nature plants

Brief Communication

https://doi.org/10.1038/s41477-023-01357-5

A wheat kinase and immune receptor form host-specificity barriers against the blast fungus

Received: 23 November 2021

Accepted: 20 January 2023

Published online: 16 February 2023

Check for updates

Sanu Arora[1,11], Andrew Steed[1,11], Rachel Goddard[1,7,11], Kumar Gaurav[1,8,11], Tom O'Hara[1,11], Adam Schoen[2], Nidhi Rawat[2], Ahmed F. Elkot[3], Andrey V. Korolev[1], Catherine Chinoy[1], Martha H. Nicholson[1], Soichiro Asuke[4], Rea Antoniou-Kourounioti[1,9], Burkhard Steuernagel[1], Guotai Yu[1,10], Rajani Awal[1], Macarena Forner-Martínez[1], Luzie Wingen[1], Erin Baggs[5], Jonathan Clarke[1], Diane G. O. Saunders[1], Ksenia V. Krasileva[5], Yukio Tosa[4], Jonathan D. G. Jones[6], Vijay K. Tiwari[2], Brande B. H. Wulff[1,10] ✉ & Paul Nicholson[1] ✉

Since emerging in Brazil in 1985, wheat blast has spread throughout South America and recently appeared in Bangladesh and Zambia. Here we show that two wheat resistance genes, *Rwt3* and *Rwt4*, acting as host-specificity barriers against non-*Triticum* blast pathotypes encode a nucleotide-binding leucine-rich repeat immune receptor and a tandem kinase, respectively. Molecular isolation of these genes will enable study of the molecular interaction between pathogen effector and host resistance genes.

Wheat blast, caused by *Pyricularia oryzae* (syn. *Magnaporthe oryzae*) pathotype *Triticum* was first identified in Brazil in 1985 (ref. [1]). The pathogen subsequently spread to cause epidemics in other regions of Brazil and neighbouring countries, including Bolivia and Paraguay[2]. Outbreaks of wheat blast occurred in Bangladesh in 2016, and the disease was reported from Zambia in 2018 (refs. [3,4]). Wheat blast is now considered to pose a threat to global wheat production[5], and discovery and deployment of resistance genes against this pathogen are critical to mitigate its threat.

While *P. oryzae* exhibits a high level of host specificity, *Triticum* pathotypes are closely related to *Lolium* pathotypes[6]. Two pathogen genes, *PWT3* and *PWT4*, condition avirulence of different isolates of *P. oryzae* pathotype *Lolium* on wheat (*Triticum aestivum*).

The resistance genes *Rwt3* and *Rwt4* in wheat recognize respectively the *PWT3* and *PWT4* avirulence gene products to prevent infection. It was originally proposed that the epidemics in Brazil occurred due to the widespread cultivation of varieties lacking *Rwt3* that are susceptible to *Lolium* pathotypes[6]. *Lolium* pathotypes have also been associated with the occurrence of wheat blast in the United States[7,8]. Recent studies suggest that both *Triticum* and *Lolium* pathotypes emerged as part of a multi-hybrid swarm in which key host specificity determinants were re-assorted[9].

In this Brief Communication, to identify candidates for *Rwt3* and *Rwt4*, we used a Triticeae bait library (Supplementary Table 1 and Additional File 1) to capture and sequence the nucleotide-binding site leucine-rich repeat protein (NLR) complements of 320 wheat

[1]John Innes Centre, Norwich Research Park, Norwich, UK. [2]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD, USA. [3]Wheat Research Department, Field Crops Research Institute, Agricultural Research Center, Giza, Egypt. [4]Graduate School of Agricultural Science, Kobe University, Kobe, Japan. [5]Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. [6]The Sainsbury Laboratory, Norwich Research Park, Norwich, UK. [7]Present address: Limagrain UK Ltd, Lincolnshire, UK. [8]Present address: European Molecular Biology Laboratory, European Bioinformatics Institute EMBL-EBI, Hinxton, UK. [9]Present address: School of Molecular Biosciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK. [10]Present address: Center for Desert Agriculture, Biological and Environmental Science and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [11]These authors contributed equally: Sanu Arora, Andrew Steed, Rachel Goddard, Kumar Gaurav and Tom O'Hara. ✉e-mail: brande.wulff@kaust.edu.sa; paul.nicholson@jic.ac.uk

lines including 300 wheat landraces from the A.E. Watkins collection harbouring the genetic diversity existing before intensive breeding (Supplementary Table 2 and Supplementary Fig. 1). We screened seedlings of the panel with Br48, a *Triticum* pathotype strain of *P. oryzae*, transformed with either *PWT3* or *PWT4* (ref. [6]) (Supplementary Table 3 and Supplementary Figs. 2 and 3) and performed *k*-mer-based association genetics using Chinese Spring[10] and Jagger[11] as the reference genomes, respectively. This led to identification of candidate NLR genes, *TraesCS1D02G029900* and *TraesJAG1D03G00423690*, for *PWT3* and *PWT4* recognition (Fig. 1a,b and Supplementary Figs. 4 and 5) on chromosome 1D within the previously defined biparental mapping intervals of *Rwt3* and *Rwt4*, respectively[12,13].

To investigate the allelic variation of these candidate genes, we used BLASTn (ref. [14]) to query their sequences in the NLR assemblies of *Aegilops tauschii*[15], the D-genome progenitor of bread wheat and Watkins wheat landraces. For *Rwt3* NLR candidate, we identified four groups of allelic variants as well as presence/absence variation in *Ae. tauschii*; however, only presence/absence variation was observed in wheat landraces (Fig. 2a, Additional File 2, Supplementary Table 4 and Supplementary Fig. 6). For *Rwt4* NLR candidate, we identified five groups of allelic variants in *Ae. tauschii* out of which only two were observed in wheat landraces (Fig. 2f, Additional File 3, Supplementary Table 4 and Supplementary Fig. 6). Although *Rwt4* NLR candidate is allelic to an NLR in linkage with *Pm24* (ref. [16]), we did not find *Pm24* linked NLR in either *Ae. tauschii* or Watkins landraces, which is consistent with the post-domestication origin hypothesis of *Pm24* in China[16]. *Rwt4* candidate was found only in lineage 2 (L2) of *Ae. tauschii*, while *Rwt3* candidate was found only in lineage 1 (L1), which explains why we could identify only the *Rwt4* candidate, and not the *Rwt3* candidate, by phenotyping and performing association genetics on an NLR gene enrichment-sequenced *Ae. tauschii* L2 panel[15] (Supplementary Table 5 and Supplementary Figs. 6–8). The L1 origin of *Rwt3* is remarkable considering that the L1 signature in wheat is mostly concentrated around a 5 Mb region surrounding the *Rwt3* candidate[17] (Supplementary Fig. 9). This finding suggests that interaction with pathogens may have played a role in wheat evolution.

