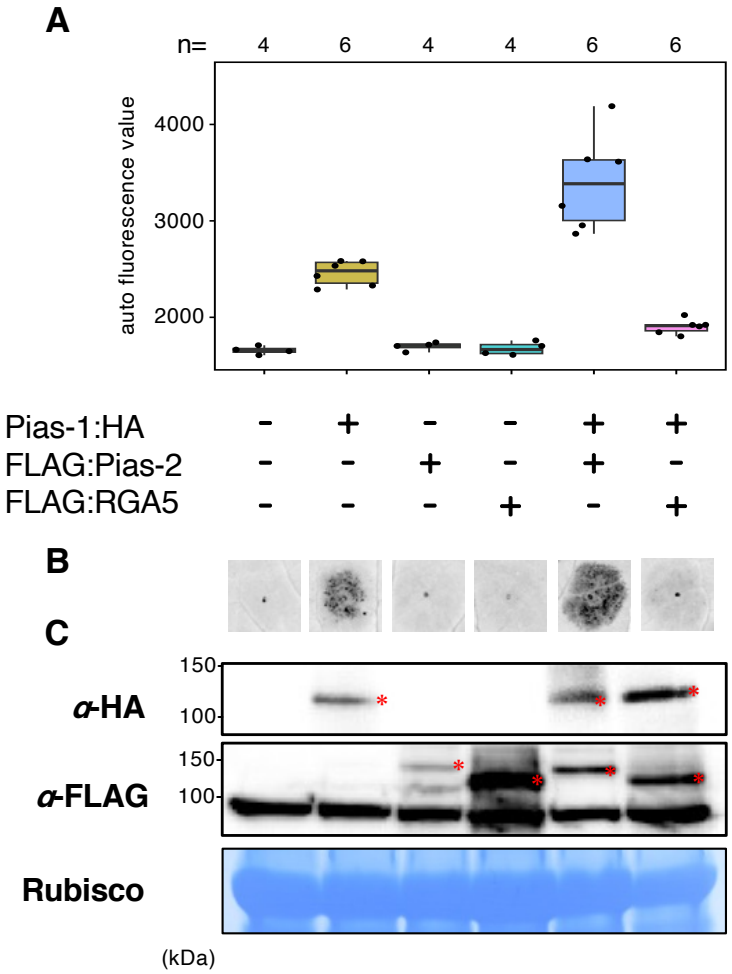


Supple. Fig. 13 (A) Comparison of ω, d_N, and d_S between *Pias-1/RGA4* and *Pias-2/RGA5*. The 'a' indicates that *Pias-2/RGA5* is significantly larger than *Pias-1/RGA4* by two-sided Welch's t-test ($p < 0.0001$). The 'b' indicates that *Pias-2/RGA5* is significantly smaller than *Pias-1/RGA4* by two-sided Welch's t-test ($p < 0.0001$). (B) Comparison between d_N, and d_S in *Pias-1/RGA4* and *Pias-2/RGA5.*



