Genomics accelerated isolation of a new stem rust avirulence gene - wheat resistance gene pair

Narayana M. Upadhyaya[1*], Rohit Mago[1*], Vinay Panwar[2], Tim Hewitt[1], Ming Luo[1], Jian Chen[1,3], Jana Sperschneider[4], Hoa Nguyen-Phuc[5], Aihua Wang[1], Diana Ortiz[1,#], Luch Hac[1], Dhara Bhatt[1], Feng Li[6], Jianping Zhang[1], Michael Ayliffe[1], Melania Figueroa[1], Kostya Kanyuka[2], Jeffrey G. Ellis[1], Peter N. Dodds[1**]

[1] Commonwealth Scientific and Industrial Research Organisation, Agriculture and Food, GPO Box1700, Canberra, ACT 2601, Australia

[2] Biointeractions and Crop Protection, Rothamsted Research, Harpenden, AL5 2JQ, United Kingdom

[3] Research School of Biology, Australian National University, Canberra ACT 2601, Australia

[4] Biological Data Science Institute, The Australian National University, Canberra, ACT 2601, Australia

[5] Department of Ecology and Evolutionary Biology, Vietnam National University, 227 Nguyen Van Cu Street, District 5, HCMC, Vietnam

[6] Department of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota 55905, USA

* These authors contributed equally

** author for correspondence: peter.dodds@csiro.au

current address: Génétique et Amélioration des Fruits et Légumes (GAFL), INRA, Domaine Saint Maurice, 63 Allée des Chênes, CS60094, 84143 Montfavet Cedex, France

Stem rust caused by the fungus *Puccinia graminis* f. sp. *tritici* (*Pgt*) is a devastating disease of the global staple crop wheat. Although this disease was largely controlled by genetic resistance in the latter half of the 20th century, new strains of *Pgt* with increased virulence, such as Ug99, have evolved by somatic hybridisation and mutation [1,2]. These newly emerged strains have caused significant losses in Africa and other regions and their continued spread threatens global wheat production. Breeding for disease resistance provides the most cost-effective control of wheat rust diseases [3]. A number of race-specific rust resistance genes have been characterised in wheat and most encode immune receptors of the nucleotide-binding leucine-rich repeat (NLR) class [4]. These receptors recognize pathogen effector proteins often known as avirulence (Avr) proteins [5]. However, only two *Avr* genes have been identified in *Pgt* to date, *AvrSr35* and *AvrSr50* [6,7] and none in other cereal rusts, which hinders efforts to understand the evolution of virulence in rust populations [8,9]. The *Sr27* resistance gene was first identified in a wheat line carrying an introgression of the 3R chromosome from Imperial rye [10]. Although not deployed widely in wheat, *Sr27* is widespread in the artificial crop species Triticosecale (triticale) which is a wheat-rye hybrid and is a host for *Pgt* [11,12]. *Sr27* is effective against Ug99 [13] and other recently emerged *Pgt* strains [14,15]. Here we identify both the *Sr27* gene in wheat and the corresponding *AvrSr27* gene in *Pgt* and show that virulence to *Sr27* can arise experimentally and in the field through deletion mutations, copy number variation and expression level polymorphisms at the *AvrSr27* locus.

The wheat stem rust isolate Pgt21-0 is avirulent on *Sr27*, but derivatives of this strain with virulence to *Sr27* have been detected in the field [16,17], suggesting that it may be heterozygous for *AvrSr27*. We therefore selected for spontaneous virulent rust mutants after inoculation of Pgt21-0 onto the resistant triticale cultivar 'Coorong' carrying *Sr27*. Three large single pustules were selected, purified in isolation and reinoculated onto Coorong to confirm their virulent phenotype (Fig. 1A and Supplementary Fig. 1). We generated Illumina sequence data from genomic DNA of each mutant and mapped the reads to the Pgt21-0 haplotype-resolved genome assembly [2] to identify potential mutations. We first screened for SNPs causing amino-acid changes in secreted protein genes, but found no genes with such mutations in more than one mutant line. Next we looked for loss of read coverage as an indicator of potential deletion mutations. Each of the mutant lines showed a large number of genes with zero read coverage, all located on one end of chromosome 2B. Visualisation of read mapping depth to chromosome 2B revealed the three mutants each contained independent and overlapping deletions (Fig. 1B). The largest was 1.5Mbp in size and corresponds to the whole of the short arm of this

chromosome. The smallest common deleted region was 196kbp and contains 50 annotated genes, five of which are predicted to encode secreted proteins (Fig. 1C).

We previously sequenced two Australian *Pgt* isolates (Pgt34-2,12 and Pgt34-2,12,13) that are clonally derived from Pgt21-0 and had evolved virulence for *Sr27* in the field [16,17]. Analysis of read coverage revealed that these two isolates each contained a small deletion of 13 kbp in the same region of chromosome 2B that was deleted in the three spontaneous mutants. This small deletion spanned two of the secreted protein genes (PGT21_006532 and PGT21_006593) along with a single adjacent gene on the distal side (Fig. 1C and Supplementary Fig. 2). No other genes in this region contained any changes in these two isolates. We confirmed the presence of this deletion in the genome of Pgt34-2,12 by PCR amplification of the deletion boundaries (Supplementary Fig. 3). We also examined sequence data previously generated from seven South African isolates [18] that are part of the same clonal lineage as Pgt21-0 [2,19]. Pgt21-0 migrated to Australia from South Africa and was first detected in 1954, and these two branches of this lineage have since evolved in isolation. Four of the South African isolates examined are virulent on *Sr27* [20] and must have evolved this phenotype independently of the Australian isolates. Sequence analysis revealed that three of these virulent lines each contain an identical ~14kbp deletion that overlaps with the deletion in Pgt34-2,12 and spans the two candidate secreted protein genes plus an additional gene on the proximal side, while the fourth isolate contains an independent deletion of 10 kbp that spans just the two secreted protein genes (Fig. 1C and Supplementary Fig. 2). Importantly all five of the avirulent isolates of this lineage contain a similar sequence to Pgt21-0 in this region (Fig 1C). A phylogenetic analysis of these clonal isolates places each unique deletion event in a separate branch, consistent with three independent mutations to virulence on *Sr27* during the diversification of this lineage (Supplementary Fig. 4). The deletion of these two secreted protein gene candidates in three independent field-derived virulent lineages provides strong evidence that at least one of these genes confers the *AvrSr27* avirulence phenotype.

The two *AvrSr27* candidate genes are closely related to each other and encode predicted secreted proteins of 144 amino acids designated AvrSr27-1 and AvrSr27-2 (Fig. 2A). A single copy of a related gene (PGT21_007034) occurs at the alternate (virulence) allele on chromosome 2A (designated *avrSr27-3*) and the three protein variants differ at 35 amino acid positions. These proteins show no similarity to any proteins from other rust species. To test the avirulence function of these proteins, we generated recombinant *Barley stripe mosaic virus* (BSMV) [21] expressing the *AvrSr27* candidates without their predicted signal peptides. We

found that BSMV expressing any of the three *AvrSr27* variants was unable to infect the triticale line Coorong, while the control virus vector (BSMV:mcs4d) carrying a 275-nt noncoding DNA fragment was highly infective (Fig. 2B, Supplementary Fig. 5a). All BSMV strains could infect the rust susceptible triticale line Rongcoo. We further tested these strains for infection of an EMS derived mutant of Coorong that had lost *Sr27* resistance (see below). All were able to infect this line, indicating that recognition was specifically conferred by the *Sr27* resistance gene (Fig. 2B, Supplementary Fig. 5a). Furthermore, BSMV strains expressing PGT21_006334 (another secreted protein candidate in the M1 196kbp deletion but outside the small deletions) were fully able to infect Coorong indicating that recognition was specific to the *AvrSr27* candidates. We also tested *AvrSr27* recognition in the wheat line Chinese Spring WRT238.5, containing the 3RS translocation carrying *Sr27*. BSMV strains expressing *AvrSr27* variants could not infect WRT238.5 containing *Sr27* but could infect the susceptible line Chinese Spring (Supplementary Fig. 5b).

It was unexpected that the *avrSr27-3* variant from the virulence allele also conferred recognition by *Sr27* in these assays. This allelic variant did not contain any sequence changes in any of the *Sr27*-virulent isolates examined. Thus, there must be some reason apart from the amino acid differences in the mature protein that explains the lack of an avirulence phenotype conferred by this allele in *Pgt*. One possibility is that expression level differences between the variants may explain this phenotype. We therefore analysed Illumina RNAseq data from an infection time course and isolated haustoria of Pgt21-0 [6] mapped to the Pgt21-0 annotated genome [2]. This analysis showed that the *AvrSr27-2* variant (PGT21_006593) was expressed approximately four-fold higher than the other two variants (Fig. 3A) and therefore likely contributes most to avirulence during infection of *Sr27* plants. Notably, we found that the *avrSr27-3* variant on chromosome 2A (retained in virulent isolates) accounts for only about 15% of the total expression from this locus. We hypothesise that this could be below a threshold required to confer an avirulence phenotype in the absence of the other two variants encoded on chromosome 2A. Consistent with this, RT-PCR analysis confirmed that the combined expression at the *AvrSr27* locus was substantially reduced in the virulent *Pgt* mutant #1 compared to the wildtype avirulent Pgt21-0 (Supplementary Fig. 6). Effector expression level polymorphisms rather than protein sequence variation were also observed to underlie virulence in some *Phytophthora spp.* isolates [22,23]. Cluster analysis of expression profiles of all secreted protein genes in Pgt21-0 revealed eight clusters with different expression profiles in germinated spore, haustoria and in infected leaves (Fig. 3B). Clusters 2, 3 and 7 are characterised by

upregulated expression in haustoria relative to germinated spores and early expression during plant infection. Cluster 6 includes haustorially enriched genes that are induced late during infection, while genes in clusters 1, 5 and 8 are preferentially expressed in germinated spores and cluster 4 included genes with relatively stable expression. The three known Pgt *Avr* genes shared a similar expression pattern and appear in clusters 2 (*AvrSr35*) and 7 (*AvrSr27* and *AvrSr50*). This shared expression profile is similar to that observed for six confirmed *Avr* genes identified in the flax rust pathogen *Melampsora lini* [24], supporting the use of expression profiling to prioritise candidate genes for avirulence function in rust fungi.