To functionally validate the *Rwt3* NLR candidate, we screened a TILLING population of Jagger[18] and found three lines each carrying a functional mutation in this gene (Supplementary Fig. 10). One line, M217, is homozygous for a mutation causing a premature stop codon, whereas another, M698, is homozygous for a mis-sense mutation (G241E) predicted to cause functional aberration in the protein (Fig. 2b and Supplementary Table 6). In both the leaf and head assays of these mutants using Br48 + *PWT3*, a loss of the wild-type resistance was observed (Fig. 2c,d). The third line, M1164, is heterozygous for another deleterious mis-sense mutation (E492K) (Fig. 2b and Supplementary Table 6). In both the leaf and head assays of the segregating progeny of M1164 using Br48 + *PWT3*, those homozygous for the mutation were found to be susceptible while the others were resistant (Fig. 2c,d and Supplementary Fig. 11). Virus-induced gene silencing (VIGS) was performed to confirm the role of the *Rwt3* NLR candidate. Resistance to Br48 + *PWT3* was lost in wheat cv. Jagger following silencing with barley stripe mosaic virus (BSMV) expressing a 400 bp *Rwt3* gene fragment (Fig. 2e and Supplementary Fig. 12). The clear loss of function observed in three independently derived TILLING mutants, the co-segregation of the M1164 mutation with susceptibility and the effect of VIGS show that the *Rwt3* NLR candidate is required for resistance to *P. oryzae* expressing the *PWT3* effector.

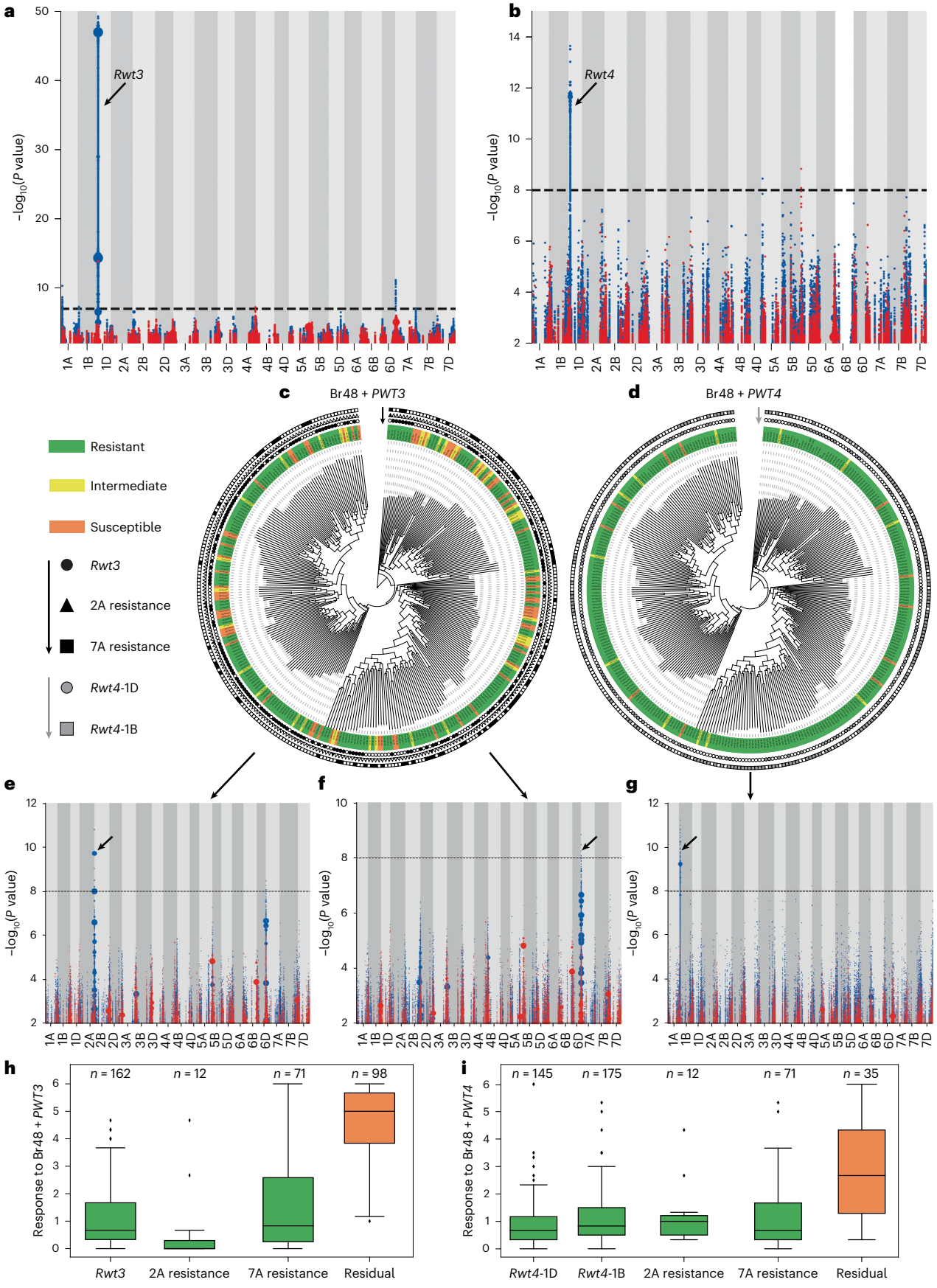
As remarked earlier, the identified *Rwt4* NLR candidate is adjacent to an allele of a wheat tandem kinase (WTK), designated *Pm24*, which confers resistance against powdery mildew[16] (Fig. 2f). Therefore, we tested both the identified *Rwt4* NLR candidate (Supplementary Fig. 5) and the linked *Pm24* allele, *TraesJAG1D03G00423590* (ref. [11]) (Supplementary Fig. 13), as candidates for *Rwt4* using the Cadenza TILLING resource[19]. For the NLR candidate, we tested four lines (two heterozygous and two homozygous) carrying mutations predicted to cause premature stop codons and three additional lines (two heterozygous and one homozygous) carrying mis-sense mutations predicted to have a significant impact on tertiary structure (Supplementary Fig. 5a and Supplementary Table 6). Neither the homozygous nor any progeny of the heterozygous mutants for this candidate showed an increase in susceptibility relative to the wild type Cadenza in either leaf or head assays with Br48 + *PWT4* (Supplementary Fig. 14). For the linked *Pm24* allele, we tested three lines (one homozygous and two heterozygous) carrying mutations that result in premature stop codons (Fig. 2g and Supplementary Table 6). In both the leaf and head assays of the homozygous line M0159 using Br48 + *PWT4*, a clear increase in susceptibility compared with the wild type was observed (Fig. 2h,i). In the leaf and head assays of the segregating progeny of heterozygous mutants (M0971 and M1103) using Br48 + *PWT4*, those homozygous for the mutation were found to be susceptible while all others were resistant (Fig. 2h,i and Supplementary Fig. 15). VIGS was performed to confirm the role of the *Rwt4* WTK candidate. Resistance to Br48 + *PWT4* was lost in wheat cv. Jagger following silencing with a 400 bp gene fragment of *Rwt4* WTK (Fig. 2j and Supplementary Fig. 16). These results show that the linked WTK, and not the identified NLR candidate, is required for resistance to *P. oryzae* expressing the *PWT4* effector. The finding that *WTK* alleles, *Pm24* and *Rwt4*, are involved in resistance to two unrelated fungal pathogens suggests that it may be a broad-spectrum component of disease resistance.