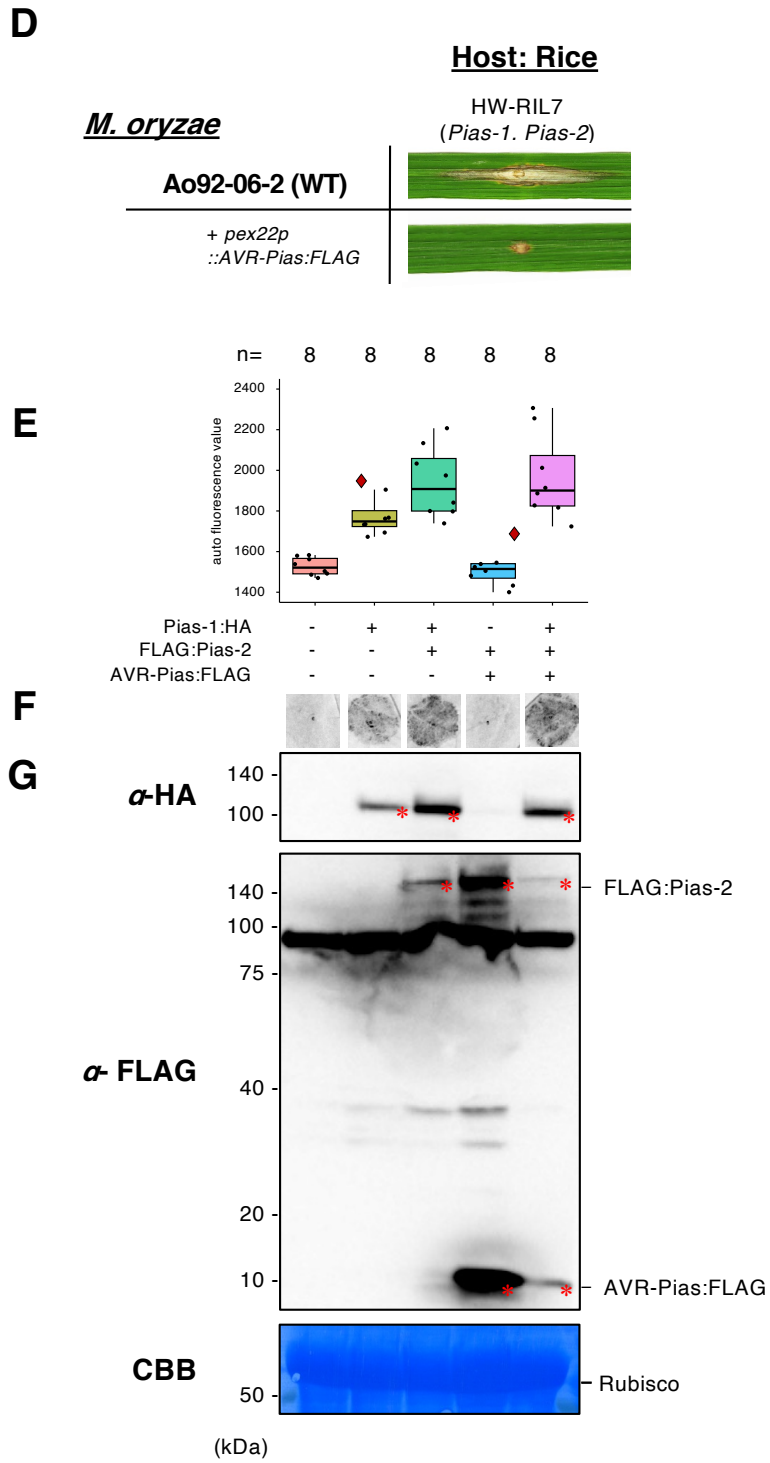
Suppl. Fig. 14 HR-like cell death after overexpression of helper NLRs of the RGA4/Pias-1 lineage in *N. benthamiana.*

(A) Boxplots of autofluorescence values after transient expression of RGA4/Pias-1 homologs. GUS vector was included as a control. The number of spots inoculated with *A. tumefaciens* are indicated at the top of the boxplot. Box bounds represent the 25th and 75th percentile, center line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). (B) Representative image of Pias-1/RGA4-meditated HR in *N. benthamiana*. (C) Immunoblot analysis of Pias-1:HA, RGA4:HA, RGA4-Oau:HA (cloned from *O. australiensis* accession W0008), RGA4-Oru:HA (cloned from *O. rufipogon* accession W1943), and RGA4-Ogr:HA (cloned from *O. granulata* accession W0067B) proteins detected by anti-HA (*α*-HA) antibody. Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of molecular size marker are indicated on the left (kDa)



Suppl. Fig. 15 HR-like cell death after overexpression of Pias-1:HA helper and Pias-1/RGA5 sensors in *N. benthamiana*.

