Supplementary Information for

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

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

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

This PDF file includes:

- Figures S1 to S19
- Tables S1 to S7
- Legends for Datasets S1 to S6
- SI References

Other supplementary materials for this manuscript include the following:

- Datasets S1 to S6
SI Materials and Methods

**Immunoblot analysis** Proteins were expressed in *Nicotiana benthamiana* leaves and extracted from leaf tissue (approximately 100 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 extracts were subjected to SDS-PAGE, followed by immunoblotting using anti-FLAG-HRP antibody, anti-HA-HRP antibody, or anti-AVR-Pia antibody. Anti-mouse IgG-HRP was used as a secondary antibody following incubation with anti-AVR-Pia antibody. Antibody-antigen complexes were detected using a luminescent image analyzer (ImageQuant LAS-4000) (Cytiva, Tokyo, Japan). Equal loading of proteins in the PAGE gel was confirmed by Coomassie blue staining.

**Yeast two-hybrid assay** To examine the protein–protein interactions between Pias-1:CC182 and Pias-2:CC177, a yeast two-hybrid assay was performed as described previously (1). The CC domain of Pias-1 (Pias-1:CC182) was amplified by PCR, digested with EcoRI and BamHI, and cloned into pGADT7 (prey) or pGBKT7 (bait) vector (Clontech, Madison, WI, USA) that had been linearized by digestion with EcoRI and BamHI by in-fusion cloning. The CC domain of Pias-2 (Pias-2:CC177) was amplified by PCR, digested with *Sfi*I, and cloned into pGADT7 or pGBKT7 vector that had been linearized by digestion with *Sfi*I. GFP (as negative control) was amplified by PCR and cloned into pGADT7 or pGBKT7 vectors that had been linearized by digestion with EcoRI and BamHI by in-fusion cloning. The various combinations of bait and prey vectors were transformed into yeast strain AH109 using the PEG/LiAc method. To detect protein–protein interactions, a 10-fold dilution series (×1, ×10⁻¹, ×10⁻²) of yeast cells (×1: OD₆₀₀ = 1.0) was spotted onto basal medium lacking Trp, Leu, Ade, and His (−HTLA) but containing 5-bromo-4-chloro-3-indolyl α-D-galactopyranoside (Clontech, Madison, WI, USA). As a control, yeast growth on basal medium lacking Trp and Leu (−TL) was also checked. To assess the protein accumulation in yeast cells, each transformant was propagated in liquid basal medium lacking Trp and Leu with gentle shaking at 30°C overnight. Yeast cells from 10 mL medium were collected, and 100 mg of yeast cells was treated with 400 µL of 0.3 N NaOH for 15 min at room temperature. The resulting yeast extracts were used for immunoblot analysis using anti-Myc (MBL, Woburn, MA, United States) for bait proteins and anti-HA (Roche, Switzerland) for prey proteins. The primers used to generate the constructs are listed in Dataset S5.

**Co-immunoprecipitation assay** Expression constructs for two types of helper-NLRs, Pias-1 (MHD) and RGA4 (MHD) (pCambial300S-“helper-NLR”:HA), were generated by PCR and cloned into the binary vector pCambial300S (http://www.cambia.org) that had been linearized by digestion with *Pst*I and *Spe*I by in-fusion cloning. The resulting vectors were introduced into *A. tumefaciens* (strain GV3101). Pias-1(MHD):HA and FLAG:Pias-2 or FLAG:GFP were transiently co-expressed in *N. benthamiana*. Similarly, RGA4 (MHD):HA and FLAG:RGA5 or FLAG:GFP were transiently co-expressed in *N. benthamiana*. The co-expressed proteins were extracted with 50 mM sodium phosphate buffer containing 10 % (v/v) glycerol and 150 mM NaCl. Extracted proteins were incubated with FLAG-agarose beads in the presence of protease inhibitor, 10 mM DTT, and 2% (w/v) blocking reagent (Cytiva, Tokyo, Japan) at 4°C for 1 h. The beads were washed with the same buffer four times, and then bound proteins were extracted with the same buffer containing 1% (w/v) SDS by boiling. Bound fractions were analyzed by immunoblotting using anti-FLAG and anti-HA antibodies. The primers used to generate constructs are listed in Dataset S5.
Fig. S1. The RaIDeN 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” using *DISCOVAR* software. (D) Gene prediction of WRC17, the resistant parent, using RNA-seq data. Reference genome sequence of “WRC17” from STEP3

RNA-seq based Gene prediction

- Using “Hiul7” aligner and assembly of the transcript was performed by “StringTie”

| Contig | Expressed genes | Gene A | Gene B | Gene C |
|--------|-----------------|--------|--------|--------|
|        |                 |        |        |        |

(E) Mapping of short reads of susceptible parent and RILs to the predicted genes of WRC17.