We also examined the *AvrSr27* locus in the reference sequence of Ug99 [2], which is avirulent on *Sr27* plants [13]. The A genome haplotype of Ug99 is shared with 21-0 and contains an identical sequence to the virulence allele *avrSr27-3* (PGTUg99_032104). The C genome haplotype (contig #00000029) encodes two proteins, one identical to AvrSr27-2 (not annotated) and one with a single amino acid difference from AvrSr27-1 (PGTUg99_024168). However, this genome region is distinguished from the Pgt21-0 chromosome 2B haplotype by the presence of two large insertions (7 kbp and 9 kbp) and a duplication of about 1.5 kbp within the intergenic region between the *AvrSr27* paralogs (Supplementary Fig. 7). Hence the Ug99 strain is also likely heterozygous for avirulence on *Sr27* plants.

We used a MutRenSeq [25] approach to identify the *Sr27* gene. Seeds of the triticale variety Coorong carrying *Sr27* were treated with EMS and 1960 M2 families were screened for susceptible mutants by infection with the avirulent Pgt21-0. A total of 35 putative mutants were identified and 27 were subsequently confirmed in the following generation as homozygous susceptible (Fig. 4A). NLR-gene capture and sequencing was performed on wildtype Coorong and four confirmed susceptible mutants (M2, M3, M4 and M6). Sequence reads from Coorong were assembled and reads from all lines aligned to this reference to identify sequence changes in the mutant lines. One contig (#5723) of 1126 bp was found that contained a mutation in three of the four mutants; one (M2) with a full deletion of the sequence and two (M3, M4) with single base changes causing amino acid substitutions Q264R and G209S (in the conserved p-loop). This contig contained a coiled-coil (CC) domain and p-loop motif, but not the rest of the NB-ARC or LRR domains, suggesting that it represented only part of a full length NLR gene. To identify the remainder of the gene, the contig was aligned to the *Triticum aestivum* cv. Chinese Spring reference (CSv1) assembly IWGSC RefSeq v1.0 [26]. The top hit (93.6%) was to the 5'end of a high confidence gene (TraesCS6B01G464400) predicted on chromosome 6B and functionally annotated as a disease resistance gene. The full gene sequence of

TraesCS6B01G464400 was then aligned back to the Coorong RenSeq *de novo* assembly which detected an additional contig (#2413) aligning to the 3' region of this gene with 93.8% identity. Contig #2413 contained both an NB-ARC and LRR domain, and inspection of the mutant line read alignments confirmed that it was also deleted in mutant M2 and identified an additional single base change in mutant M6 (Supplementary Fig. 8). PCR amplification from genomic DNA confirmed that these two contigs are derived from the same gene in Coorong. A transcript assembled from RNAseq data of the cultivar Coorong matches sequence of both contigs and encodes a full-length CC-NB-LRR protein of 956 amino acids with a single intron within the coding sequence (Fig. 4B). Amplification from the four mutants confirmed the nucleotide changes detected by Renseq in each line. Four additional mutants were also examined by PCR amplification and one (M3) contained a single amino acid change (T211I) in the p-loop while the gene sequence failed to amplify from the other three (M7, M8, M9) indicating that they contained deletions of this region. The eight independent mutants containing deletions or missense mutations in this gene provided strong evidence that it confers *Sr27* resistance. Among known wheat resistance proteins, Sr27 is most closely related to Sr13 (90% amino acid identity; Supplementary Fig. 9,10).

To confirm the function of the *Sr27* candidate gene, we first used a wheat protoplast transfection assay to co-express the gene with the *AvrSr27* variants identified above as well as the reporter gene luciferase [27,28]. Recognition of a *Pgt* avirulence protein by the corresponding disease resistance protein in this assay leads to cell death and therefore a reduction in the expression of the co-transformed luciferase reporter gene which is detected by its bioluminescence. Co-expression of the *Sr27* candidate gene with each of the three *AvrSr27* gene variants led to a strong reduction in luciferase activity compared to the *Sr27* or *AvrSr27* genes alone (Fig. 4C). This was a specific recognition response as no loss of reporter gene expression was seen when *Sr27* was co-expressed with the unrelated *Pgt* avirulence gene *AvrSr50* [6], nor when *AvrSr27-1* was co-expressed with the unrelated wheat resistance gene *Sr50* [29]. This result also confirmed that the protein encoded by the *avrSr27-3* virulence allele in Pgt21-0 could be recognised by *Sr27* as observed in the virus-mediated protein overexpression assay (Fig. 2B). Likewise, co-expression of *Sr27* and the three *AvrSr27* variants by agrobacterium-mediated transient transformation in *Nicotiana benthamiana* resulted in cell death induction (Figure 4D, Supplementary Fig.11). Lastly, we generated transgenic wheat lines expressing the *Sr27* gene and found that they showed strong resistance to a wheat stem rust isolate avirulent on *Sr27* (Figure 4E and Table S1), but were fully susceptible to strains of

leaf rust (*P. triticina*) and stripe rust (*P. striiformis*) (Table S1) confirming that this gene confers specific resistance to stem rust in wheat.

Sr27 is effective against Ug99 and other *Pgt* strains causing recent epidemics and could provide effective protection against wheat stem rust, particularly if used in combination with other resistance genes. However, the current *Sr27*-containing wheat line with the rye 3RS translocation introgression have been reported to show a yield penalty [30]. Therefore it may be necessary to generate smaller 3RS translocation segments in wheat to alleviate this linkage drag effect. This approach was successful for generating smaller segments of the *Thinopyrum ponticum* chromosome carrying *Sr26* and *SrB* (now *Sr61*) [31]. Alternatively, *Sr27* may be used as part of a resistance gene cassette in a transgenic approach to provide durable resistance through pyramiding multiple effective resistance genes at a single locus [3,32]. The feasibility of this approach was recently demonstrated by the generation of transgenic wheat expressing a gene stack of five cloned *Sr* genes [28]. The identification of *AvrSr27* allows for confirmation of *Sr27* transgene function in such a multi-resistant line through transient protoplast expression, as demonstrated for *Sr50* and *Sr35* [28]. The *AvrSr27* gene identification also provides the capacity to detect virulence genotypes in *Pgt* populations through sequencing approaches. Together with *AvrSr35* and *AvrSr50*, this represents the beginnings of a molecular diagnostic toolbox for effective prediction of virulence phenotypes of *Pgt* isolates from genotype sequence data. Given the observed role of gene-expression polymorphisms in virulence on *Sr27*, RNA-based detection [33] may be a valuable approach. This work also highlights that field-based expression data as well as genome sequence may be required for accurate phenotype prediction. This would facilitate regional deployment of the most appropriate stem rust resistance genes and their combinations for effective disease control.

METHODS

Rust isolates and plant material

Australian *Pgt* strains Pgt21-0 and 98-1,2,(3),(5),6 have been described previously [2,16,34]. Triticale variety Coorong contains *Sr27* [11] while Rongcoo is a line of triticale susceptible to Australian *Pgt* isolates [35]. The wheat line CS WRT238.5 contains a 3RS translocation carrying the *Sr27* resistance gene derived from Imperial Rye in the Chinese Spring background [36]. Triticale and wheat plants were grown in a growth chamber at 23°C with a 16-hour light period and 50% humidity.

Pgt virulent mutant selection

Seedlings of the triticale cultivar Coorong at the 2-3 leaf stage (20 pots [15cm] with 6-8 plants per pot) were inoculated with Pgt21-0 (~250 µL) using talc as a carrier (1:4; spores:talc). Inoculated plants were incubated in a humid chamber maintained at 23°C for 48 hrs. After this time plants were moved to glasshouse maintained at 23°C/18°C (day/ night) under natural daylight. Plants were screened twice, at 14 and 20 days post inoculation, for any susceptible infection sites showing large pustule development typical of a susceptible interaction (infection type 3,4). More than 10 mutant pustules were detected and three of these were collected and re-inoculated onto the susceptible wheat line Morocco for amplification in isolation to prevent cross contamination between isolates. After amplification, the virulence of each mutant on *Sr27* was confirmed by reinfection of Coorong plants. One mutant line was tested on the full Australian differential set for *Pgt* and showed an identical virulence profile to Pgt21-0 except for a single additional virulence for *Sr27*. No other isolates virulent on Coorong were present in the laboratory prior to the mutation screen ruling out contamination as a source of the virulent pustules.