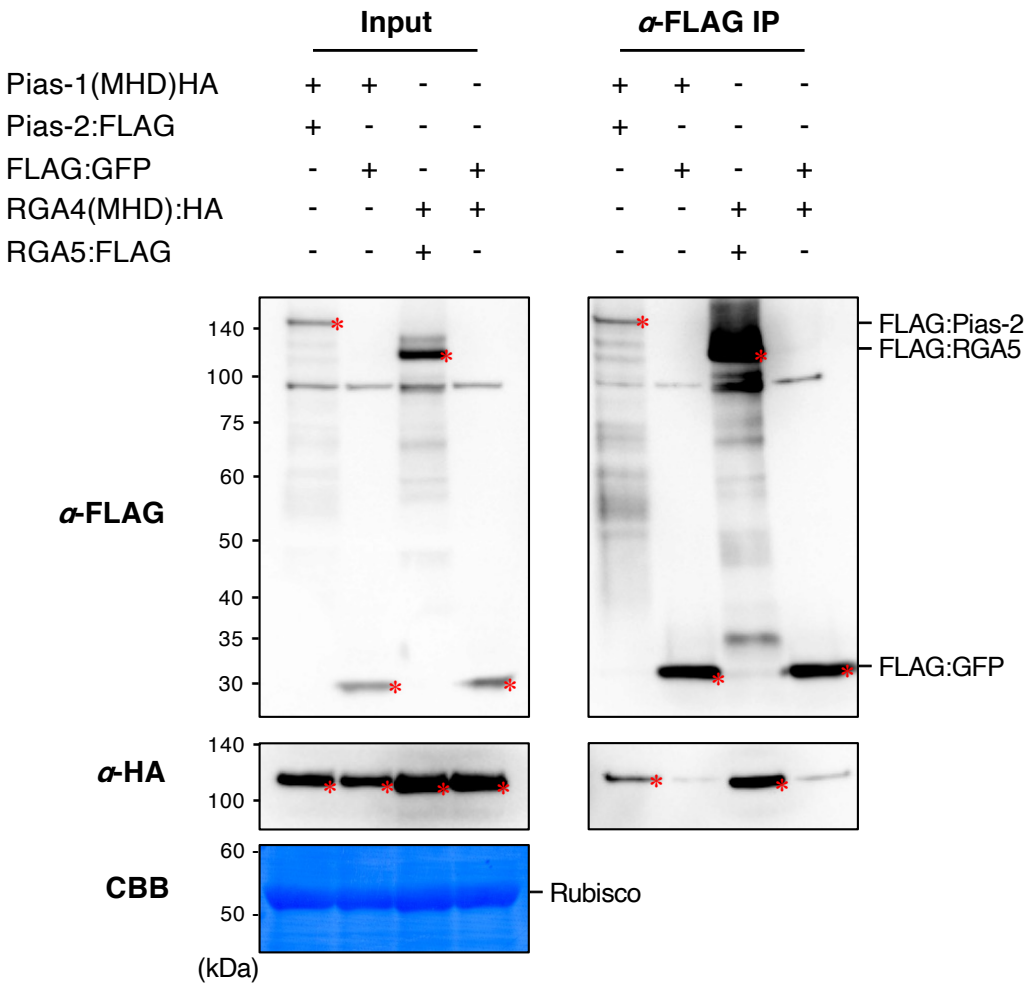
(A) Boxplots of autofluorescence values after transient expression of Pias-1:HA, FLAG:Pias-2, and FLAG:RGA5 separately or in combination in *N. benthamiana*. The number of spots inoculated with *A. tumefaciens* are indicated at the top of the boxplot. Box bounds represent the 25th and 75th percentile, center line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). (B) Representative image of HR after transient expression of Pias-1:HA, FLAG:Pias-2, and FLAG:RGA5 separately or in combination in *N. benthamiana*. (C) Immunoblot analysis of Pias-1:HA protein detected by anti-HA (α-HA) antibody and FLAG:Pias-2 and FLAG:RGA5 proteins detected by anti-FLAG antibody. Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of molecular size marker are indicated on the left (kDa).



Suppl. Fig. 15 (continued). HR-like cell death after overexpression of Pias-1 helper, Pias-2 sensor, and AVR-Pias in *N. benthamiana*.