We developed Kompetitive Allele-Specific PCR (KASP) markers for *Rwt3* and *Rwt4* (Supplementary Table 7) and tested them on the core 300 Watkins lines obtaining a validation rate of 97% and 99% based on their corresponding in silico predictions (Supplementary Table 8). Based on

Fig. 1 | Genetic identification of resistances to the blast fungus by *k*-mer-based association mapping on an *R*-gene enriched sequencing panel of wheat landraces. a,**b**, *k*-mers associated with resistance to (Br48 + *PWT3* mapped to Chinese Spring (**a**) and Br48 + *PWT4* mapped to Jagger (**b**). **c**,**d**, *k*-mer-based phylogeny of wheat landraces showing the phenotype of an accession after inoculation with: Br48 + *PWT3* (**c**) and Br48 + *PWT4* (**d**), and the predicted presence of the putative resistances. Phenotype of an accession after inoculation with a blast isolate is indicated by the colour used to highlight the label of that accession, while the presence and absence of allele-specific polymorphisms is indicated by filled symbols with black/grey and white, respectively. **e**,**f**, *k*-mers significantly associated with resistance to Br48 + *PWT3* in the absence of the *Rwt3* candidate gene on chromosome 1D leads to the identification of a resistance on chromosome 2A when mapped to the assembly of wheat cultivar SY Mattis (**e**), and chromosome 7A when mapped to wheat cultivar Jagger (**f**). **g**, *k*-mers significantly associated with resistance to Br48 + *PWT4* in the absence of *Rwt4* candidate gene on chromosome 1D leads to the identification of a resistance on a region of chromosome 1B containing the homoeologue of *Rwt4* when mapped to Jagger. **h**, Box plots showing variation for resistance to Br48 + *PWT3* in Watkins lines predicted to carry *Rwt3*, 2A resistance, 7A resistance and none of them. **i**, Box plots showing variation for resistance to Br48 + *PWT4* in Watkins lines predicted to carry *Rwt4*-1D, *Rwt4*-1B, 2A resistance, 7A resistance and none of them. In the box plots of **h** and **i**, boxes denote the interquartile range with the horizontal bar inside representing the median. Whiskers extend to 1.5 times the interquartile range, values outside of which are considered outliers and shown as individual points. Number of Watkins lines belonging to each class is indicated at the top of the corresponding box plot. In the association plots of **a** and **b** and **e**–**g**: points on the *y* axis depict *k*-mers positively associated with resistance in blue and negatively associated with resistance in red. Point size is proportional to the number of *k*-mers. The association score is defined as the $-\log_{10}$ of the *P* value obtained using the likelihood ratio test for nested models. The threshold of significant association scores is adjusted for multiple comparisons using the Bonferroni approximation. Arrows indicate regions with significant association scores.

https://doi.org/10.1038/s41477-023-01357-5

the in silico predictions, *Rwt3* is present only in 148 of the 192 Watkins lines resistant to Br48 + *PWT3* (Fig. 1c and Supplementary Table 8), while *Rwt4* is present in only 141 of the 271 Watkins lines resistant to Br48 + *PWT4* (Fig. 1d and Supplementary Table 8). This suggests that there are other resistance genes in the Watkins panel recognizing *PWT3*, *PWT4* or additional effectors in Br48. We re-ran GWAS with the leaf assay