DNA isolation and sequencing

DNA was extracted from urediniospores of mutant *Pgt* lines as described [16], quality assessed with a Nanodrop Spectrophotometer (Thermo Scientific, Wilmington, DE) and the concentration quantified using a broad-range assay in Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). DNA library preparation and Illumina sequencing were performed by the Australian Genome Research Facility (AGRF) on the HiSeq2500 (250 bp PE reads, mutants M1 and M2) or MiSeq (300bp PE reads, mutant M3) platforms and about 20 million reads

obtained for each isolate. Sequence reads were imported to CLC Genomics Workbench (CLCGW) version 10.0.1 or later (QIAGEN), filtered and trimmed to remove low quality ends, sequencing adapters and low- quality reads (Trim using quality score 0.01, maximum number of ambiguities 2). Reads from the three mutants as well as from the parent Pgt21-0 were mapped to the karyon-phased chromosome level Pgt21-0 reference genome [2] using high stringency settings (similarity fraction 0.98 and length fraction 0.95). Variant calling of each sample against the reference was performed using the "Basic Variant Detection" program with parameters: ignore non-specific matches; minimum coverage 10; significance 1.0%; minimum variant count 2; include broken pairs. The program "Compare variants within group" in CLCGW was used to identify variants specific to each sample as well as shared variants. Non-synonymous variants in secreted protein genes were predicted using the CLCGW tool "Amino Acid Changes" and curated manually by visual inspection of read mapping tracks in CLCGW. Read coverage statistics (mean read depth and percent coverage) for annotated genes were extracted using the "Create Statistics for Target Regions" program.

Phylogenetic analyses

For whole-genome SNP calling and phylogenetic analysis we used Illumina DNA sequence data from six Australian isolates [16], seven South African isolates and one Czech isolate [18] which we previously showed to belong to a single clonal lineage [2]. FreeBayes v. 1.1.0 [37]was used to call biallelic SNP variants across all samples simultaneously using trimmed reads aligned to the Pgt21-0 genome assembly [2]. VCF files were filtered using vcffilter in vcflib (v1.0.0-rc1) with the parameters: QUAL > 20 & QUAL / AO > 10 & SAF > 0 & SAR > 0 & RPR > 1 & RPL > 1 & AC > 0. VCF files were converted to multiple sequence alignment in PHYLIP format using the vcf2phylip script (*https://zenodo.org/record/1257058#.XNnE845Kh3g*) and R-package ips/phyloch wrappings (*http://www.christophheibl.de/Rpackages.html*). Phylogenetic trees were constructed using the maximum likelihood criterion in RAxML v. 8.2.1.pthread [38], assuming unlinked loci and using 500 bootstrap replicates with a general time reversible model. Convergence and posterior bootstopping (bootstrapping and convergence criterion) were confirmed with the *-I* parameter in RAxML and also with R-packages ape [39], ips/phyloch and phangorn [40]. Dendrograms were drawn using ggplot2 [41] and ggbio [42] R-packages. For phylogenetic analyses of Sr27 and related resistance proteins, protein sequences were aligned by CLUSTAL and a Maximum likelihood tree constructed using MEGA version X [43].

Virus-mediated *in planta* overexpression of candidate AvrSr27 proteins

The *Barley stripe mosaic virus* (BSMV) mediated protein overexpression system known as "BSMV VOX" [44,45] comprising three T-DNA binary plasmids, pCaBS-α, pCaBS-β and pCa-γb2A-LIC, was used in this study. The coding sequence of AvrSr27 variants with the first 27aa of the predicted signal peptide excluded and replaced by a single methionine start codon was amplified by PCR using primer pairs 2-210.F and 2-210.R (*AvrSr27-1*, and *AvrSr27-3*) and 2-210.F and 5-210.R (*AvrSr27-2*) (Supplementary Table 2) and cloned into pCa-γb2A-LIC to generate BSMV:AvrSr27-1, -2, -3. BSMV:mcs4d [46], which contains 275-bp non-coding sequence corresponding to the multiple cloning site sequence of the vector pBluescript II KS, was included as a negative control. All constructs were verified by Sanger sequencing. Transformation of the BSMV binary plasmids into *A. tumefaciens* strain GV3101 (pMP90) and agroinfiltration into 3-4 week old *N. benthamiana* plants was done as described previously [47]. Sap from the infiltrated *N. benthamiana* leaves was used to mechanically inoculate leaves of two-leaf stage triticale or wheat plants. Post-inoculation plants were maintained in a contained environment room with day/night temperatures of 23°C/20°C at around 65% relative humidity and a 16 hr photoperiod with light intensity of approximately 180 $\mu molm^{-2}s^{-1}$. The presence/absence of virus symptoms in subsequently emerged leaves due to systemic infection of the virus was assessed weekly from 7 to 21 days post-infection. The numbers of inoculated plants showing systemic infection symptoms were recorded for each treatment. Each inoculation experiment involved at least 10 individual plants per BSMV VOX construct and was repeated at least once with similar results.

Expression analysis

RNA reads from germinated spores and haustorial tissue (100 bp paired-end) of Pgt21-0 were obtained from NCBI BioProject PRJNA253722 [16] while reads from an infection timecourse [6] were obtained from NCBI BioProject PRJNA415866. Salmon 1.1.0 [48] was used to align reads to the *Pgt* 21-0 transcripts and to estimate transcript abundances (TPMs) in each sample (salmon index –keepDuplicates and salmon quant –validateMappings). We used tximport and DESeq2 [49] to assess gene differential expression and to obtain the regularized logarithm transformation (rlog). Secreted protein genes showing differential expression (padj < 0.1 and -1.5 < logFC > 1.5) for at least one time point during infection versus germinated spores were selected with DESeq2 and k-means clustering using Euclidean distance was performed on the relative rlog transformed counts (average rlog for biological replicates) of the 1,352

differentially expressed genes in the secretome. Rather than absolute expression we used the amount by which each gene deviates in a specific sample from the gene's average across all samples. For RT-PCR confirmation of expression, RNA was extracted from infected leaves 5 days post inoculation using the Qiagen RNAEasy kit and treated with DNAse (Turbo DNAse kit). One microgram of total RNA was used for 1st strand synthesis with SuperScript III and oligo(dT) (Invitrogen) according to the manufacturer's instructions (20 µl reaction) and diluted to 100 µl in water. PCRs were performed with gene specific primers (Supplementary Table 2) using GoTaq polymerase with following cycles: 95°C 3min; 95°C 30sec, 58°C 40sec, 72°C 50sec 29 cycles; 72°C 5min.

Isolation of *Sr27* mutants and NLR gene capture sequencing

Seeds of triticale variety Coorong carrying *Sr27* were subject to EMS mutagenesis as described [50]. In short, 2000 seeds were treated for 12 hrs with 0.3% Ethylmethanesulfonate (EMS) then washed thoroughly with water and sown in large pots (40 seeds per 30cm pot). Single heads from each M1 plant were threshed separately and M2 families from each plant were sown in a tray (30 M2 families per tray) along with resistant and susceptible controls. One-week seedlings were inoculated with Pgt21-0 and scored for their phenotype 2 weeks post inoculation. Seed was harvested from susceptible plants and retested in the following generation by rust inoculation. DNA was extracted from wild-type Coorong triticale and four homozygous susceptible mutants as described [51]. NLR gene capture was performed by Arbor Biosciences (Ann Arbor, USA) following the myBaits protocol with version 2 of the Triticeae bait library available at github.com/steuernb/MutantHunter. Library construction followed the TruSeq RNA Protocol v2. All enriched libraries were sequenced by a HiSeq 2500 (Illumina) using 250 bp paired-end reads.

NLR read assembly and identification of mutations

Primary read sequencing data from wild-type and mutants were first trimmed for quality using Trimmomatic v0.38 [52] with the parameters *ILLUMINACLIP:novogene_indexed_adapters.fa:2:30:10:8:TRUE, LEADING:28, TRAILING:28, MINLEN:20*. Data from wild-type was then *de novo* assembled using CLCGW v11.0.1 with a length fraction of 0.95 and a similarity fraction of 0.98 and the remaining parameters as default. Contigs shorter than 1kb were omitted from the final assembly using a custom script. An annotation of NBS-LRR motifs was created using the program NLR-Parser [25]. Trimmed data of each mutant and wild type was mapped to the *de novo* assembly using

BWA v0.7.15 [53]. Samtools v1.7.0 [54] was used for processing of resulting SAM files to retain only reads mapping in a proper pair with parameter *-f 2*, then removing duplicates and generating pileup files using parameters *-BQ0* and *-aa*. SNV calling and subsequent candidate identification was performed using the following scripts from the MuTrigo pipeline (https://github.com/TC-Hewitt/MuTrigo). Contig regions with high levels of SNPs in the wildtype data indicating poor assembly or read alignment were detected using *Noisefinder.pyc* with default parameters and masked prior to downstream analysis. Potentially mutated nucleotide positions were recorded from pileup files using *SNPlogger.pyc* with parameter *-d 20*. *SNPtracker.pyc* was used to retrieve contigs containing polymorphisms in two or more mutants using default parameters plus parameter *-s C\>T G\>A indel*. This translates to considering only polymorphisms with a minimum of 80% mutant allele frequency and selecting only for insertions, deletions, or C-to-T or G-to-A SNVs that do not share an identical position with another mutant or wild-type. Candidate gene contigs were aligned to the chromosome scale reference assembly of *Secale cereal*e inbred line 'Lo7' [55] and the *Triticum aestivum* cv. Chinese Spring reference (CSv1) assembly IWGSC RefSeq v1.0 [26] using BLAST v2.7.1.