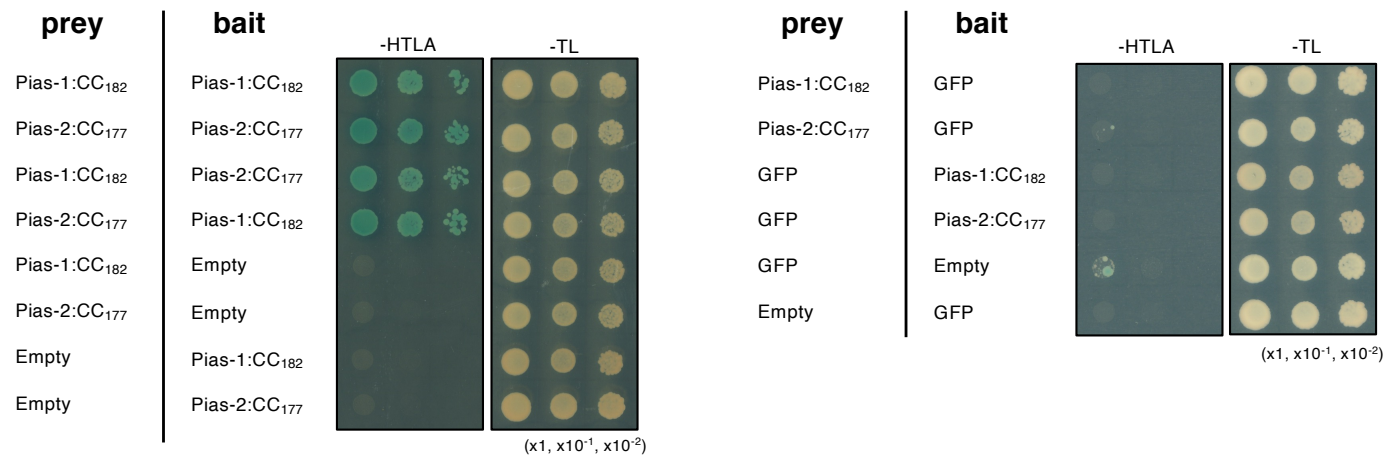
(D) The rice line HW-RIL7 with *Pias* recognizes the *M. oryzae* Ao-92-06-2 isolate with *AVR-Pias:FLAG* (Ao92-06-2+*pex22p*:*AVR-Pias:FLAG*) and shows resistance to this isolate. (E) Boxplots of autofluorescence values after transient expression of Pias-1:HA, FLAG:Pias-2, and AVR-Pias:FLAG separately or in combination in *N. benthamiana*. The number of spots inoculated with *A. tumefaciens* are indicated above the boxplot Box bounds represent the 25th and 75th percentile, center line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). (F) Representative image of HR after transient expression of Pias-1:HA, FLAG:Pias-2, and FLAG:RGA5 separately or in combination in *N. benthamiana*. (G) Immunoblot analysis of Pias-1:HA protein detected by anti-HA (α-HA) antibody and FLAG:Pias-2 and AVR-Pias:FLAG proteins detected by anti-FLAG antibody. Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of the molecular size marker are indicated on the left (kDa).