https://doi.org/10.1038/s41477-023-01357-5

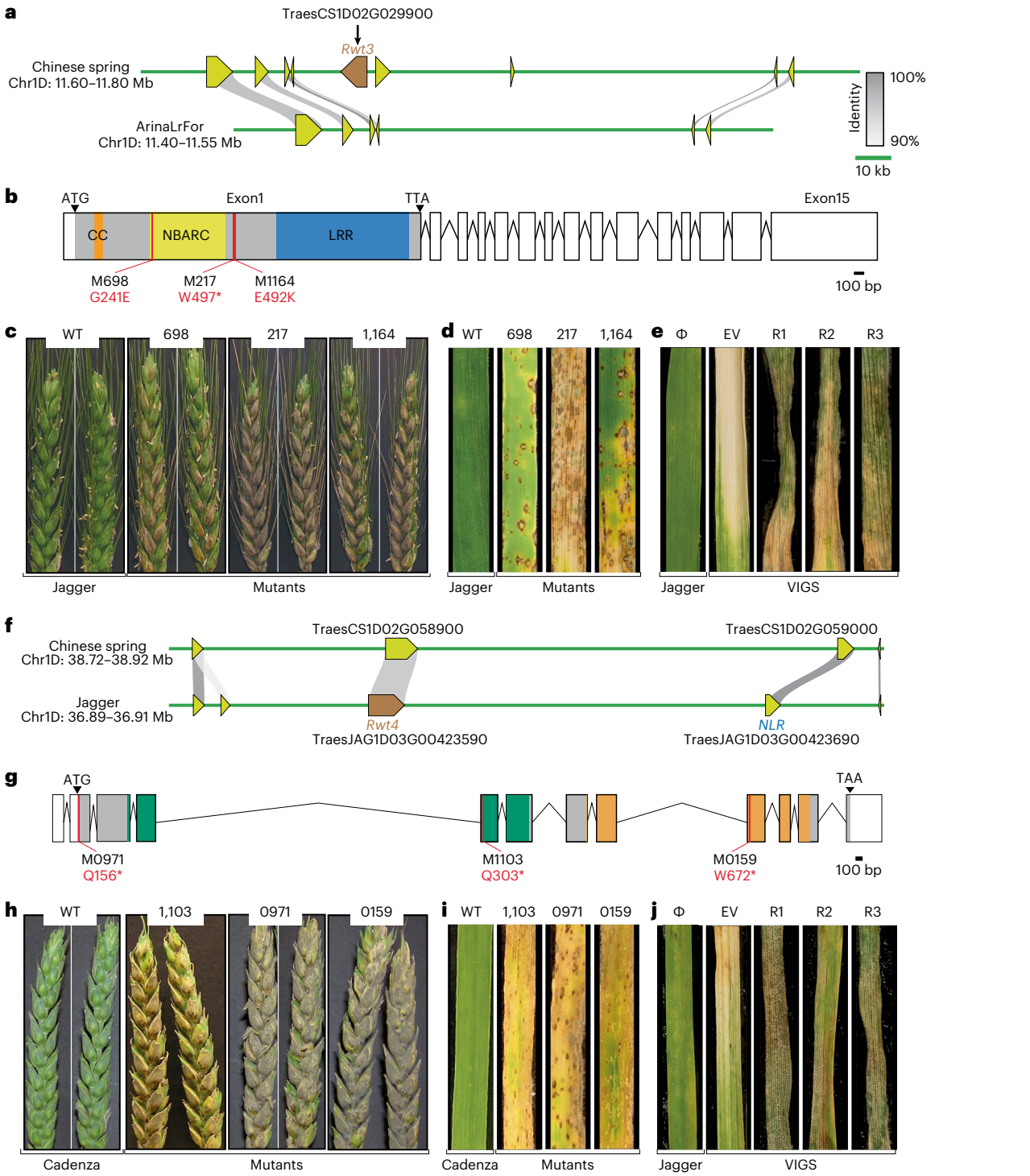


Fig. 2 | Genetic and functional characterization of *Rwt3* and *Rwt4*. **a**, Gene-based collinearity analysis of the two haplotypes linked to *Rwt3* identified in the wheat pangenome. **b**, Structure of the NLR candidate gene for *Rwt3*. The predicted 1,069-amino-acid protein has domains with homology to a CC, NB-ARC and LRRs. **c**,**d**, Wheat blast head (**c**) and detached leaf (**d**) assays for the *Rwt3* Jagger mutants and wild type with Br48 + *PWT3*. **e**, Leaf segments from plants subjected to VIGS with non-virus control (Φ), empty vector (EV) and *Rwt3* target (R1, R2 and R3) and super-infected with Br48 + *PWT3*. **f**, Gene-based collinearity analysis of the two haplotypes linked to *Rwt4* identified in the wheat pangenome. **g**, Structure of the WTK candidate gene for *Rwt4*. The predicted protein of 895 amino acids has domains with homology to a WTK (shown with green and orange colours). **h**,**i**, Wheat blast head (**h**) and detached leaf (**i**) assays for the *Rwt4* Cadenza mutants and wild type with Br48 + *PWT4*. **j**, Leaf segments from plants subjected to VIGS with non-virus control (Φ), empty vector (EV) and *Rwt4* target (R1, R2 and R3) and super-infected with Br48 + *PWT4*.

disease phenotype of Br48 + *PWT4*, restricted to the Watkins lines not containing *Rwt4*. Using Jagger[11] as the reference genome, we obtained a clear peak on chromosome 1B in the region homoeologous to that on 1D containing *Rwt4* (Fig. 1g), indicating that *Rwt4* has a homoeologue on chromosome 1B that provides resistance to *P. oryzae* expressing the *PWT4* effector. We followed the same protocol and re-ran the GWAS

with the leaf assay disease phenotype of Br48 + *PWT3*, restricted to the Watkins lines not containing *Rwt3*. This identified a clear peak on chromosome 2A using Mattis[11] as the reference genome (Fig. 1e) and another on chromosome 7A using Jagger[11] as the reference genome (Fig. 1f). A resistance termed *Rmg2* located on chromosome 7A has previously been identified in the cultivar Thatcher[20] and a resistance termed *Rmg7* has been reported on the distal region of the long arm of chromosome 2A of tetraploid wheat[21]. In both instances the resistances were identified using the same isolate, Br48, as used in our work suggesting that the resistances identified on chromosomes 2A and 7A may correspond to *Rmg7* and *Rmg2* reported previously. Watkins lines carrying either of these two resistance loci showed resistance to both Br48 + *PWT3* and Br48 + *PWT4* (Fig. 1h,i). Additionally, Watkins lines carrying the 7A resistance showed similar levels of resistance to both Br48 and Br48 + *PWT3* (Supplementary Fig. 17). These observations suggested that the 7A and 2A resistance loci interact with Br48 and supported their characterization as *Rmg2* and *Rmg7*, respectively. Based on the in silico predictions, at least one of *Rwt3*, 2A and 7A resistance loci is present in 177 of the 192 Watkins lines resistant to Br48 + *PWT3* (Fig. 1c and Supplementary Table 8), while at least one of *Rwt4*-1D, *Rwt4*-1B, 2A and 7A resistance loci is present in 251 of the 271 Watkins lines resistant to Br48 + *PWT4* (Fig. 1d and Supplementary Table 8).