Confirming the full length candidate *Sr27* sequence

Primers Sr27c5723ExtF1 and Sr27c2413ExtR1 (Supplementary Table 2) were designed based on the genomic sequence of the two non-overlapping contigs #2413 and #5723 and were used to amplify the non-overlapping region between these contigs. A *de novo* RNA transcript assembly was generated from Illumina RNAseq data from Coorong seedlings infected with *Pgt* [16] (downloaded from NCBI-SRA SAMN07836894, PRJNA415866) using CLCGW v11.0.1 (similarity fraction of 0.98 and the remaining parameters as default). A full-length RNA transcript matching the genomic sequence and including the 5' and 3' UTR regions was identified and the primers Sr27F and Sr27R1 were used for amplification for the full-length gene from the wildtype and mutants for subsequent sequence comparison. All PCRs were performed using Phusion high fidelity DNA polymerase (NEB, USA) according to manufacturer's instructions.

Generation of expression constructs

AvrSr27 gene sequences with the first 27aa of the signal peptide excluded and replaced by a single methionine start codon were amplified from cDNA from purified haustoria of Pgt21-0 using Phusion high-fidelity DNA polymerase (Thermo Scientific) and cloned into

pENTR™/D-TOPO® according to manufacturer's (Invitrogen) instructions. Full length Sr27 cDNA was also cloned into pENTR™/D-TOPO®. For wheat protoplast transfection *Sr27* (coding sequence including introns) and *AvrSr27* sequences were inserted into the p35s-pUbi-GTW-GFP [56] using Gateway® Technology (Life Techonologies™). Plasmids encoding GFP [57] and Luciferase [27] were previously described. Primers used for PCR are shown in Supplementary Table 2. LR Clonase was used to introduce inserts from entry clones Gateway-compatible pGADT7 and pGBKT7 (Clontech) vectors for yeast-two-hybrid assays [58], and pAM-PAT vectors [59,60] for *in planta* transient expression driven by the CaMV 35S promoter. All plasmids were confirmed by sequencing and analysed using Vector NTI Advance (Life Technologies) or CodonCode Aligner V.4.0.4 (CodonCode Corporation) software.

Transient expression assays

Wheat protoplast isolation and transformation were performed as described [57]. High quality plasmid DNA was isolated by using Qiagen Endo-free Plasmid Maxi kits (Cat#12362). DNA concentrations were adjusted to a range from 1µg/µl to 4µg/µl and 10 µg of each plasmid was used in co-transformation experiments. After 24 hours incubation in the dark at 23°C, luciferase activity was quantified as described [28]. *N. benthamiana* plants were grown in a growth chamber at 23°C with 16 h light period. Agrobacterium cultures containing the *Sr27* and *AvrSr27* expression vectors were grown overnight at 28°C in LB media with appropriate antibiotic selections. The cells were pelleted and resuspended in infiltration mix (10 mM MES pH 5.6, 10 mM $MgCl_2$, 150 µM acetosyringone) to an OD_{600} = 0.5 or 1.0, followed by incubation at room temperature for 2 h. Cultures were infiltrated into leaves of 4-week-old tobacco with a 1-ml syringe. For documentation of cell death, leaves were photographed 2-5 days after infiltration

Wheat transformation

Scutellum tissue from wheat cultivar Fielder was transformed with Agrobacterium strain AGL0 containing the Ubi-Sr27 construct using hygromycin (30mg/mL) for selection as previously described [61,62]. T0 lines were screened by PCR with primers Sr27c5723F/R and three positive lines were identified (PC311.3, PC311.17 and PC311.18). T1 seed was harvested and eleven or twelve progeny of each line were grown in a growth cabinet (23°C, 16 h light). Two-week-old seedlings were inoculated with *Pgt* race 98-1,2,(3),(5),6 to which *Sr27* confers resistance, and rust reactions were assessed after 10-15 days. Additional T1 progeny of each

line were further tested by inoculation with a leaf rust isolate of pathotype 76-1,3,7,9,10,12,13 and a stripe rust isolate of pathotype 134 E16 A+17+33+.

References

1. Singh, R.P., *et al.* Emergence and Spread of New Races of Wheat Stem Rust Fungus: Continued Threat to Food Security and Prospects of Genetic Control. *Phytopathology* **105**, 872-84 (2015).
2. Li, F., *et al.* Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nature Communications* **10** (2019).
3. Ellis, J.G., Lagudah, E.S., Spielmeyer, W. & Dodds, P.N. The past, present and future of breeding rust resistant wheat. *Frontiers in Plant Science* **5** (2014).
4. Periyannan, S., Milne, R.J., Figueroa, M., Lagudah, E.S. & Dodds, P.N. An overview of genetic rust resistance: From broad to specific mechanisms. *Plos Pathogens* **13** (2017).
5. Garnica, D.P., Nemri, A., Upadhyaya, N.M., Rathjen, J.P. & Dodds, P.N. The Ins and Outs of Rust Haustoria. *Plos Pathogens* **10** (2014).
6. Chen, J., *et al.* Loss of AvrSr50 by somatic exchange in stem rust leads to virulence for Sr50 resistance in wheat. *Science* **358**, 1607-1610 (2017).
7. Salcedo, A., *et al.* Variation in the AvrSr35 gene determines *Sr35* resistance against wheat stem rust race Ug99. *Science* **358**, 1604-1606 (2017).
8. Bakkeren, G. & Szabo, L.J. Progress on Molecular Genetics and Manipulation of Rust Fungi. *Phytopathology* **0**, PHYTO-07-19-0228-IA (2020).
9. Figueroa, M., Dodds, P.N. & Henningsen, E.C. Evolution of virulence in rust fungi - multiple solutions to one problem. *Curr Opin Plant Biol* **56**, 20-27 (2020).
10. McIntosh, R.A., Wellings, C.R., Park, R.F. Wheat rusts: an atls of resistance genes. 1995, Melbourne: CSIRO publications.
11. McIntosh, R.A., Luig, N.H., Milne, D.L. & Cusick, J. Vulnerability of triticales to wheat stem rust. *Canadian Journal of Plant Pathology* **5**, 61-69 (1983).
12. Zhang, J., Wellings, C.R., McIntosh, R.A. & Park, R.F. Seedling resistances to rust diseases in international triticale germplasm. *Crop & Pasture Science* **61**, 1036-1048 (2010).
13. Jin, Y., Pretorius, Z.A. & Singh, R.P. New virulence within race TTKS (Ug99) of the stem rust pathogen and effective resistance genes. *Phytopathology* **97**, S137-S137 (2007).
14. Olivera, P., *et al.* Phenotypic and genotypic characterization of race TKTTF of *Puccinia graminis* f. sp. *tritici* that caused a wheat stem rust epidemic in southern Ethiopia in 2013–14. *Phytopathology* **105**, 917-928 (2015).
15. Olivera, P.D., *et al.* Presence of a Sexual Population of Puccinia graminis f. sp. tritici in Georgia Provides a Hotspot for Genotypic and Phenotypic Diversity. *Phytopathology* **109**, 2152-2160 (2019).
16. Upadhyaya, N.M., *et al.* Comparative genomics of Australian isolates of the wheat stem rust pathogen Puccinia graminis f. sp tritici reveals extensive polymorphism in candidate effector genes. *Frontiers in Plant Science* **5** (2015).
17. Zhang, J., Zhang, P., Karaoglu, H. & Park, R.F. Molecular Characterization of Australian Isolates of Puccinia graminis f. sp. tritici Supports Long-Term Clonality but also Reveals Cryptic Genetic Variation. *Phytopathology™* **107**, 1032-1038 (2017).
18. Lewis, C.M., *et al.* Potential for re-emergence of wheat stem rust in the United Kingdom. *Communications Biology* **1**, 13 (2018).
19. Visser, B., *et al.* Microsatellite Analysis and Urediniospore Dispersal Simulations Support the Movement of *Puccinia graminis* f. sp. *tritici* from Southern Africa to Australia. *Phytopathology* **109**, 133-144 (2019).
20. Visser, B., Herselman, L. & Pretorius, Z.A. Genetic comparison of Ug99 with selected South African races of Puccinia graminis f.sp tritici. *Molecular Plant Pathology* **10**, 213-222 (2009).
21. Lee, W.S., Hammond-Kosack, K.E. & Kanyuka, K. Barley Stripe Mosaic Virus-Mediated Tools for Investigating Gene Function in Cereal Plants and Their Pathogens: Virus-Induced Gene