Candidate genes

14 839

No. of RILs

Lesion area (mm²)

0 10 20 30 40 50

NGS reads

RNA-seq

Fragmentation & sequencing reads

R-parent “WRC17”

S parental “Hitomebore”

#22 #25 #34 #53

DNA extract

RNA extract

Hitomebore (S)

Susceptible (S)-RILs

WRC17 (Keiboba)

#22 #25 #34 #37 #39 #53

Reference genome sequence of “WRC17” from STEP3

aligned RNA-seq read

Expressed genes

Lesion A

Lesion B

Lesion C

Contig

Candidate

genes

PIW17-1/PIW17-2

PIW17-1/piw17-2

piw17-1/PIW17-2

piw17-1/piw17-2

Lesion area (mm²)

0 10 20 30 40 50

No. of RILs

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)

0 10 20 30 40 50

Lesion area (mm²)
**Fig. S2. Mapping of NLRs 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 Dataset S5). CNL-04 and CNL-05, CNL-06 and CNL-07, and CNL-08, NL-02, 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.
Fig. S2 (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.
Fig. S3. 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.
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.
Fig. S4. 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.
Fig. S5. 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 (2).
Fig. S5 (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 (3).
**Fig. S6. 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 (SI Appendix, Table S4 and S5). Next, the isolate 2012-1 was used to inoculate barley (*Hordeum vulgare*) cultivar ‘Nigate’, which is highly susceptible to *M. oryzae* (4), and the infected leaves were subjected to RNA-seq (5), 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 (SI Appendix, Table S4). 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*. 

---

(Raw text continues as in the image)
Fig. S7. 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.
Fig. S8. 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.
Fig. S9. 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.
Fig. S9 (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 (2).
Fig. S9 (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 (2).
Fig. S10. Positions of the junction region (this study) and the CID motif (6).

Amino acid sequence alignment of RGA5, Pias-2 and LOC_Os12g18360 and the positions of the junction region (this study) and the CID motif as described by Bailey et al. (6). The aqua and orange boxes indicate the junction region and CID motif, respectively. The red lines indicate LRR in the junction region.
Fig. S11. The *O. punctata* W1582 junction -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 junction-ID and the W1582 junction-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 junction region, respectively.
Fig. S11 (Continued).
Fig. S12. The junction + 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).
Fig. S13.
(A) Comparison of $\omega$, $d_N$, 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.
Fig. S14. 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-mediated 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).
Fig. S15. 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).
**Fig. S15. (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).
Fig. S16. 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

Fig. S16 (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:CC1-182) and/or Pias-2 1-177 (Pias-2:CC1-177) 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.
Fig. S17. 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).
Fig. S18. 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*. 

| M. oryzae          | Host: Rice |
|--------------------|-----------|
|                    | HW-RIL7 (Pias-1, Pias-2) | Sasanishiki (RGA4, RGA5) | Sas1493 (rga4, RGA5) |
| Ao92-06-2 (WT)     | S         | S                     | -                     |
| + *pex22p::AVR-Pia*| S         | R                     | S                     |
| + *pex22p::AVR-Pias*| R       | S                     | -                     |
Fig. S19. 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.
Table S1. Summary of Illumina sequencing results used for RaIDeN methods

| Line | Whole genome sequencing | Purpose |
|------|-------------------------|---------|
|      | R-parental line         |         |
|      | WRC17                   |         |
|      | Miseq                   | No      |
|      | Total no. of paired-end (PE) reads after filtering | 36,047,742 |
|      | Max length of short reads | 250 |
|      | Purpose | for De novo assembly |
|      | S-parental line         |         |
|      | Hitomebore              |         |
|      | Hiseq4000               | Yes     |
|      | Total no. of paired-end (PE) reads after filtering | 32,016,561 |
|      | Max length of short reads | 126 |
|      | Purpose | for RaIDeN |
|      | S-RILs                  |         |
|      | #22                     | Hiseq4000 | Yes |
|      | #25                     | Hiseq4000 | Yes |
|      | #34                     | Hiseq4000 | Yes |
|      | #37                     | Hiseq4000 | Yes |
|      | #39                     | Hiseq4000 | Yes |
|      | #53                     | Hiseq4000 | Yes |
|      | RNAseq                  |         |
|      | M. oryzae infected leaf of WRC17 | Nextseq500 | Yes |
|      | Total no. of paired-end (PE) reads after filtering | 39,744,107 |
|      | Max length of short reads | 75 |
|      | Purpose | for RaIDeN |
Table S2. Summary metrics of genome assembly of the resistance rice line WRC17. 
Genome assembly was performed by DISCOVAR software.

| Metric                                           | Value  |
|--------------------------------------------------|--------|
| 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  |
Table S3. 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 (7).