We designed a KASP marker for the *Rwt4*-1B homoeologue (Supplementary Tables 7 and 9) that, along with those for *Rwt3* and *Rwt4*-1D, should enable wheat breeders to ensure that cultivars maintain host-specificity barriers. It was recently shown that, even though the pandemic clonal lineages tend to dominate *Triticum* pathotypes, sexual recombination between *Triticum* and non-*Triticum* pathotypes is possible and can raise the adaptive potential of *Triticum* pathotypes in the absence of host-specificity barriers[22,23]. This suggests that *PWT4*, which is naturally found in *Lolium* pathotypes isolated from *Avena* but not in *Triticum* pathotypes[6], can potentially be gained by *Triticum* pathotypes through sexual recombination. Gain of *PWT4* by *Triticum* pathotypes would suppress the resistance conferred by *Rmg8* against *AVR-Rmg8* in the absence of *Rwt4* (ref. [24]) (Supplementary Fig. 18). Therefore, the possibility of either gain of *PWT4* or the loss-of-function mutations in *AVR-Rmg8* threatens one of the few reported resistances that show effectiveness against *Triticum* pathotypes at both the seedling and head stage[25].

So far, 11 postulated resistance genes against *P. oryzae* have been identified in wheat and none of these has been cloned[26]. We used isolates of *P. oryzae* differing in a single effector to screen genome sequenced/characterized diversity panels of *Ae. tauschii* and hexaploid wheat to clone *Rwt3* and *Rwt4* resistance genes that confer resistance in wheat to *P. oryzae* isolates carrying *PWT3* and *PWT4*, respectively. The isolation of effector and resistance gene pairs will enable study of their interaction and potential for engineering as well as provide leads to identify genes effective against the *Triticum* pathotype of *P. oryzae*.

Methods

Watkins panel configuration

Using the SSR genotype data from Wingen et al. (2014) (ref. [27]), a core set of 300 genetically diverse wheat landraces with spring growth habit were selected from the Watkins collection (Supplementary Fig. 1 and Supplementary Table 2) along with 20 non-Watkins lines. The DNA was extracted following a modified CTAB protocol[28]. The seeds of these lines are available from the Germplasm Resources Unit (www.seedstor.ac.uk) under wheat resistance gene enrichment (WREN) sequencing collection (WREN0001- WREN0320).

Phenotyping of *Ae. tauschii* and Watkins panels with wheat blast isolates

The *M. oryzae* pathotype *Triticum* (MoT) isolate Br48 and the transformed isolates Br48 + *PWT3* and Br48 + *PWT4* (ref. [6]) were grown on complete medium agar. A conidial suspension of 0.3–0.4 × 10^6 conidia per millilitre was used for all inoculations. Detached seedling assays with the *Ae. tauschii* and Watkins panels were carried out as described by Goddard et al. (2020) (ref. [29]) and scored for disease symptoms using a 0–6 scale (Supplementary Figs. 2, 3 and 7 and Supplementary Tables 3 and 5). Resistance at the heading stage was assessed according to Goddard et al. (2020) (ref. [29]). Heads of *Ae. tauschii* and wheat were scored using a 0–6 scale (Supplementary Fig. 2e and 2f, respectively).

Bait library design for the Watkins panel

Two bait libraries were used for the capture of the immune receptors from the Watkins panel (1) NLR Triticeae bait library V3 (https://github.com/steuernb/MutantHunter/), including 275 genes conserved in grasses[30] and (2) a new bait library that included NLRs extracted from the genomes of *T. turgidum* cv. Svevo and cv. Kronos and *T. dicoccoides* cv. Zavitan, and only those genes that had <50% coverage by previously designed baits were used. To remove redundancies, NLR sequences were passed through CD-HIT (v4.6.8-2017-0621 -c 0.9 -G 0 -aS 0.9 -p 1). This bait design also included wheat domestication genes *VRN1A* (AY747598), *Wx1* (AY050174), *Q* (AY702956), *Rht-b1* (JX993615), *Rht-d1* (HE585643), *NAM-B1* (MG587710) and wheat orthologues of known immune signalling components ICS1, NPR1, NDR1, EDS1, PAD4, SRFR1, SAG101, RAR1, SGT1, HSP90.2, HSP90.4, RIN4, ADR1 and PBS1 extracted through BioMart (Supplementary Table 1 and Additional File 1). The bait probes were designed by Arbor Bioscience and filtered with their Repeat Mask pipeline, which removed the baits that were >50% repeat masked and any non-NLR baits with more than three hits in the wheat genome. To balance for the low copy number genes, baits derived from domestication genes were multiplied 10× and those derived from immune signalling genes were 3× compared with the baits derived from NLRs.

Library construction and sequencing of the Watkins panel

Illumina libraries with an average insert size of 700 bp were enriched by Arbor Biosciences, as previously described[31], and sequenced on an Illumina HiSeq with either 150 or 250 paired end (PE) reads at Novogene, China to generate an average of 3.82 Gb per accession (Supplementary Table 2). The raw reads were trimmed using Trimmomatic v0.2 (ref. [32]) and de novo assembled with the CLC Assembly Cell (http://www.clcbio.com/products/clc-assembly-cell/) using word size (-w = 64) with standard parameters.

Generating Watkins *k*-mer presence/absence matrix and its phylogeny

A presence/absence matrix of *k*-mers ($k = 51$) was constructed from trimmed raw data using Jellyfish[33] as described in Arora et al. (2019) (ref. [15]). *k*-mers occurring in fewer than four lines or in all but three or fewer lines were removed during the construction of the matrix. From the *k*-mer matrix generated with Watkins RenSeq data, 5,310 randomly extracted *k*-mers were used to build an unweighted pair group method with arithmetic mean (UPGMA) tree with 100 bootstraps.

k-mer-based association mapping

For the reference genomes of *T. aestivum*—Chinese Spring[10], Jagger[11] and Mattis[11]—and of *Ae. tauschii* AY61 (ref. [34]), NLRs were predicted using NLR-Annotator[35] and their sequences along with 3 kb sequence from both upstream and downstream region (if available) were extracted using samtools (version 1.9) to create the corresponding reference NLR assemblies. The disease phenotypes were averaged across the replicates after removing the non-numerical values and the mean phenotype scores multiplied by −1 so that a higher value represents a higher resistance. For those *k*-mers of a reference NLR assembly whose presence/absence in the panel correlates with the phenotype, that is, the absolute value of Pearson's correlation obtained was higher than 0.1, a *P* value was assigned using linear regression while taking the three most significant principal component analysis dimensions as covariates to

control for the population structure. A stringent cut-off of 8, based on Bonferroni adjustment[17] to a *P* value of 0.05, was chosen for Watkins RenSeq association mapping (Fig. 1), while a cut-off of 7 was chosen for *Ae. tauschii* L2 RenSeq association mapping (Supplementary Fig. 8).