Silencing, Host-Mediated Gene Silencing, and Virus-Mediated Overexpression of Heterologous Protein. *Plant Physiology* **160**, 582-590 (2012).
22. Qutob, D., *et al.* Copy number variation and transcriptional polymorphisms of Phytophthora sojae RXLR effector genes Avr1a and Avr3a. *PLoS ONE* **4**, e5066 (2009).
23. Pais, M., *et al.* Gene expression polymorphism underpins evasion of host immunity in an asexual lineage of the Irish potato famine pathogen. *BMC Evolutionary Biology* **18**, 93 (2018).
24. Wu, W., *et al.* Flax rust infection transcriptomics reveals a transcriptional profile that may be indicative for rust Avr genes. *PLOS ONE* **14**, e0226106 (2019).
25. Steuernagel, B., *et al.* Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat Biotechnol* (2016).
26. Appels, R., *et al.* Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, eaar7191 (2018).
27. Saur, I.M.L., Bauer, S., Lu, X.L. & Schulze-Lefert, P. A cell death assay in barley and wheat protoplasts for identification and validation of matching pathogen AVR effector and plant NLR immune receptors. *Plant Methods* **15** (2019).
28. Luo, M., *et al.* A five-transgene cassette confers broad-spectrum resistance to a fungal rust pathogen in wheat. *Nat Biotechnol* **https://doi.org/10.1038/s41587-020-00770-x** (2020).
29. Mago, R., *et al.* The wheat Sr50 gene reveals rich diversity at a cereal disease resistance locus. *Nature Plants* **1** (2015).
30. Marais, G.F. An evaluation of three Sr27-carrying wheat × rye translocations. *South African Journal of Plant and Soil* **18**, 135-136 (2001).
31. Mago, R., *et al.* Transfer of stem rust resistance gene SrB from Thinopyrum ponticum into wheat and development of a closely linked PCR-based marker. *Theor Appl Genet* **132**, 371-382 (2019).
32. Wulff, B.B., Horvath, D.M. & Ward, E.R. Improving immunity in crops: new tactics in an old game. *Curr Opin Plant Biol* **14**, 468-76 (2011).
33. Hubbard, A., *et al.* Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. *Genome Biology* **16**, 23 (2015).
34. Park, R.F. Stem rust of wheat in Australia. *Australian Journal of Agricultural Research* **58**, 558-566. (2007).
35. Singh, S.J. & Mcintosh, R.A. Allelism of 2 Genes for Stem Rust Resistance in Triticale. *Euphytica* **38**, 185-189 (1988).
36. Acosta, A.C. The transfer of stem rust rsistance from rye to wheat. *Dissertation Abstracts* **23**, 34-35 (1962).
37. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv*, 1207.3907 (2012).
38. Stamatakis, A., Ludwig, T. & Meier, H. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**, 456-463 (2005).
39. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290 (2004).
40. Heibl, C. PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic software packages. *Avaliable online at: http://www. christophheibl. de/Rpackages. html* (2008).
41. Wickham, H. ggplot2: elegant graphics for data analysis. 2016: Springer.
42. Yin, T., Cook, D. & Lawrence, M. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biology* **13**, R77 (2012).
43. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35**, 1547-1549 (2018).

44. Lee, W.S., Rudd, J.J. & Kanyuka, K. Virus induced gene silencing (VIGS) for functional analysis of wheat genes involved in Zymoseptoria tritici susceptibility and resistance. *Fungal Genetics and Biology* **79**, 84-88 (2015).
45. Franco-Orozco, B*., et al.* A new proteinaceous pathogen-associated molecular pattern (PAMP) identified in Ascomycete fungi induces cell death in Solanaceae. *New Phytologist* **214**, 1657-1672 (2017).
46. Hen-Avivi, S*., et al.* A Metabolic Gene Cluster in the Wheat W1 and the Barley Cer-cqu Loci Determines beta-Diketone Biosynthesis and Glaucousness. *Plant Cell* **28**, 1440-1460 (2016).
47. Lee, W.S., Rudd, J.J., Hammond-Kosack, K.E. & Kanyuka, K. Mycosphaerella graminicola LysM Effector-Mediated Stealth Pathogenesis Subverts Recognition Through Both CERK1 and CEBiP Homologues in Wheat. *Molecular Plant-Microbe Interactions* **27**, 236-243 (2014).
48. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* **14**, 417-419 (2017).
49. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15** (2014).
50. Mago, R*., et al.* Generation of Loss-of-Function Mutants for Wheat Rust Disease Resistance Gene Cloning. *Wheat Rust Diseases: Methods and Protocols* **1659**, 199-205 (2017).
51. Mago, R*., et al.* Development of PCR markers for the selection of wheat stem rust resistance genes Sr24 and Sr26 in diverse wheat germplasm. *Theoretical and Applied Genetics* **111**, 496-504 (2005).
52. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
53. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics* **25**, 1754-1760 (2009).
54. Li, H*., et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
55. Rabanus-Wallace, M.T*., et al.* Chromosome-scale genome assembly provides insights into rye biology, evolution, and agronomic potential. *bioRxiv*, 2019.12.11.869693 (2019).
56. Akamatsu, A*., et al.* An OsCEBiP/OsCERK1-OsRacGEF1-OsRac1 Module Is an Essential Early Component of Chitin-Induced Rice Immunity. *Cell Host & Microbe* **13**, 465-476 (2013).
57. Arndell, T*., et al.* gRNA validation for wheat genome editing with the CRISPR-Cas9 system. *BMC Biotechnology* **19**, 71 (2019).
58. Bemoux, M*., et al.* Structural and Functional Analysis of a Plant Resistance Protein TIR Domain Reveals Interfaces for Self-Association, Signaling, and Autoregulation. *Cell Host & Microbe* **9**, 200-211 (2011).
59. Cesari, S*., et al.* Cytosolic activation of cell death and stem rust resistance by cereal MLA-family CC-NLR proteins. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 10204-10209 (2016).
60. Cesari, S*., et al.* The rice resistance protein pair RGA4/RGA5 recognizes the Magnaporthe oryzae effectors AVR-Pia and AVR1-CO39 by direct binding. *Plant Cell* **25**, 1463-81 (2013).
61. Richardson, T., Thistleton, J., Higgins, T.J., Howitt, C. & Ayliffe, M. Efficient Agrobacterium transformation of elite wheat germplasm without selection. *Plant Cell, Tissue and Organ Culture (PCTOC)* **119**, 647-659 (2014).
62. Ishida, Y., Tsunashima, M., Hiei, Y. & Komari, T. Wheat (Triticum aestivum L.) Transformation Using Immature Embryos, in *Agrobacterium Protocols, Vol 1, 3rd Edition*. 2015, Springer: New York. p. 189-198.
63. Sperschneider, J*., et al.* The stem rust fungus *Puccinia graminis* f. sp. *tritici* induces centromeric small RNAs during late infection that direct genome-wide DNA methylation. *BioRXiv*, 469338 (2020).

Acknowledgements. Work described here was supported by funding from the 2Blades foundation. JC was supported by a Chinese Scholarship Council (CSC) postgraduate fellowship. JS was supported by an Australian Research Council Discovery Early Career Researcher Award 2019 (DE190100066). KK and VP acknowledge financial support by the Institute Strategic Program Grant 'Designing Future Wheat' (BB/P016855/1) from the Biotechnology and Biological Sciences Research Council of the UK (BBSRC). HN-P was supported by USDA-NIFA grant #2018-67013-27819.

Author contributions. P.N.D. conceptualized the project, acquired funding and supervised the work. N.M.U, R.M., V.P., M.L., A.W., D.O., J.C., J.Z., D.B., M.A. and L.H. acquired experimental data. N.M.U., R.M., J.S., H.N-P., T.H., M.F., K.K., J.G.E, and P.N.D, conducted data analysis. P.N.D. drafted the manuscript and all authors contributed to review and editing.

Competing interests. The authors declare no competing interests.

Data availability. All sequence data from this study are available in NCBI under BioProject PRJNA695305 (Sr27 MutRenSeq data) and PRJNA698655 (Pgt21-0 mutants). This includes raw data for Figure 2B, S2 and S8. Figure 3 is derived from raw sequence data available in NCBI BioProject PRJNA253722 and PRJNA415866. All other relevant data is available upon request from the corresponding author.

Code availability. Scripts and files for MutRenSeq analysis are available at https://github.com/TC-Hewitt/MuTrigo.

589 **Figure legends**

590 **Fig. 1 *Pgt* mutants with virulence to *Sr27*.**

591 **a,** Rust strains Pgt21-0 and a spontaneous mutant (21m1) selected for virulence on *Sr27* were
592 inoculated onto seedlings of triticale line Coorong containing *Sr27*. Image taken 14 days post
593 inoculation. **b,** Read coverage graphs for chromosome 2B (orange bar) of Illumina DNA
594 sequence reads from Pgt21-0 and three *Sr27*-virulent mutants of Pgt21-0 (21m1, 21m2 and
595 21m3). Y-axes represents absolute read depth (0 to 100). Position on the chromosome is
596 indicated in 1 Mbp intervals from the 5' end with the centromere (CM) position [63] indicated.
597 Approximate sizes of the deleted regions in mutant lines are shown in kbp. **c,** Heat map showing
598 percentage of read coverage for all annotated genes in the *AvrSr27* region of chromosome 2B
599 in Illumina reads from Pgt21-0, the spontaneous mutants 21m1, 21m2 and 21m3, four
600 Australian field isolates (98-1,2,3,5,6; 194-1,2,3,5,6; 34,2,12 and 34,2,12,13) and seven South
601 African isolates (SA-01 to 07). The virulence phenotype of each strain on *Sr27* is indicated,
602 with avirulent isolates listed red and virulent ones in blue. Genes encoding secreted proteins
603 are indicated in bold, and the orange box indicates the two genes included in the deletion region.

604 **Fig. 2 The *AvrSr27* locus encodes two related secreted proteins.**

605 **a,** Schematic of gene arrangement at the *AvrSr27* locus and amino acid alignment of AvrSr27
606 protein variants from Pgt21-0. AvrSr27-1 and AvrSr27-2 are encoded at the avirulence allele
607 on chromosome 2B, while avrSr27-3 is encoded at the virulence allele on chromosome 2A.
608 Amino acid sequence (single letter code) of AvrSr27-1 is shown with identical residues in other
609 variants indicated by a dot. The predicted signal peptide region is underlined and in bold. **b,**
610 Infection of triticale lines Coorong (contains *Sr27*), Rongcoo (susceptible) and a Coorong
611 derived mutant that has lost *Sr27* (M1) with the *BSMV* expression vector encoding the predicted
612 mature AvrSr27-1, -2, or -3 proteins, or containing a non-coding multiple cloning site (empty
613 vector). Representative images of the infection phenotypes are shown for each treatment with
614 the numbers of systemically infected plants out of the total number of plants inoculated under
615 each image. Data represent total numbers from two to four independent experiments each
616 involving 10-12 plants per treatment.