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     |

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
Table S4. 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 No. | 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* assembly                       |
| **S-pathogen for HW-RIL7** |          |                     |                                  |                                                   |                          |                                              |
| Ina87T-56A | r1       | Miseq               | Yes                              | 1,776,902                                         | 125                      | for *RalDeN*                                 |
| Ao92-06-2  | r2       | Miseq               | Yes                              | 2,691,934                                         | 125                      | for *RalDeN*                                 |
| 85-141     | r3       | Miseq               | Yes                              | 3,034,450                                         | 125                      | for *RalDeN*                                 |
| SL91-48D   | r4       | Miseq               | Yes                              | 3,108,538                                         | 125                      | for *RalDeN*                                 |
| 2403-1     | r5       | Miseq               | Yes                              | 3,388,670                                         | 125                      | for *RalDeN*                                 |
| H98-315-1  | r6       | Miseq               | Yes                              | 2,933,673                                         | 125                      | for *RalDeN*                                 |
| 0423-1     | r7       | Miseq               | Yes                              | 2,996,542                                         | 125                      | for *RalDeN*                                 |
| Ina85-182  | r8       | Miseq               | Yes                              | 3,148,657                                         | 125                      | for *RalDeN*                                 |
| **RNAseq** |          |                     |                                  |                                                   |                          |                                              |
| *M. oryzae* isolate 2012-1 inoculated to the leaf of *Hordeum vulgare* var. Nigrate |          |                     |                                                  |                                                   |                          |                                              |
| 24H after inoculation | Nextseq500      | Yes                         | 15,658,123                             | 75                                                 |                          | for *RalDeN* and expression analysis         |
| 48H after inoculation | Nextseq500      | Yes                         | 15,770,613                             | 75                                                 |                          | for *RalDeN* and expression analysis         |
Table S5. 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        | 41.0     |
| assembled genome (Mb)    |          |
Table S6. A survey of 51 A-genome *Oryza* accessions with DUF761 ID revealed that the *O. punctata* junction-ID-like downstream sequence is widely conserved.

BLASTN search was performed on each genome sequence using *O. punctata* W1582 junction-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 “junction”-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              |
Table S7. 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               |
| P19          | -            | 0.1                  |
Dataset S1. Summary of 853 candidate genes for the rice blast resistance of WRC17 as selected by RaIDeN pipeline.

Dataset S2. Summary of whole genome sequence data used in this study.

Dataset S3. Summary of Pia/Pias orthologs.

Dataset S4. Prediction of Integrated Domains (IDs) using the whole genome sequences of Oryza accessions.

Dataset S5. List of primers used in this study

Dataset S6. The DNA-seq and RNA-seq data from this study

SI References

1. H. Kanzaki et al., Arms race co-evolution of Magnaporthe oryzae AVR-Pik and rice Pik genes driven by their physical interactions. Plant J. 72, 894-907 (2012).
2. J. T. Robinson et al., Integrative genomics viewer. Nat. Biotechnol. 29, 24-26. (2011).
3. E. C. Martin et al., LRRpredictor-A New LRR Motif Detection Method for Irregular Motifs of Plant NLR Proteins Using an Ensemble of Classifiers. Genes (Basel). 11, 286. (2020).
4. G. S. Hyon et al., Characterization of interactions between barley and various host-specific subgroups of Magnaporthe oryzae and M. grisea. J. Gen. Plant Pathol. 78, 237-246. (2012).
5. M. Shimizu et al, RNA-Seq of in planta-expressed Magnaporthe oryzae genes identifies MoSVP as a highly expressed gene required for pathogenicity at the initial stage of infection. Mol. Plant Pathol. 20, 1682-1695 (2019).
6. P. C. Bailey et al., Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. Genome Biol. 19, 23 (2018).
7. P. Li et al., RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. BMC Genomics. 17, 852 (2016).
8. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. Nat. Methods. 12, 357-360. (2015).
9. M. Pertea M et al., StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290-295. (2015).
10. Q. Zhao et al., Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. Nat. Genet. 50, 278-284 (2018).
11. H. Du et al., Sequencing and de novo assembly of a near complete indica rice genome. Nat. Commun. 8, 15324 (2017).
12. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. Nature. 436, 793-800 (2005).
13. Z. Wu et al., De novo genome assembly of Oryza granulata reveals rapid genome expansion and adaptive evolution. Commun. Biol. 1, 84 (2018).
14. A. V. Zimin et al., Hybrid assembly of the large and highly repetitive genome of Aegilops tauschii, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. Genome Res. 27, 787-792 (2017).
15. M. Mascher et al., A chromosome conformation capture ordered sequence of the barley genome. Nature. 544, 427-433 (2017).
16. J. L. Bennetzen et al., Reference genome sequence of the model plant Setaria. Nat. Biotechnol. 30, 555-561 (2012).