In silico gene structure prediction

The *Rwt3* NLR candidate gene transcript is 5,937 bp. Only one of the 15 annotated exons (grey-coloured exon in Fig. 2b) appears to be translated into protein. This exon encodes a protein of 1,069 amino acids with a coiled-coil (CC) domain, a nucleotide-binding (NB-ARC) domain and several leucine-rich repeats (LRRs) motifs at the C-terminus (Supplementary Fig. 4). The *Rwt4* NLR candidate gene has three exons encoding 1,038 amino acids with domains having homology to a CC domain, two NB-ARC domains and two LRRs at the C-terminus (Supplementary Fig. 5). The *Rwt4* WTK candidate has an open reading frame of 2,688 bp that has 12 predicted exons that encode a protein of 895 amino acids with putative tandem protein kinase domains (Fig. 2g and Supplementary Fig. 13). Domains were predicted by the National Center for Biotechnology Information (NCBI) and Pfam databases. The gene structure of both *Rwt3* and *Rwt4* NLR candidate genes was consistent with that predicted using complementary DNA RenSeq data of Watkins lines.

Identification and phenotyping of Cadenza TILLING mutants to test the function of *Rwt4*

Cadenza TILLING lines[19] for the NLR candidate for *Rwt4* were identified within the wheat Ensembl database (https://plants.ensembl.org/Triticum_aestivum/) for the gene TraesCS1D02G059000. Lines containing mutations leading to premature stop codons and those for which the 'sorting intolerant from tolerant' (SIFT) score was 0.0 or 0.01 were selected for phenotyping. For the *Rwt4* kinase candidate gene, Cadenza TILLING lines were identified for the gene TraesCS1D02G058900. Details of the mutations present in the Cadenza TILLING lines are provided in Supplementary Table 6.

Identification and phenotyping of Jagger TILLING mutants to test the function of *Rwt3*

For selecting mutations in the *Rwt3* candidate gene (TraesCS1D02G029900), TILLING was performed in wheat cultivar Jagger[18] using genome-specific primer pairs (Supplementary Fig. 10). The effects of the mutations on the predicted protein were analysed using SnapGene software (version 5.0.7 from GSL Biotech). The effects of mis-sense mutations were determined using Protein Variation Effect Analyzer (PROVEAN) v1.1 software[36]. Selected lines were phenotyped as described above. Details of the mutations are provided in Supplementary Table 6.

KASP analysis and sequencing of TILLING lines to confirm mutations

KASP (LGC Genomics) was performed to confirm mutations where suitable PCR primers could be designed. Alternatively, the region containing the mutation was amplified and purified products were sequenced by Eurofins Genomics. Sequence analysis was performed with Geneious Prime software.

VIGS using the BSMV

A 400 bp fragment of *Rwt3* and *Rwt4* (Supplementary Figs. 12 and 16) was cloned into the BSMV vector pCa-cbLIC via ligation independent cloning as described previously[37].

KASP marker design to detect *Rwt3* and *Rwt4* in wheat cultivars and Watkins collection

The regions differentiating *Rwt3* and *Rwt4* from their alternate allelic variants identified in *Ae. tauschii* and wheat were used to design KASP markers (Supplementary Table 7). KASP markers were validated on the sequenced Watkins panel of 300 lines (Supplementary Table 8), following which KASP marker analysis was performed on the entire Watkins panel (~900) to understand the distribution of these genes in the landrace collection (Supplementary Table 9).

Characterization of the resistance identified on chromosome 7A

A set of Watkins lines were genotyped as carrying either *Rwt3* or the 7A resistance or having neither or both resistances. All lines were phenotyped in leaf assays using isolates Br48 and Br48 + *PWT3*. Lines lacking either resistance were susceptible to both isolates (Supplementary Fig. 17). Lines carrying either the 7A resistance alone or both the 7A resistance and *Rwt3* showed a similar level of resistance to both Br48 and Br48 + *PWT3*.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The RenSeq 150 bp paired-end Illumina sequences (raw data) for the 300 Watkins and 21 non-Watkins lines (including Anahuac) and the cDNA RenSeq data of 6 Watkins lines are available from NCBI study number PRJNA760793. The *k*-mer matrix and the CLC assemblies of the 300 Watkins and 21 non-Watkins lines are available from Zenodo under the following DOIs: https://doi.org/10.5281/zenodo.5557564, https://doi.org/10.5281/zenodo.5557685, https://doi.org/10.5281/zenodo.5557721, https://doi.org/10.5281/zenodo.5557827, https://doi.org/10.5281/zenodo.5557838 and https://doi.org/10.5281/zenodo.5655720. The genomic sequences of *Rwt3* and *Rwt4* are available from the wheat Ensembl databases as *TraesCS1D02G029900* from https://plants.ensembl.org/Triticum_aestivum/ and *TraesJAG1D03G00423590* from https://plants.ensembl.org/Triticum_aestivum_jagger/, respectively.

Code availability

Scripts for the Watkins *k*-mer matrix generation, phylogenetic tree construction and *k*-mer-based association mapping can be found in the repository https://github.com/arorasanu/watkins_renseq.