617 **Fig. 3 Differential expression of *AvrSr27* alleles.**

a, Expression levels of the three *AvrSr27* variants in haustoria and germinated spores of Pgt21-0 and in infected plants at 5, 6, and 7 days post infection. Expression levels were derived from mapping RNAseq reads from each of the tissue types (3 biological replicates) to annotated genes of Pgt21-0 and shown as transcripts per million (y-axis). Significant differences are indicated (* $p<0.05$; ** $p<0.01$). **b,** Expression profiling of all secreted proteins of Pgt21-0. Clusters with different expression patterns (numbered 1 to 8) are indicated. Positions of *AvrSr27*, *AvrSr35* and *AvrSr50* are indicated. Blue colour intensity indicates relative expression levels (relative rlog transformed counts; darker colour indicates higher expression) in haustoria (H), germinated spores (GS), and infected leaves 2, 3, 4, 5, 6, or 7 days post infection.

Fig. 4 Identification of the *Sr27* resistance gene.

a, Infection phenotype of Coorong (left) and the EMS-derived susceptible mutant #1 (right) inoculated with Pgt21-0 9 days after infection. **b,** Schematic diagram of the Sr27 protein with the amino acid changes in mutants M1, M3, M4 and M6 indicated. **c,** Luciferase activity (luminescence units y-axis) detected in protoplasts co-expressing *Sr27* or *AvrSr27* variants 1 to 3 alone, *Sr27* and *AvrSr27* variants in combination, *Sr50* plus *AvrSr27-1* or *Sr27* plus *AvrSr50*. **d,** Transient expression in *N. benthamiana* of Sr27:YFP and YFP:AvrSr27 variants alone or in combination. YFP alone and the autoactive coiled-coil domain of Sr50 (CCSr50-YFP) were used as negative and positive controls respectively. Images taken 4 days post-infiltration. **e,** Phenotypes of selected T1 lines and control plants infected with stem rust race 98-1,2,3,5,6, and scored 10 days post infection. Image shows an infected leaf from the susceptible Chinese Spring, WRT238.5 containing the native *Sr27* gene, the susceptible parent Fielder and representative T1 plants from two transgenic lines (#3 and #17) containing the *Sr27* transgene and a non-transgenic line recovered from tissue culture (#16).

Supplementary Figures:

Supplementary Fig. 1 Infection phenotypes of Pgt21-0 and three spontaneous mutants (M1, M2 and M3) on Triticale lines Coorong (contains *Sr27*), Rongcoo (rust susceptible) and a mutant line derived from Coorong with a loss of *Sr27* resistance gene (Sr27Mutant1). Image was taken 14 days after inoculation of seedling leaves.

Supplementary Fig. 2 Field isolates of *Pgt* with virulence for *Sr27* contain small deletions on chromosome 2B. **a,** Illumina read coverage graphs for the *AvrSr27* region of chromosome 2B (orange bar) for Pgt21-0, three *Sr27*-virulent mutants of Pgt21-0 and nine field isolates of the same clonal lineage; 34,2,12 and 34,2,12,13 from Australia and SA01 to SA07 from South Africa. Isolates avirulent on *Sr27* are listed in red and virulent isolates in blue. Position on the chromosome in kbp is indicated above the graphs. **b,** Close-up of read coverage graphs in boxed region of (A). Approximate sizes of the deleted regions in virulent field isolates are shown in kbp. Positions of Pgt21_006532 and PGT21_006593 on chromosome 2B reference are indicated.

Supplementary Fig. 3 Confirmation of the 13.2 kbp deletion in *Sr27*-virulent rust strain 34-2-12. The positions of primers P423, P424 and P426 on chromosome 2B around the *AvrSr27* locus are indicated (blue arrow heads) relative to the boundaries of the deletion region in 34-2-12 inferred from the genomic sequence reads. PCR amplification products from genomic DNA of Pgt21-0 and 34-2,12 are shown after separation on a 1% agarose gel. The primers P383 and P353 are designed to amplify a fragment of the *AvrSr50* gene as a control region that is identical in both isolates.

Supplementary Fig. 4 Independent deletions of *AvrSr27* in the race 21 clonal lineage of *Pgt*. Phylogenetic analysis of *Pgt* isolates of the race 21 lineage from South Africa, Australia and the Czech Republic (indicated in colour key) using a RAxML model and biallelic SNPs called against the full dikaryotic genome of *Pgt*21-0. Blue arrowheads indicate where three independent mutations leading to virulence on *Sr27* occurred within this lineage. Scale bar indicates number of nucleotide substitutions per site. Nodes with bootstrap values greater than 85% are indicated by green triangles.

Supplementary Fig. 5 Infection of Triticale lines with BSMV constructs. **a,** RT-PCR assay to check the accumulation of BSMV in Coorong, Rongcoo and *Sr27* mutant plants. RNA

extracted from leaf samples collected at 14 days post BSMV inoculation was amplified using primers flanking the cloning site in BSMV. +/- indicates the presence or absence of virus symptoms in plants challenged with the respective BSMV construct and buffer control. pDNA of the respective BSMV constructs used as positive control. **b,** Infection of wheat lines Chinese Spring (CS) and CS WRT238.5 carrying *Sr27* on an introgressed 3RS chromosome segment with the *BSMV* expression vector encoding *AvrSr27-1*, *-2*, *-3*, or a non-coding multiple cloning site (empty). Y-axis indicates the proportion of inoculated plants that demonstrated systemic viral infection symptoms.

Supplementary Fig. 6 Differential expression of *AvrSr27* alleles. RT-PCR analysis of gene expression in Pgt21-0 and the *Sr27* virulent mutant M1. RT-PCR was performed on three replicate samples (R1 to R3) of RNA extracted from wheat infected with Pgt21-0 or the virulent mutant line (M1). Primers used targeted the transcripts from genes *PGT21-006334* (included in the deleted region of mutant 1, left lanes, primers P373/341), *AvrSr50* (not deleted in mutant 1, middle lanes, primers P383/P351) or all three *AvrSr27* variants (two of which are deleted in mutant 1, right lanes, primers P420/P414/P415).

Supplementary Fig. 7 Schematic alignment of *AvrSr27* locus haplotypes from the A, B and C genomes (blue, orange and green respectively) of Pgt21-0 and Ug99. Regions of high sequence similarity (>95%) between the haplotypes are indicated by grey shading. The positions of *AvrSr27* coding sequences are indicated by the dark blue boxes. *AvrSr27-1** in haplotype C encodes a single amino acid change compared to *AvrSr27-1* on chromosome 2B. Chromosome and contig positions of the selected genome segments are indicated (bp).

Supplementary Fig. 8 Detection of mutations in a candidate *Sr27* gene by NB-LRR capture and sequencing. Two contigs (#5723 and #2413) assembled from wildtype Coorong contain the 5' and 3' regions of this gene. Graphs show the read coverage per base of reads derived from the NB-LRR capture library from Coorong (wild-type) and four mutants (2,3,4 and 6) mapped to these two contigs. The positions of single nucleotide changes are shown by the coloured bars with the specific change indicated. Mutant M2 produced no reads specific to these contigs and therefore likely contains a deletion. The positions of conserved motifs are indicated under the graphs (numbered colored bars).

Supplementary Fig. 9 Maximum likelihood phylogenetic tree comparing Sr27 amino acid sequence to proteins sequences of known wheat resistance proteins. Scale bar shows amino acid sequence divergence.

Supplementary Fig. 10 Amino acid sequence alignment of the Sr27 and Sr13a resistance proteins.

Supplementary Fig. 11 Immunoblot showing protein expression of Sr27-YFP and YFP-AvrSr27 protein constructs detected using anti-GFP antibodies. Ponceau red staining of filter indicates equal loading of protein extracts.

720 Supplementary Table 1 Phenotypes of transgenic wheat lines with stem, leaf and stripe rust
721 infection.

722

T1 Family	Transgene	Stem rust: 98-1,2,3,5,6			Stripe rust: 134E16A+17+33+			Leaf rust: 76-1,3,7,9,10,12,13+Lr37		
		No. T1 seedlings	R	S	No. T1 seedlings	R	S	No. T1 seedlings	R	S
PC311-3	+	11	11	0	4	0	4	6	0	6
PC311-16	-	11	0	11	4	0	4	6	0	6
PC311-17	+	12	12	0	5	0	5	5	0	5
PC311-18	+	11	10	1	4	0	4	6	0	6
PC311-Control	-	11	0	11	5	0	5	6	0	6

723

724

Supplementary Table 2 Primer sequences used

Primer name	Sequence 5'> 3'	Target
Sr27F	CCTGTTCGATCACTGGTCG	Sr27 forward
Sr27F2	GTGAAGATGGTCTGCATTGTTGGATCG	Sr27
Sr27R1	GATGGTATATACCGTGGTCCGACAAAT	Sr27 reverse
Sr27R2	CGGAGGTTAAGCGGCGGAGA	Sr27
Sr27R3	GGTTTTGTGTGCATAGTTTACCAAGAG	Sr27
Sr27c5723F	ATGAACCCAAGGCTGAGTTG	Sr27
Sr27c5723R	ACATGCAAATAAGGGCTTCC	Sr27
Sr27c2413F	GTAAGGCTCCAAGGATGCAG	Sr27
Sr27c2413R	TAAGTTTCCCGACGGAATTG	Sr27
Sr27c5723ExtF1	AAGAACAAGTGAAGATGGTCTGC	Sr27 3' end contig 5723
Sr27c2413ExtR1	TTCTGTAGAATAGTTGTCTTGAGTGCTC	Sr27 5' end contig 2413
Sr27c2413ExtF1	TACGAGGAAGCGAAAAACAGC	Sr27
Sr27cSeq1R	ATCTTCTCAGTGAAGCCATC	Sr27
Sr27cSeq2R	AAATGTGACTTGATACATCTG	Sr27
Sr27cSeq3F	ATAGTGATTGACGACATATGG	Sr27
Sr27cSeq4F	GATTCATTCGACAAGAAGGT	Sr27
P341	ATTCAGATTTAAGAGTCTTGATTGAGTCCCCATG	PGT21_006334 reverse
P373	CACCATGCAATTAGCCAGTGTCTTATGTG	PGT21_006334 forward
P351	GTCTTCCTACCTGTGTTGGCGCCTTGCAAAATG	AvrSr50 reverse primer
P383	CACCATGATGCATTCAATTATCTTTCAAACACTCC	AvrSr50 forward primer
P414	TTACCATCTTCTGTGACACTCTGGG	AvrSr27-2 reverse
P415	TTACCATCTGCTGTGACACTCTGG	AvrSr27-1/-3 reverse
P416	CACCATGGCAATGACACCACATCACCAAAGCAAT	AvrSr27-1/-2 27 aa signal peptide clipped CDS 5' with CACCATG for directional Topo cloning
P417	CACCATGGCAATGACACCACATCACCAAATCAAT	avrSr27-3 27 aa signal peptide clipped CDS 5' with CACCATG for directional Topo cloning
P418	CCATCTTCTGTGACACTCTGGGCTTG	AvrSr27-2 CDS 3' without stop codon
P419	CCATCTGCTGTGACACTCTGGGCTTG	AvrSr27-1/-3 CDS 3' without stop codon
P420	CACCATGCATTACATCACCCCCATAATCCTT	AvrSr27-1/-2/-3 forward
P423	AAGTGGATAACGTACTCTGCACAAC	upstream of 5' deletion boundary in 34M1/34M2 mutants
P424	AGTGACTGCAATTCACCAATATTTCG	Internal to deletion region in 34M1/34M2 mutants
P426	AACATTCAGTGCGAGGAATGGGGAG	downstream of 3' boundary region in 34M1/34M2 mutants
2-210.F	CCAACCCAGGACCGTTGATGGCAATGACACCACATCACCAAA	Amplification of AvrSr27-1, -2, and -3 for cloning into BSMV VOX vector
2-210.R	AACCACCACCACCGTTACCATCTGCTGTGACACTCTG	Amplification of AvrSr27-1 and -3 for cloning into BSMV VOX vector
5-210.R	AACCACCACCACCGTTACCATCTTCTGTGACACTCTGG	Amplification of AvrSr27-2 for cloning into BSMV VOX vector

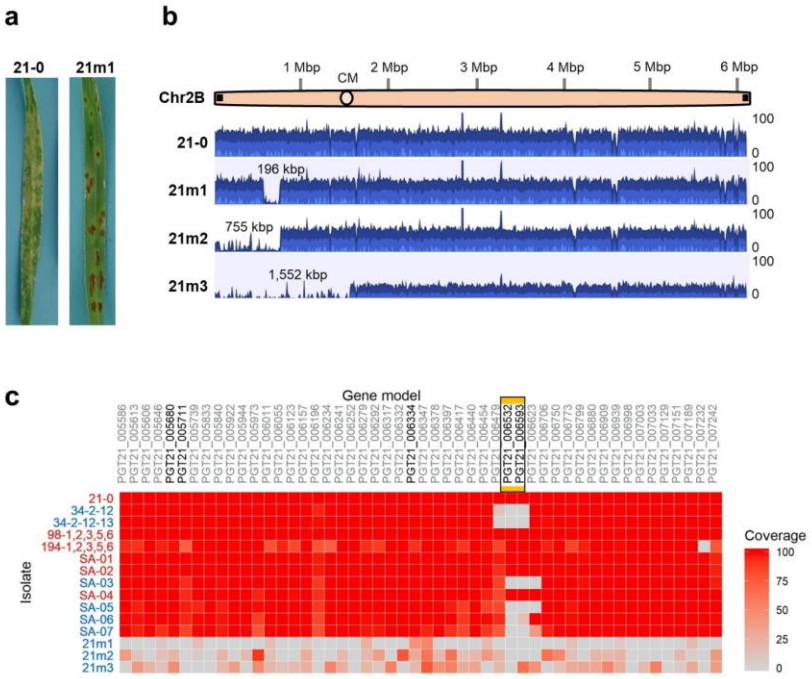

Fig. 1 *Pgt* mutants with virulence to *Sr27*.

a, Rust strains Pgt21-0 and a spontaneous mutant (M1) selected for virulence on *Sr27* were inoculated onto seedlings of triticale line Coorong containing *Sr27*. Image taken 14 days post inoculation. **b,** Read coverage graphs for chromosome 2B (orange bar) of Illumina DNA sequence reads from Pgt21-0 and three *Sr27*-virulent mutants of Pgt21-0 (M1, M2 and M3). Y-axes represents absolute read depth (0 to 100). Position on the chromosome is indicated in 1 Mbp intervals from the 5' end with the centromere (CM) position indicated (Sperschneider et al 2020). Approximate sizes of the deleted regions in M1, M2 and M3 are shown in kbp. **c,** Heat map showing percentage of read coverage for all annotated genes in the *AvrSr27* region of chromosome 2B in Illumina reads from Pgt21-0, the spontaneous mutants M1, M2 and M3, four Australian field isolates (98-1,2,3,5,6; 194-1,2,3,5,6; 34,2,12 and 34,2,12,13) and seven South African isolates (SA-01 to 07). The virulence phenotype of each strain on *Sr27* is indicated, with avirulent isolates listed red and virulent ones in blue. Genes encoding secreted proteins are indicated in bold, and the orange box indicates the two genes included in the deletion region.

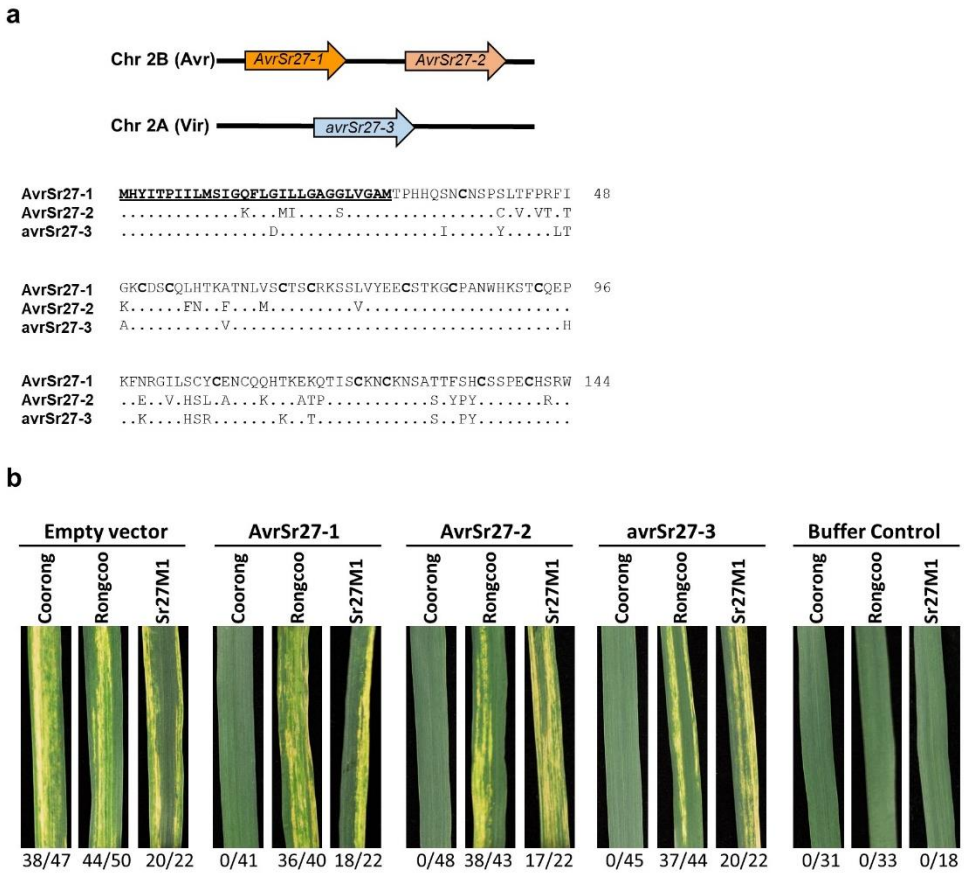

Fig. 2 The *AvrSr27* locus encodes two related secreted proteins.

a, Schematic of gene arrangement at the *AvrSr27* locus and amino acid alignment of AvrSr27 protein variants from Pgt21-0. AvrSr27-1 and AvrSr27-2 are encoded at the avirulence allele on chromosome 2B, while avrSr27-3 is encoded at the virulence allele on chromosome 2A. Amino acid sequence (single letter code) of AvrSr27-1 is shown with identical residues in other variants indicated by a dot. The predicted signal peptide region is underlined and in bold. **b,** Infection of triticale lines Coorong (contains *Sr27*), Rongcoo (susceptible) and a Coorong derived mutant that has lost *Sr27* (M1) with the *BSMV* expression vector encoding the predicted mature AvrSr27-1, -2, or -3 proteins, or containing a non-coding multiple cloning site (empty vector). **c,** Infection of wheat lines Chinese Spring (CS) and CS WRT238.5 (carrying *Sr27* on an introgressed 3RS chromosome segment) with the *BSMV* expression vector encoding AvrSr27-1, -2, -3, or a non-coding multiple cloning site (empty vector). Representative images of the infection phenotypes are shown for each treatment with the numbers of systemically infected plants out of the total number of plants inoculated under each image. Data represent total numbers from two to four independent experiments each involving 10-12 plants per treatment

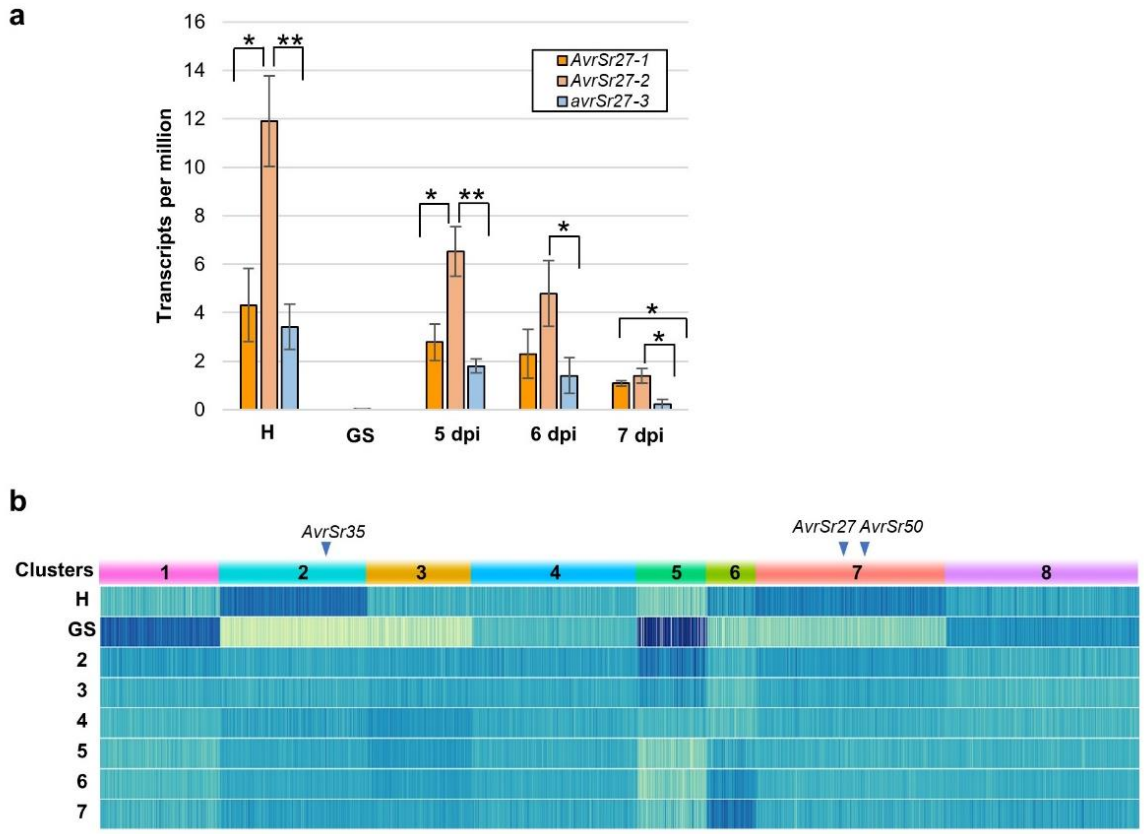

Fig. 3 Differential expression of *AvrSr27* alleles.

a, Expression levels of the three *AvrSr27* variants in haustoria and germinated spores of Pgt21-0 and in infected plants at 5, 6 and 7 days post infection. Expression levels were derived from mapping RNAseq reads from each of the tissue types (3 biological replicates) to annotated genes of Pgt21-0 and shown as transcripts per million (y-axis). Significant differences are indicated (* $p<0.05$; ** $p<0.01$). **b,** Expression profiling of all secreted proteins of Pgt21-0. Clusters with different expression patterns (numbered 1 to 8) are indicated. Positions of *AvrSr27*, *AvrSr35* and *AvrSr50* are indicated. Blue colour intensity indicates relative expression levels (relative rlog transformed counts; darker colour indicates higher expression) in haustoria (H), germinated spores (GS), and infected leaves 2, 3, 4, 5, 6, or 7 days post infection.

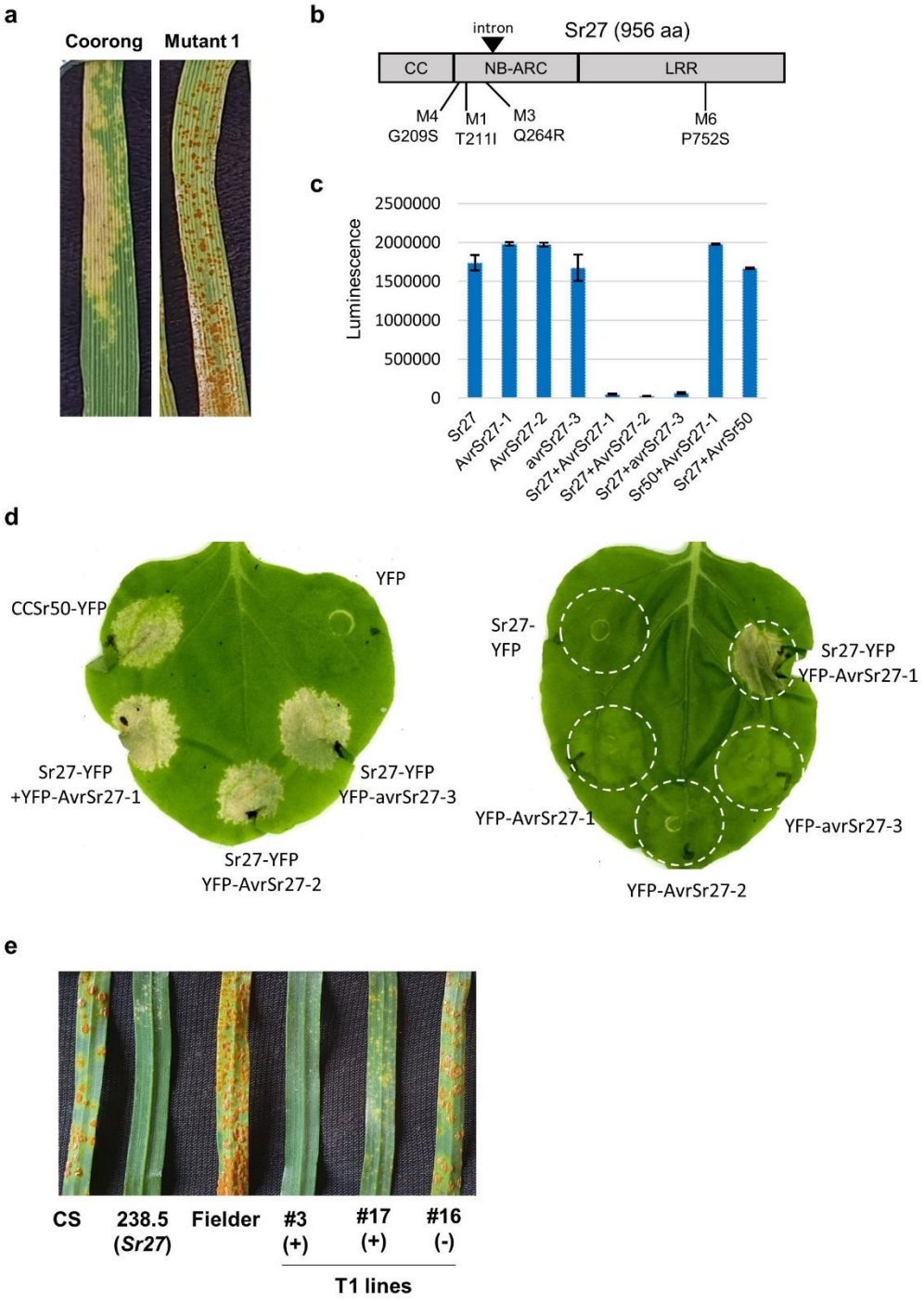

Fig. 4 Identification of the *Sr27* resistance gene.

a, Infection phenotype of Coorong (left) and the EMS-derived susceptible mutant #1 (right) inoculated with Pgt21-0 9 days after infection. **b,** Schematic diagram of the Sr27 protein with the amino acid changes in mutants M1, M3, M4 and M6 indicated. **c,** Luciferase activity (luminescence units y-axis) detected in protoplasts co-expressing *Sr27* or *AvrSr27* variants 1 to 3 alone, *Sr27* and *AvrSr27* variants in combination, *Sr50* plus *AvrSr27-1* or *Sr27* plus *AvrSr50*. **d,** Transient expression in *N. benthamiana* of Sr27:YFP and YFP:AvrSr27 variants alone or in combination. YFP alone and the autoactive coiled-coil domain of Sr50 (CCSr50-YFP) were used as negative and positive controls respectively. Images taken 4 days post-infiltration. **e,** Phenotypes of selected T1 lines and control plants infected with stem rust race

98-1,2,3,5,6, and scored 10 days post infection. Image shows an infected leaf from the susceptible Chinese Spring, WRT238.5 containing the native *Sr27* gene, the susceptible parent Fielder and representative T1 plants from two transgenic lines (#3 and #17) containing the *Sr27* transgene and a non-transgenic line recovered from tissue culture (#16).