A

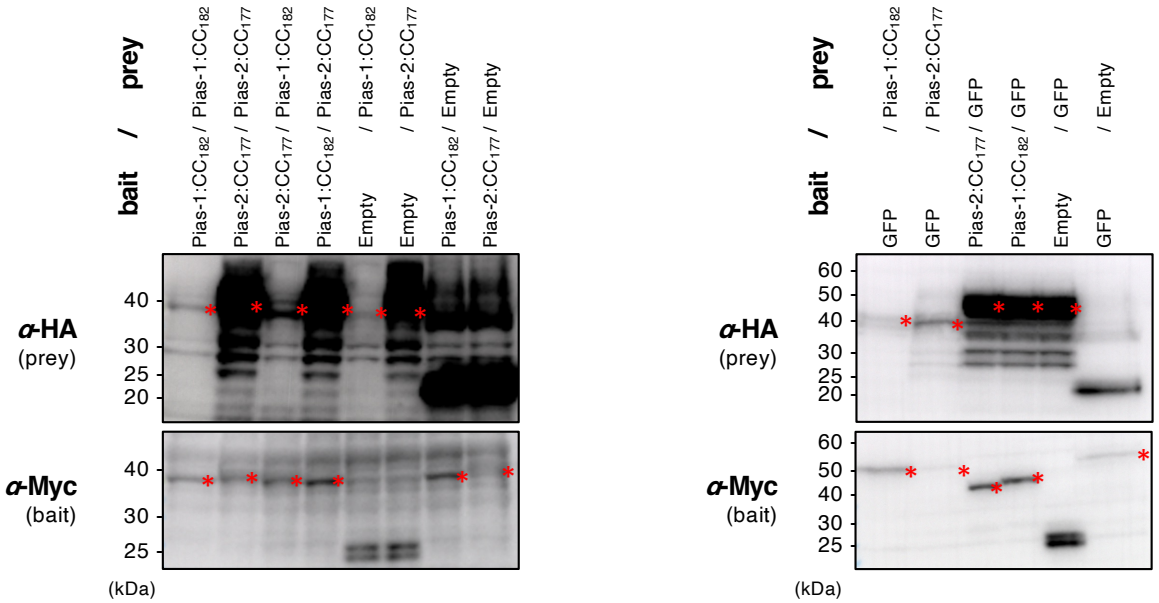


Suppl. Fig. 16. Co-immunoprecipitation shows that Pias-1 interacts with Pias-2.
(A) Co-immunoprecipitation (Co-IP) of Pias-1 (MHD mutant):HA with FLAG:Pias-2 or FLAG:GFP (negative control), as well as Co-IP of RGA4 (MHD mutant):HA with FLAG:RGA5 or FLAG:GFP (negative control) were performed. Pias-1 (MHD mutant):HA and FLAG:Pias-2 or FLAG:GFP were transiently co-expressed in *N. benthamiana.* Similarly, RGA4 (MHD mutant):HA and FLAG:RGA5 or FLAG:GFP were transiently co-expressed in *N. benthamiana*. Instead of the wild-type Pias-1 and RGA4, their MHD mutants (TYG to MHD in ARC2 subdomain as described in Cesari et al. 2014) were used to avoid HR-like cell death that reduces protein accumulation. We judge the weak bands detected by anti-HA antibody after Co-IP in Pias-1(MHD):HA+FLAG:GFP (lane 2) and RGA4(MHD):HA+FLAG:GFP (lane 4) are non-specific. Bound fractions were analyzed by immunoblotting using anti-FLAG and anti-HA antibodies. Coomassie blue staining of the Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of the molecular size marker are indicated on the left (kDa). We obtained similar results in three independent experiments.

B



C



Suppl. Fig. 16 (continued). Yeast two-hybrid assays indicate that the Pias-1 CC domain and Pias-2 CC domain homo- and heterodimerize.

(B) The Pias-1 and Pias-2 CC domains form homo- and heterocomplexes. A dilution series of yeast cells expressing a GAL4-AD and GAL4-BD fusion of Pias-1 1-182 (Pias-1:CC_{1-182}) and/or Pias-2 1-177 (Pias-2:CC_{1-177}) on selective media lacking His, Trp, Leu and Ade (–HTLA) and non-selective media lacking Trp and Leu (–TL) with 5-bromo-4-chloro-3-indolyl α-D-galactopyranoside (X-a-gal). GFP was used as the negative control. Photographs were taken 5 days after growth. (C). Immunoblot analysis confirms the protein production in the Y2H assay shown in (B). The bait protein was tagged with the Myc epitope and the prey protein with the HA epitope. The protein bands expressed from each vector are marked by red asterisks. The positions of the molecular size marker are indicated on the right (kDa). We obtained similar results in three independent experiments. We attempted Y2H assay using the full-length Pias-1 and Pias-2 constructs. However, we could not detect interactions between the full-length Pias-1 and Pias-2.

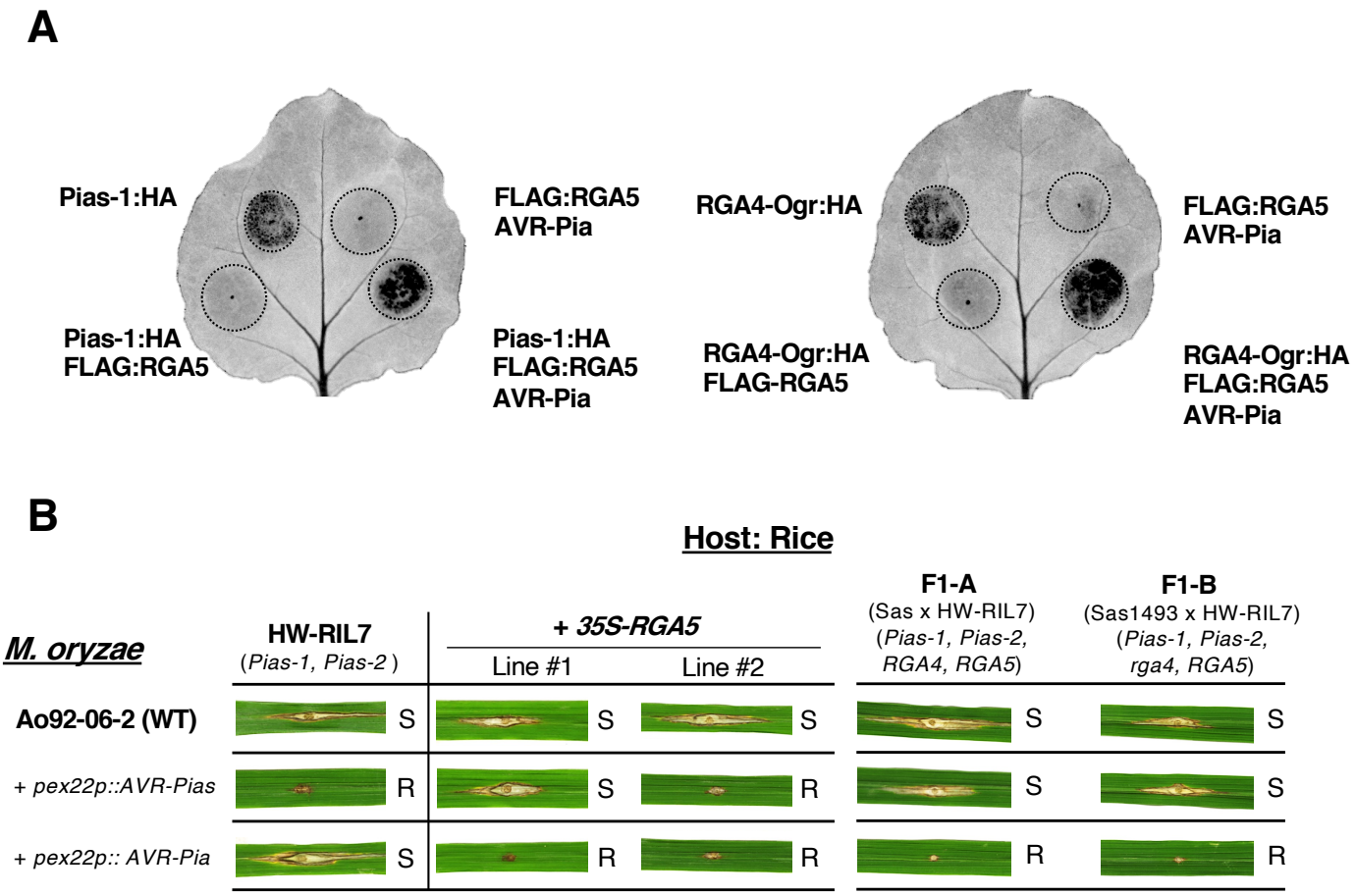
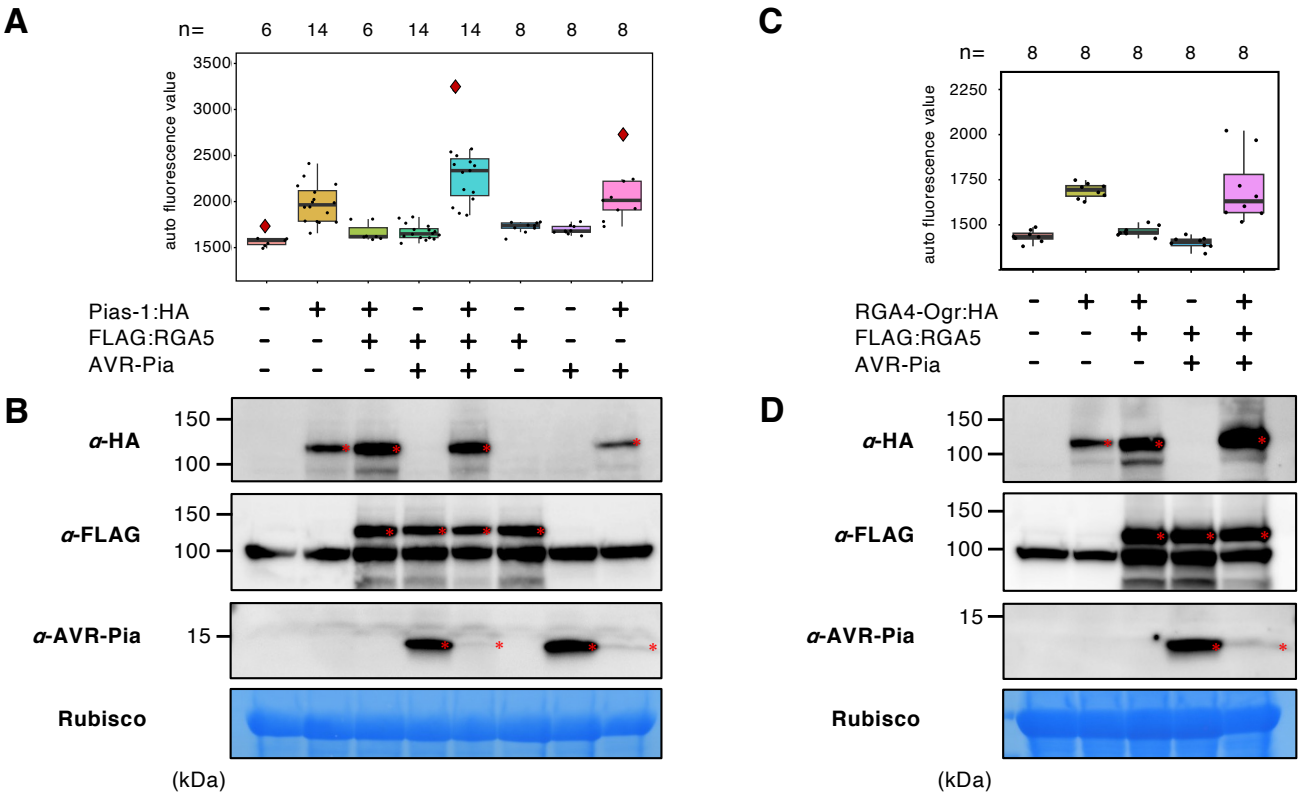
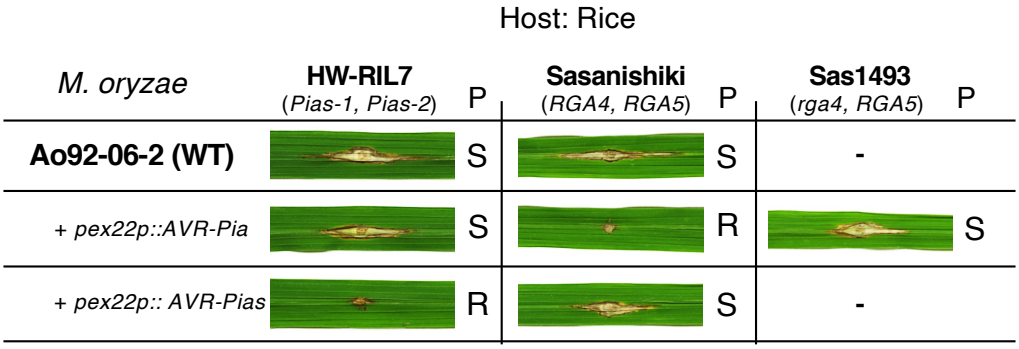


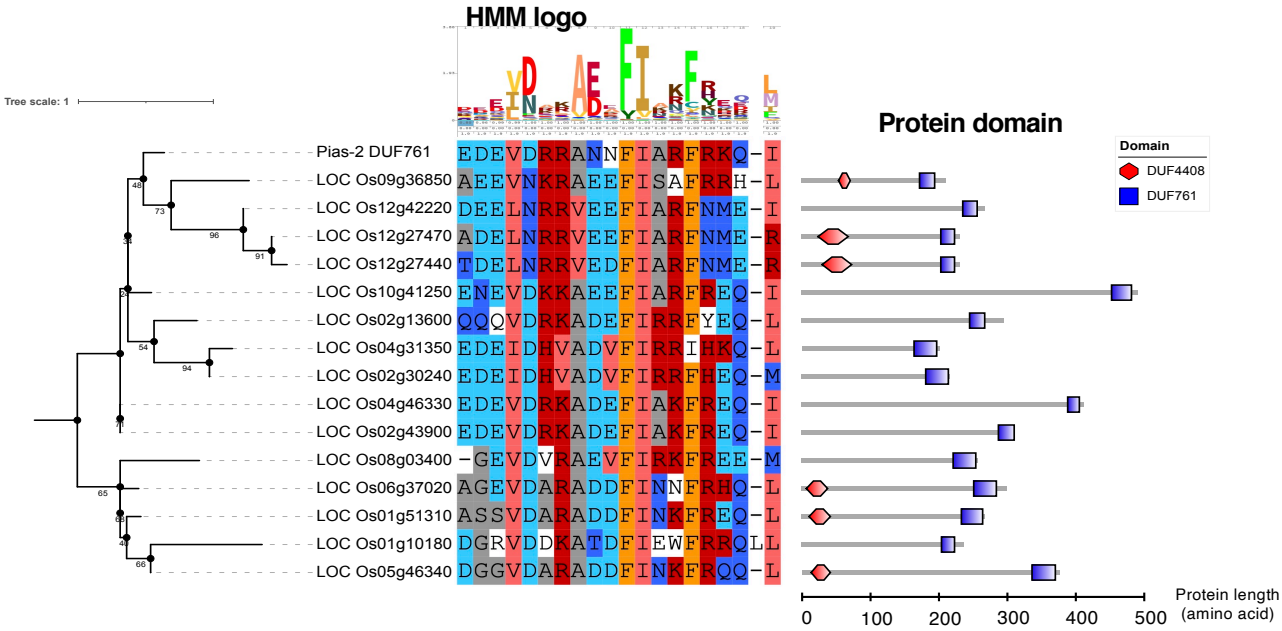
Figure 4. The NLR helper Pias-1 is functionally conserved.
(A) Representative images of *N. benthamiana* leaves after agroinfiltration with Pias-1:HA, Pias-1:HA/FLAG:RGA5, FLAG:RGA5/AVR-Pia, and Pias-1:HA/FLAG:RGA5/AVR-Pia (left) and RGA4-Ogr:HA derived from *O. granulata*, RGA4-Ogr:HA/FLAG:RGA5, FLAG:RGA5/AVR-Pia, and RGA4-Ogr:HA/FLAG:RGA5/AVR-Pia (right). Autofluorescence under UV light is shown. (B) Pias-1 cooperates with RGA5 to recognize AVR-Pia and induces resistance in rice. The rice line HW-RIL7 with *Pias* (*Pias-1* and *Pias-2*) recognizes the Ao-92-06-2 strain with *AVR-Pias* (Ao92-06-2+*pex22p*:*AVR-Pias*) and induces resistance. However, HW-RIL7 cannot recognize the Ao02-06-2 strain with *AVR-Pia* (Ao92-06-2+*pex22p*:*AVR-Pia*). Two lines (Line #1 and #2) contain the *35S-RGA5* transgene in the HW-RIL7 background. F1-A is a progeny derived from a cross between Sasanishiki with *Pia* (*RGA4* and *RGA5*) and HW-RIL7. F1-B is a progeny derived from a cross between a Sasanishiki mutant (Sas1493) with *pia* (*rga4* and *RGA5*) and HW-RIL7.



Suppl. Fig. 17. HR-like cell death caused by Pias1 and RGA4-Ogr expression is suppressed by RGA5 expression, and additional AVR-Pia expression induces cell death.
(A) Boxplots of autofluorescence values after transient expression of Pias-1:HA, FLAG:RGA5, and AVR-Pia separately or in combination in *N. benthamiana* leaves. The number of spots inoculated with *A. tumefaciens* are indicated at the top of the boxplot. Box bounds represent the 25th and 75th percentile, center line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). (B) Immunoblot analysis of Pias-1:HA protein detected by anti-HA (*α*-HA) antibody, FLAG:RGA5 protein detected by anti-FLAG antibody, and AVR-Pia detected by anti-AVR-Pia antibody. Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of molecular size markers are indicated on the left (kDa). (C) Boxplots of autofluorescence values after transient expression of RGA4-Ogr:HA, FLAG:RGA5, and AVR-Pia separately or in combination in *N. benthamiana* leaves.The number of spots inoculated with *A. tumefaciens* are indicated at the top of the boxplot. Box bounds represent the 25th and 75th percentile, center line represents the median, and whiskers indicate the range of the maximum or minimum data within a 1.5 x interquartile range (IQR). D) Immunoblot analysis of RGA4-Ogr:HA protein detected by anti-HA (*α*-HA) antibody, FLAG:RGA5 protein detected by anti-FLAG (*α*-FLAG) antibody, and AVR-Pia detected by anti-AVR-Pia (*α*-AVR-Pia) antibody. Coomassie blue staining of Rubisco small subunit shows equal protein loading. The protein bands expressed from the constructs are marked by red asterisks. The positions of molecular size markers are indicated on the left (kDa).



Suppl. Fig. 18. Results of inoculation of HW-RIL7, Sasanishiki and Sas1493 rice plants with *M. oryzae* isolates Ao92-06-2, Ao92-06-2 + *pex22p::AVR-Pia* and Ao92-06-2 + *pex22p::AVR-Pias*.



Suppl. Fig. 19. The DUF761-containing gene family in rice.
A BLASTP search was performed using the amino acid sequence of the DUF761 domain of Pias-2 as a query. Fifteen DUF761 domain-containing genes were retrieved using a cutoff e-value < 10. The middle panel shows the amino acid sequence alignment and the right panel shows the domain structure of each protein.

Suppl. Table 11. List of NLRs and AVR used in the cell death assay.
OD_{600} of *Agrobacterium tumefaciens* used for agroinfiltration to *N. benthamiana* leaves is shown.

Protein name	Epitope tag	Density (OD_{600})
Pias-1	C-terminal HA	0.2
RGA4	C-terminal HA	0.2
RGA4-Oru	C-terminal HA	0.2
RGA4-Oau	C-terminal HA	0.2
RGA4-Ogr	C-terminal HA	0.2
Pias-2	N-terminal FLAG	0.4
RGA5	N-terminal FLAG	0.4
AVR-Pia	Untagged	0.3
AVR-Pias	C-terminal FLAG	0.3