References

1. Igarashi, S., Utiamada, C. M., Igarashi, L. C., Kazuma, A. H. & Lopes, R. S. *Pyricularia* in wheat: 1. Occurrence of *Pyricularia* sp. in Paraná State. *Fitopatol. Bras.* **11**, 351–352 (1986).
2. Cruz, C. D. & Valent, B. Wheat blast disease: danger on the move. *Trop. Plant Pathol.* **42**, 210–222 (2017).
3. Malaker, P. K. et al. First report of wheat blast caused by *Magnaporthe oryzae* pathotype *Triticum* in Bangladesh. *Plant Dis.* **100**, 2330 (2016).
4. Tembo, B. et al. Detection and characterization of fungus (*Magnaporthe oryzae* pathotype *Triticum*) causing wheat blast disease on rain-fed grown wheat (*Triticum aestivum* L.) in Zambia. *PLoS ONE* **15**, e0238724 (2020).
5. Singh, P. K. et al. Wheat blast: a disease spreading by intercontinental jumps and its management strategies. *Front. Plant Sci.* **12**, 1467 (2021).
6. Inoue, Y. et al. Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science* **357**, 80–83 (2017).
7. Rush, M. C., Lindberg, G. D. & Carver, R. B. Blast: a serious new disease of forage grasses in Louisiana. *Phytopathology* **62**, 806 (1972).
8. Farman, M. et al. The *Lolium* pathotype of *Magnaporthe oryzae* recovered from a single blasted wheat plant in the United States. *Plant Dis.* **101**, 684–692 (2017).

9. Rahnama, M. et al. Recombination of standing variation in a multi-hybrid swarm drove adaptive radiation in a fungal pathogen and gave rise to two pandemic plant diseases. Preprint at *bioRxiv* https://doi.org/10.1101/2021.11.24.469688 (2021).
10. International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, eaar7191 (2018)..
11. Walkowiak, S. et al. Multiple wheat genomes reveal global variation in modern breeding. *Nature* **588**, 277–283 (2020).
12. Vy, T. T. P. et al. Genetic analysis of host-pathogen incompatibility between *Lolium* isolates of *Pyricularia oryzae* and wheat. *J. Gen. Plant Pathol.* **80**, 59–65 (2014).
13. Hirata, K., Tosa, Y., Nakayashiki, H. & Mayama, S. Significance of *PWT4–Rwt4* interaction in the species specificity of *Avena* isolates of *Magnaporthe oryzae* on wheat. *J. Gen. Plant Pathol.* **71**, 340–344 (2005).
14. Zhang, Z. et al. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**, 203–214 (2000).
15. Arora, S. et al. Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* **37**, 139–143 (2019).
16. Lu, P. et al. A rare gain of function mutation in a wheat tandem kinase confers resistance to powdery mildew. *Nat. Commun.* **11**, 1–11 (2020).
17. Gaurav, K. et al. Population genomic analysis of *Aegilops tauschii* identifies targets for bread wheat improvement. *Nat. Biotechnol.* **40**, 422–431 (2022).
18. Rawat, N. et al. A TILLING resource for hard red winter wheat variety Jagger. *Crop Sci.* **59**, 1666–1671 (2019).
19. Krasileva, K. V. et al. Uncovering hidden variation in polyploid wheat. *Proc. Natl Acad. Sci. USA* **114**, E913–E921 (2017).
20. Zhan, S. W., Mayama, S. & Tosa, Y. Identification of two genes for resistance to *Triticum* isolates of *Magnaporthe oryzae* in wheat. *Genome* **51**, 216–221 (2008).
21. Tagle, A. G., Chuma, I. & Tosa, Y. *Rmg7*, a new gene for resistance to *Triticum* isolates of *Pyricularia oryzae* identified in tetraploid wheat. *Phytopathology* **105**, 495–499 (2015).
22. Latorre, S. M. et al. A pandemic clonal lineage of the wheat blast fungus. Preprint at *bioRxiv* https://doi.org/10.1101/2022.06.06.494979 (2022).
23. Barragan, A. C. et al. Wild grass isolates of *Magnaporthe* (syn. *Pyricularia*) spp. from Germany can cause blast disease on cereal crops. Preprint at *bioRxiv* https://doi.org/10.1101/2022.08.29.505667 (2022).
24. Inoue, Y., Vy, T. T. P., Tani, D. & Tosa, Y. Suppression of wheat blast resistance by an effector of *Pyricularia oryzae* is counteracted by a host specificity resistance gene in wheat. *N. Phytol.* **229**, 488–500 (2021).
25. Anh, V. L. et al. Rmg8, a new gene for resistance to *Triticum* isolates of *Pyricularia oryzae* in hexaploid wheat. *Phytopathology* **105**, 1568–1572 (2015).
26. Singh, P. K. et al. Wheat blast: a disease spreading by intercontinental jumps and its management strategies. *Front. Plant Sci.* **12**, 710707 (2021).
27. Wingen, L. U. et al. Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat. *Theor. Appl. Genet.* **127**, 1831–1842 (2014).
28. Yu, G., Hatta, A., Periyannan, S., Lagudah, E. & Wulff, B. B. H. Isolation of wheat genomic DNA for gene mapping and cloning. *Methods Mol. Biol.* **1659**, 207–213 (2017).
29. Goddard, R. et al. Dissecting the genetic basis of wheat blast resistance in the Brazilian wheat cultivar BR 18-Terena. *BMC Plant Biol.* **20**, 1–15 (2020).
30. Marcussen, T. et al. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**, 1250092 (2014).
31. Steuernagel, B., Witek, K., Jones, J. D. G. & Wulff, B. B. H. MutRenSeq: a method for rapid cloning of plant disease resistance genes. *Methods Mol. Biol.* **1659**, 215–229 (2017).
32. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
33. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics* **27**, 764–770 (2011).
34. Zhou, Y. et al. Introgressing the *Aegilops tauschii* genome into wheat as a basis for cereal improvement. *Nat. Plants* **7**, 774–786 (2021).
35. Steuernagel, B. et al. The NLR-Annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiol.* **183**, 468–482 (2020).
36. Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* **7**, e46688 (2012).
37. Corredor-Moreno, P. et al. The branched-chain amino acid aminotransferase TaBCAT1 modulates amino acid metabolism and positively regulates wheat rust susceptibility. *Plant Cell* **33**, 1728–1747 (2021).

Acknowledgements

The high-performance computing resources and services used in this work were supported by the Norwich Bioscience Institutes Partnership (NBIP) Computing infrastructure for Science (CiS) group alongside the Earlham Institute (EI) scientific computing group. We are grateful to the John Innes Centre (JIC) Horticultural Services for plant husbandry; EI for providing open access to the Kronos genome; R. Goram at the JIC genotyping platform for the KASP genotyping. This research was financed by the Biotechnology and Biological Sciences Research Council (BBSRC) Designing Future Wheat Cross-Institute Strategic Programme to B.B.H.W. and P.N. (BBS/E/J/000PR9780); a John Innes Centre Institute Strategic Grant to B.B.H.W.; Science, Technology & Innovation Funding Authority (STDF), Egypt-UK Newton-Mosharafa Institutional Links award, Project ID (30718) to A.F.E. and B.B.H.W.; the Gordon and Betty Moore Foundation through grant GBMF4725 to the Two Blades Foundation; and the Gatsby Charitable Foundation to JDGJ; National Science Foundation (award number 1943155) and USDA NIFA (award numbers 2020-67013-32558 and 2020-67013-31460) to N.R. and V.K.T.; European Research Commission grant (ERC-2016-STG-716233-MIREDI) to K.V.K. and BBSRC Norwich Research Park Doctoral Training Grant (BB/M011216/1) for supporting E.B.

Author contributions

This work was conceived by P.N., J.C. and B.B.H.W. Watkins panel configuration, DNA extraction and sequence acquisition (B.B.H.W., L.W., M.F.-M., R.A., S. Arora, G.Y., A.F.E. and J.D.G.J.). Bait library design (K.V.K., E.B. and B.S.), *k*-mer matrix construction and association mapping (S. Arora and K.G.), candidate genes discovery and analysis (S. Arora and K.G.), phylogenetic analysis (S. Arora and K.G.), replication of association mapping results (R.A.-K.), blast isolates (Y.T. and S. Asuke), phenotyping of diversity panels and TILLING mutants (A. Steed, R.G., T.O., P.N., C.C. and M.H.N.), KASP marker design and analysis (S. Arora, A. Steed, K.G. and P.N.), Jagger mutant identification (V.K.T., A. Schoen and N.R.), mutant confirmation and segregation (A. Steed, R.G. and P.N.), cDNA RenSeq data (S. Arora and A.F.E.), VIGS design and construction (D.G.O.S. and A.V.K.), drafted manuscript (S. Arora, P.N., K.G., R.G., A. Steed, V.K.T., Y.T., A. Schoen,

N.R. and K.V.K.) and designed figures (S. Arora, P.N., K.G., R.G., A. Steed, A. Schoen and L.W.).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41477-023-01357-5.

Correspondence and requests for materials should be addressed to Brande B. H. Wulff or Paul Nicholson.

Peer review information *Nature Plants* thanks Zhiyong Liu, Guo-Liang Wang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Paul Nicholson
Brande Wulff

Last updated by author(s): 03-1-2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used during data collection.
Data analysis	Scripts for the Watkins k-mer matrix generation, phylogenetic tree construction and k-mer based association mapping can be found in the repository https://github.com/arorasanu/watkins_renseq. The above custom scripts were written and tested in Python 3.6.13 alongwith following Python modules: numpy (v1.17.0) pandas (v0.23.0) Biopython (v1.78) scikit-learn (v0.24.2) statsmodels (v0.12.2) bitarray (v2.3.0) matplotlib (v3.3.0) Additionally, the following softwares were used for data analysis in the study: BLAST+ command-line tools for alignment version 2.2.28 CD-HIT (v4.6.8) BioMart (https://ensembl.org/info/data/biomart) Trimmomatic (v0.2) CLC assembly cell (https://www.qiagenbioinformatics.com/products/clc-assembly-cell/) Jellyfish (v2.2.6) NLR-Annotator (https://github.com/steuernb/NLR-Annotator) samtools (v1.9) SnapGene (v5.0.7)

PROVEAN (v1.1)
Geneious Prime (v2022.2.2)
iTOL (https://itol.embl.de/)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The RenSeq 150-bp paired-end Illumina sequences (raw data) for the 300 Watkins and 21 non-Watkins accessions (including Anahuac) and the cDNA RenSeq data of 6 Watkins accessions are available from NCBI study number PRJNA760793. The k-mer matrix and the CLC assemblies of the 300 Watkins and 21 non-Watkins accessions are available from Zenodo under the DOIs: 10.5281/zenodo.5557564, 10.5281/zenodo.5557685, 10.5281/zenodo.5557721, 10.5281/zenodo.5557827, 10.5281/zenodo.5557838 and 10.5281/zenodo.5655720. The genomic sequences of Rwt3 and Rwt4 are available from the wheat Ensembl databases as TraesCS1D02G029900 from https://plants.ensembl.org/Triticum_aestivum/ and TraesJAG1D03G00423590 from https://plants.ensembl.org/Triticum_aestivum_jagger/, respectively.

NLR and Kinase domains were predicted using NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and Pfam (http://pfam.xfam.org) databases.

Cadenza TILLING lines for Rwt4 candidate genes were identified using the Plant Ensembl database (http://plants.ensembl.org/Triticum_aestivum/)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No specific sample size experiment was done. A core set of 300 genetically diverse wheat landraces was chosen from the Watkins collection and the number was decided based on the available funding for sequencing.
Data exclusions	No deliberate data exclusion was done.
Replication	In the case of Watkins diversity panel, Br48+Pwt3 and Br48+Pwt4 phenotypes were scored for three replicates per genotype (Supplementary Table 3). In the case of Aegilops tauschii diversity panel, Br48+Pwt3 and Br48+Pwt4 phenotypes were scored for five replicates per genotype while Br48 phenotype was scored for three replicates per genotype (Supplementary Table 5).
Randomization	No deliberate randomization was imposed during the phenotype experiments. The phenotyping was done in a controlled environment where all of the exact same methods were used and scored by experienced researchers. The phenotypes for independent plants of the same genotype were generally consistent (see Supplementary Tables 3 and 5) and resulted in clear GWAS peaks (i) (in the case of Watkins diversity panel) around the loci that had been previously mapped by biparental genetics for Rwt3 and Rwt4, (ii) (in the case of Rwt4) around the same locus using two completely different diversity panels for wheat and Aegilops tauschii, and/or (iii) for which we confirmed the function of candidate genes, thus validating our methods and conclusions.
Blinding	The person doing the phenotyping did not have access to the genotype data. So in retrospect, the data collection was blinded.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging