HAL open science

The stem rust effector protein AvrSr50 escapes Sr50 recognition by a substitution in a single surface-exposed residue

Diana Ortiz, Jian Chen, Megan Outram, Isabel Saur, Narayana Upadhyaya, Rohit Mago, Daniel Ericsson, Stella Cesari, Chunhong Chen, Simon Williams, et al.

To cite this version:

Diana Ortiz, Jian Chen, Megan Outram, Isabel Saur, Narayana Upadhyaya, et al.. The stem rust effector protein AvrSr50 escapes Sr50 recognition by a substitution in a single surface-exposed residue. New Phytologist, 2022, 234 (2), pp.592-606. 10.1111/nph.18011 . hal-03610628

HAL Id: hal-03610628

https://hal.inrae.fr/hal-03610628

Submitted on 16 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

New Phytologist

Research

The stem rust effector protein AvrSr50 escapes Sr50 recognition by a substitution in a single surface-exposed residue

Diana Ortiz[1,2]*, Jian Chen[1,3]*, Megan A. Outram[3]*, Isabel M.L. Saur[4,5,6], Narayana M. Upadhyaya[1], Rohit Mago[1], Daniel J. Ericsson[3,7], Stella Cesari[8], Chunhong Chen[1], Simon J. Williams[3] and Peter N. Dodds[1]

[1]Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation, Canberra, ACT 2601, Australia; [2]National Research Institute for Agriculture, Food and Environment, Genetics and Breeding of Fruit and Vegetables Unit, Montfavet 84143, France; [3]Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia; [4]Department of Plant–Microbe Interactions, Max Planck Institute for Plant Breeding Research, Cologne 50829, Germany; [5]University of Plant Sciences, University of Cologne, Cologne 50674, Germany; [6]Cluster of Excellence on Plant Sciences, Cologne 50674, Germany; [7]Australian Synchrotron, Macromolecular Crystallography,Clayton, Vic. 3168, Australia; [8]PHIM Plant Health Institute, Université de Montpellier, INRAE, CIRAD, Institut Agro, IRD, Montpellier 34980, France

Authors for correspondence:
Simon J. Williams
Email: *simon.williams@anu.edu.au*

Peter N. Dodds
Email: *peter.dodds@csiro.au*

Received: *26 July 2021*
Accepted: *12 January 2022*

New Phytologist (2022)
doi: 10.1111/nph.18011

Key words: AvrSr50, effector evolution, effector structure, resistance breakdown, Sr50, stem rust, wheat resistance.

Summary

- Pathogen effectors are crucial players during plant colonisation and infection. Plant resistance mostly relies on effector recognition to activate defence responses. Understanding how effector proteins escape from plant surveillance is important for plant breeding and resistance deployment.
- Here we examined the role of genetic diversity of the stem rust (*Puccinia graminis* f. sp. *tritici* (*Pgt*)) *AvrSr50* gene in determining recognition by the corresponding wheat *Sr50* resistance gene. We solved the crystal structure of a natural variant of AvrSr50 and used site-directed mutagenesis and transient expression assays to dissect the molecular mechanisms explaining gain of virulence.
- We report that AvrSr50 can escape recognition by Sr50 through different mechanisms including DNA insertion, stop codon loss or by amino-acid variation involving a single substitution of the AvrSr50 surface-exposed residue Q121. We also report structural homology of AvrSr50 to cupin superfamily members and carbohydrate-binding modules indicating a potential role in binding sugar moieties.
- This study identifies key polymorphic sites present in *AvrSr50* alleles from natural stem rust populations that play important roles to escape from Sr50 recognition. This constitutes an important step to better understand *Pgt* effector evolution and to monitor AvrSr50 variants in natural rust populations.

Introduction

During a lifetime, plants may encounter many diverse microbes that can affect their growth and development in beneficial or detrimental ways. To maintain their health, plants have evolved an immune system that recognises pathogenic microbes to then trigger an arsenal of defence responses (Jones & Dangl, 2006). Plant immunity relies on extracellular and intracellular receptor proteins that work together to sense either conserved or specific virulence factors in the pathogen and to efficiently activate disease-resistance responses (Jones *et al.*, 2016; Tang *et al.*, 2017; Ngou *et al.*, 2021). Nucleotide-binding leucine-rich repeat receptors (NLRs) are largely responsible for mediating the intracellular surveillance and constitute the most abundant family of disease-resistance proteins identified so far (Jacob *et al.*, 2013). Nucleotide-binding leucine-rich repeat receptors are observed throughout aquatic and land plants and appear to have originated and rapidly diverged before the split of green algae (Gao *et al.*, 2018; Ortiz & Dodds, 2018; Shao *et al.*, 2019).

Nucleotide-binding leucine-rich repeat receptors share a common modular structure with a central nucleotide-binding site (NBS) and C-terminal leucine-rich repeat (LRR) domains, but are separated into three major groups depending on the presence of a Toll-interleukin 1 receptor (TIR), a coiled-coil (CC) or a RPW8/HeLo-like domain at the N-terminus of the proteins (Jubic *et al.*, 2019). Nucleotide-binding leucine-rich repeat receptors recognise specific pathogen secreted proteins, called effectors, through two main strategies; direct binding to the effector (Saur *et al.*, 2021) or indirectly by detecting either a modification of the effector host target protein, 'guardee', or modifications of a host protein that mimics the effector target, 'decoy' (van der Hoorn & Kamoun, 2008).

*These authors contributed equally to this work.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Stem rust disease, caused by the fungus *Puccinia graminis* f. sp. *tritici* (*Pgt*), is considered one of the main threats for wheat production and food security around the world. During the last 2 decades, the emergence of new *Pgt* virulent races, especially Ug99 detected in Uganda (1999) and virulent on most of the wheat cultivars grown in the field, has prompted the study and identification of effective stem rust resistance genes (Singh *et al.*, 2011). Two main classes of genes, race-specific (NLRs; functional from seedling to adult growth stages) or adult plant resistance (APRs; functioning mainly in the adult stage), are typically used by wheat breeders to prevent plant disease (Ellis *et al.*, 2014). With the exception of *Sr60*, which encodes a protein with two putative kinase domains (Chen *et al.*, 2020), so far, all cloned race-specific stem rust resistance genes encode NLR proteins with a N-terminal CC domain (Periyannan *et al.*, 2013; Saintenac *et al.*, 2013; Mago *et al.*, 2015; Steuernagel *et al.*, 2016; Zhang *et al.*, 2017, 2021; Chen *et al.*, 2018; Arora *et al.*, 2019; Upadhyaya *et al.*, 2021). One approach to better understand how these immune receptors work and how they should be deployed in the field, is the identification of the matching effectors and their mode of recognition.

To date three *Pgt* avirulence effector proteins, AvrSr50, AvrSr35 and AvrSr27, have been cloned and recognition by the corresponding NLR proteins Sr50, Sr35 and Sr27 has been demonstrated (Chen *et al.*, 2017; Salcedo *et al.*, 2017; Upadhyaya *et al.*, 2021). The *Sr50* resistance gene is derived from rye and is present in wheat as a translocated segment of rye chromosome 1 but has not been widely deployed in cultivated varieties (Mago *et al.*, 2015). This gene is orthologous to *Mla* (*mildew locus a*) in barley, which confers resistance to *Blumeria graminis* f. sp. *hordei* (*Bgh*). Allelic MLA proteins also bind directly to sequence-unrelated effectors from the barley powdery mildew pathogen, *B. graminis* f.sp. *hordei* (*Bgh*), indicating that direct effector recognition seems to be common in the MLA-like family of receptors (Saur *et al.*, 2019). Similarly, NLR proteins from flax also recognise flax rust avirulence proteins through direct interaction (Dodds *et al.*, 2006; Catanzariti *et al.*, 2010), suggesting that direct recognition is common in resistance to biotrophic fungi (Saur & Hückelhoven, 2021).

Structural analyses of fungal effectors can provide molecular insight into their function and interaction with host proteins, and how their sequence diversification can lead to escape from recognition. To date, the three-dimensional structures of only a few effectors from rust fungi have been determined: AvrL567 (Wang *et al.*, 2007), AvrM (Ve *et al.*, 2013) and AvrP (Zhang *et al.*, 2018) from the flax rust fungus *Melampsora lini*, and MLP124266 and MLP124017 from the poplar rust fungus *Melampsora larici-populina* (de Guillen *et al.*, 2019). For AvrL567 and AvrP, the structures were important for guiding further mutagenesis studies that identified multiple surface-exposed polymorphic residues involved in the recognition by their corresponding NLRs (Wang *et al.*, 2007; Zhang *et al.*, 2018). Currently no structures have been determined of effectors from the economically important stem rust pathogen *Pgt*.

Functional and structural characterisation of rust effector proteins is important to uncover the molecular mechanisms underlying effector gain of virulence and detection by NLR proteins. In this study, we determined the crystal structure of a virulent avrSr50 allele. Functional analysis identified a single polymorphism in a surfaced exposed amino-acid residue that enables the virulent avrSr50 variant to escape from Sr50 recognition. We further identified substantial variation in AvrSr50 alleles in a diverse collection of global *Pgt* isolates, with four virulence alleles arising from different mechanisms including insertion, amino-acid variation and introduction of a destabilising C-terminal extension fusion through stop codon loss. Sr50 confers resistance to wide array of stem rust isolates, including the Ug99 race group, and is an important source of resistance to introduce in commercial wheat crops. Collectively, our results provide valuable information about AvrSr50 diversity and loss of recognition by Sr50, which is essential for preventing the breakdown of resistance to the devastating stem rust fungus in the field.

Materials and Methods

Vector construction

For protein expression and purification, the cDNA of three AvrSr50 variants, AvrSr50^{WT}, AvrSr50^{RKQQC} and avrSr50^{QCMJC} (Chen *et al.*, 2017) from this point forwards referred to as AvrSr50-A1, AvrSr50-C and avrSr50-B6, respectively, were PCR-amplified, excluding their signal peptide as identified by SignalP-5.0 (Almagro Armenteros *et al.*, 2019), and cloned into a modified, Golden Gate-compatible, pOPIN expression vector using Golden Gate cloning (Supporting Information Table S1, Bentham *et al.*, 2021). Primers consisted of a gene-specific sequence flanked by overhangs compatible for Golden Gate cloning (Table S2) and Golden Gate digestion/ligation reactions and cycling were carried out as described previously (Iverson *et al.*, 2016). The resulting constructs contained an N-terminal His_6-tag followed by a 3C protease cleavage site to facilitate protein purification by affinity chromatography. Plasmid integrity was confirmed using Sanger sequencing.

For binding and recognition assays, all plasmids were generated using Gateway® Technology (Life Technologies™) or restriction/ligation cloning. Site-directed mutagenesis was induced using the QuickChange Lightning kit (Agilent Technologies, Santa Clara, CA, USA) or Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Waltham, MA, USA), primer sequences are detailed in Table S2. Additional natural variants of AvrSr50 were synthesised (Epoch Life Sciences, Missouri City, TX, USA) in Gateway compatible format. Gateway entry clones were generated using pENTR™/D-TOPO® and pDONR™207 vectors (Life Technologies™). Yeast-two-hybrid assays were performed with modified BD-pGBKT7 and AD-pGADT7 (Clontech, Mountain View, CA, USA) expression vectors (Bernoux *et al.*, 2011). For transient expression experiments in *Nicotiana benthamiana*, pBIN19-35S-GTW-3HA, pBIN19-35S-YFP-GTW, pBIN19-35S-HA-GTW and pAM-PAT-35S-GTW-YFP (Cesari *et al.*, 2016) vectors were used. Constructs generated for transient wheat protoplast transfection were derived from pAHC17-pUbi (Yoshida *et al.*, 2009; Okuyama *et al.*, 2011).

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

For stable wheat transformation the pVecNeo binary vector carrying pSr50-Sr50-YFPv was described by Cesari *et al.* (2016). Additional vectors containing nuclear localisation or export signals (pSr50-Sr50-YFPv-NLS, pSr50-Sr50-YFPv-nls, pSr50-Sr50-YFPv-NES and pSr50-Sr50-YFPv-nes) were generated following a similar procedure as for pSr50-Sr50-YFPv (Cesari *et al.*, 2016) except that the VenusR primer was replaced with either NESR, nesR, NLSR or nlsR (please refer to Table S2) to amplify of the YFP-NLS/nls/NES/nes fragments from the equivalent Sr33 constructs containing these fusions (Cesari *et al.*, 2016).

Heterologous expression of three AvrSr50 variants in *Escherichia coli* and protein purification

AvrSr50 proteins were expressed in T7 Express *lysY* competent *E. coli* (NEB, C3010I). Bacterial cultures were grown in Terrific Broth (TB) medium at 37°C with shaking at 220 RPM until OD_{600} was 0.6–0.8, and then induced with 500 µM isopropyl-1-thio-β-D-galactopyranoside (IPTG) and incubated for 16 h at 20°C. Cells were harvested by centrifugation and resuspended in a lysis buffer containing 50 mM HEPES pH 8.0, 300 mM NaCl, 1 mM dithiothreitol (DTT), 1 mM phenylmethanesulfonyl fluoride (PMSF) with 1 µg g^{-1} of cell pellet DNase I (Sigma). Cells were lysed using an Avestin Emulsifex C5 between 5000–10000 psi, followed by clarification of the lysate by centrifugation at 20 000 ***g*** for 35 min at 4°C. The clarified lysate was applied to a 5 ml HisTrap FF crude column (GE Healthcare, Chicago, IL, USA) and the column washed with 20 column volumes (c/v) of buffer consisting of 50 mM HEPES pH 8.0, 300 mM NaCl, 30 mM imidazole, and 1 mM DTT. The proteins were eluted using a linear gradient from 30 mM to 250 mM imidazole over 10 c/v. The fractions were visualised using Coomassie-stained SDS-PAGE, and the fractions containing the protein of interest were pooled. The N-terminal His_6-tag was removed by overnight treatment with His_6-tagged Human Rhinovirus 3C protease at 4°C in 50 mM HEPES pH 8.0, 300 mM NaCl, and 1 mM DTT. The 3C-cleaved proteins of interest were purified from the 3C cleavage mixture by passing the sample over a 5-ml HisTrap FF crude column pre-equilibrated with 50 mM HEPES pH 8.0, 300 mM NaCl and 30 mM imidazole, with 1 mM DTT. The column flowthrough was concentrated to 1 ml and purified further by size-exclusion chromatography on a Superdex 75 16/600 column (GE Healthcare) pre-equilibrated with 10 mM HEPES pH 8.0, 150 mM NaCl and 1 mM DTT. The peak fractions were analysed by Coomassie-stained SDS-PAGE and fractions containing the protein of interest were concentrated to *c.* 30 mg ml^{-1} in a 10-kDa MWCO Amicon centrifugal concentrator (Merck, Darmstadt, Germany), before flash freezing in liquid nitrogen and storage at −80°C. Protein concentration was determined from absorbance at 280 nm using an extinction coefficient calculated by the Expasy ProtParam tool; https://web.expasy.org/protparam/.

Crystallisation, data collection and structure determination

Initial screening of crystallisation conditions for AvrSr50 variants was performed in MRC2 96-well plates at 20°C using the sitting-drop vapour-diffusion method and commercially available sparse matrix screens. For screening, 400 nl drops, which consisted of 200 nl protein solution equilibrated against 200 nl of reservoir solution, were prepared using a NT8®-Drop Setter robot (Formulatrix, Bedford, MA, USA). The drops were monitored and imaged using the Rock Imager system (Formulatrix). No crystals were observed for either AvrSr50-A1 and AvrSr50-C, however, crystals of avrSr50-B6 (33 mg ml^{-1}) were observed in several different conditions. Crystals with the best morphology were observed in 3.5 M sodium formate pH 7.5 (Index C1) and 60% Tacsimate pH 7.0 (Index C5). Crystal optimisation was carried out in 24-well hanging-drop vapour-diffusion plate format, with 1 µl protein solution equilibrated against 1 µl of reservoir solution.

Before X-ray data collection, crystals were transferred into a cryoprotectant solution containing the reservoir solution with 20% glycerol. Some crystals were soaked for *c.* 60 s in reservoir solution supplemented with 1 M sodium bromide. Single-wavelength anomalous diffraction (SAD) and native datasets were collected on the MX1 (Cowieson *et al.*, 2015) and MX2 (Aragão *et al.*, 2018) beamlines at the Australian Synchrotron, respectively (Table S3). The datasets were processed in XDS (Kabsch, 2010) and scaled using AIMLESS (Evans & Murshudov, 2013) in the CCP4 suite (Winn *et al.*, 2011). For SAD phasing, the CRANK2 pipeline was used (Skubák & Pannu, 2013) in the CCP4 suite. The model was then refined using *phenix.refine* in the PHENIX package (Afonine *et al.*, 2012), and iterative model building and refinement was carried out in Coot (Emsley *et al.*, 2010). This model was used as a template for MR (McCoy *et al.*, 2007). Automatic model building was carried out with AUTOBUILD (Terwilliger *et al.*, 2007) and completed using *phenix.refine* (Afonine *et al.*, 2012) and COOT (Davis *et al.*, 2004). Structure validation was carried out using the MolProbity online server (Davis *et al.*, 2004). The coordinates and structure factors were deposited in the Protein Data Bank (Berman *et al.*, 2002) with the accession no. 7MQQ. Structures were visualised using PYMOL, and structural comparisons were carried out using the PDBeFold (Krissinel & Henrick, 2004) or the DALI server (Holm, 2020).

Transient expression and cell death induction assays in *Nicotiana benthamiana* and wheat protoplasts

Nicotiana benthamiana plants were grown in a growth chamber at 23°C with 16 h : 8 h, light : dark photoperiod. For agroinfiltration in *N. benthamiana*, pBIN19, pAM-PAT and pAHC17-pUbi vectors containing either AvrSr50 variants, Sr50 or CC-Sr50 were transformed into *Agrobacterium* strain GV3101_pMP90 (pBIN19 vectors) and GV3103 (pAM-PAT vectors) by electroporation. Single colonies were selected and grown in Luria-Bertani liquid medium containing 25 µg ml^{-1} of rifampicin, 15 µg ml^{-1} gentamicin and either 50 µg ml^{-1} of kanamycin (pBIN19 vectors) or 25 µg ml^{-1} of carbenicillin (pAM-PAT vectors) at 28°C for 24 h before infiltration. Bacteria were harvested by centrifugation and resuspended in infiltration medium (10 mM MES pH 5.6, 10 mM $MgCl_2$ and 150 µM acetosyringone). Bacteria culture densities were adjusted to different

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

OD_{600} values (35S-Sr50 = 0.1, pUbi-Sr50 = 2.0, CC-Sr50 = 0.5 and YFP-AvrSr50 = 1.0, AvrSr50 = 0.5) then mixed and co-infiltrated in 4-wk-old *N. benthamiana* plants. To evaluate AvrSr50 mutants for recognition, each mutant was infiltrated on at least 10 spots, two per leaf with a maximum two leaves per plant. Cell death induction was analysed from 48 to 96 h after infiltration.

Wheat (*Triticum aestivum* L.) seedlings of Fielder variety were grown in a growth chamber under shade at 25°C with a photoperiod of 16 h : 8 h, light : dark for 7–9 d. Isolation of wheat protoplasts was performed as described in Arndell *et al.* (2019). Briefly, the epidermis of the wheat leaves was peeled off and the leaf immersed in 0.6 M mannitol for 15 min to induce plasmolysis. Leaves were then transferred to enzyme solution containing (20 mM MES-KOH (pH 5.7), 0.8 M mannitol, 10 mM KCl, 1.5% (w/v) cellulase R10, 0.75% (w/v) macerozyme R10, 10 mM $CaCl_2$, 10 mM KCl and 0.1% (w/v) BSA) and incubated in the dark for 3 h with gentle rotation at *c.* 40 rpm. Then the solution was filtered through a 100-µM cell strainer. Protoplasts were collected by centrifugation at 80 ***g*** for 3 min and washed twice with washing solution W5 (2 mM MES-KOH (pH 5.7), 5 mM KCl, 125 mM $CaCl_2$, 154 mM NaCl), followed by gravity collection over 40 min on ice and resuspended in MMG solution (4 mM MES-KOH (pH 5.7), 0.4 M mannitol, 15 mM $MgCl_2$). Protoplasts were counted on a haemocytometer and adjusted to a concentration of *c.* 3.0×10^5 cells ml^{-1} using MMG solution. The luciferase reporter plasmid (12 µg) and the respective effector and resistance gene constructs (15 µg) were co-transfected into 200 µl of wheat protoplasts using PEG transfection solution (40% (w/v) PEG 4000, 0.2 M mannitol, 100 mM $CaCl_2$). After 16 h protoplasts were collected by centrifugation and dissolved in Luciferase Cell Culture Lysis buffer (Promega, catalogue no. E1531). Luciferase activity was measured using the Luciferase Assay System (Promega, catalogue no. E1501).

Interaction assays in yeast-two-hybrid system

AD-AvrSr50 variants were co-transformed with the full-length BD-Sr50 vector in the *Saccharomyces cerevisiae* HF7c strain and grown on synthetic dropout (SD) medium without tryptophan (W) and leucine (L). For interaction assays, three independent colonies for each combination were grown in 5 ml of liquid SD medium (−WL) for 16 h at 30°C. Yeast cultures were adjusted to OD_{600} = 1.0 and then diluted 1/10, 1/100 and 1/1000 times in sterile conditions. Aliquots (5 µl) of each sample were inoculated onto selective medium lacking tryptophan, leucine and histidine (−WLH) for interaction analysis and incubated at 30°C. Yeast growth was evaluated from 72 to 96 h after incubation.

Wheat transformation and stem rust resistance analyses

Stable transgenic wheat plants were generated by the transformation of pVecNeo binary vectors (Table S1) into the stem rust susceptible cultivar Fielder using *Agrobacterium tumefaciens-mediated* transformation system as previously described (Richardson *et al.*, 2014). For stem rust resistance analyses, 12–20 independent lines were generated for each construct and five T0 lines (Table S4) were infected with *Pgt* isolate 98–1,2,3,5,6. Transgene presence was confirmed by PCR using Sr50-specific primers (Sr50-F3 and Sr50-R3; Table S2).

Confocal microscopy

Wheat epidermal cells were observed on 10–14-d-old T1 plants and samples were prepared in perfluorodecalin. Wheat cell images were acquired using the Online Fingerprinting mode of Zen 2012 software on a Zeiss 780 confocal microscope using a LD ×40 1.1W lens. Reference spectra for Online Fingerprinting were derived from 14-channel (517–641 nm) Lambda mode data. The YFP spectra derived from samples were compared with manufacturer's data to confirm their identity. The validity of these spectra was checked periodically on experimental samples by Spectral Unmixing of Lambda mode data, thereby confirming that the residuals channel, representing fluorescence emission from uncharacterised sources, was consistently of low intensity.

Allele mining and phylogenetic analyses

AvrSr50 allelic sequences from six rust isolates (Pgt21-0, Ug99, QCMJC, RKQQC, 4a, MCCFC, SCCL) were reported previously (Chen *et al.*, 2017). Allelic sequences from a further 40 global isolates were assembled from DNA or RNA sequence data (Illumina paired-end reads, European Nucleotide Archive, PRJEB22223). Reads from each isolate (Table S5) were initially mapped at low stringency (similarity fraction 0.88, length fraction 0.7) to the reference AvrSr50 allele from Pgt21-0 using the 'Map Reads to Reference' tool in CLC Genomics Workbench 10 or later (CLCGWB; Qiagen Aarhus) and the mapped reads were extracted. This approach captured reads representing both haplotypes, which were then *de novo* assembled (CLCGWB tool) at high stringency (similarity fraction 0.98, length fraction 0.95, Global alignment) to build the two allele models for each isolate. In some cases in which the second allele did not assemble fully, unmapped reads from the high stringency assembly were reassembled with lower stringency. Accuracy of haplotype sequences were cross checked by visual inspection of read mapping tracks in the CLCGWB browser. Translated amino-acid sequences were aligned using Clustal and a maximum likelihood tree constructed using Mega v.X (Kumar *et al.*, 2018).

Results

Sr50 recognises AvrSr50 and induces stem rust resistance in the cytosol of wheat cells

Nucleotide-binding leucine-rich repeat receptors have been shown to confer resistance and recognise effector proteins in different subcellular compartments. To investigate Sr50 distribution in wheat cells, we generated stable wheat transgenic lines expressing Sr50 fused to Yellow Fluorescence Protein-venus (YFPv) under the control of its native promoter (pSr50) (Fig. 1a). T0 transgenic lines showed a typical resistance response to a *P.*

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

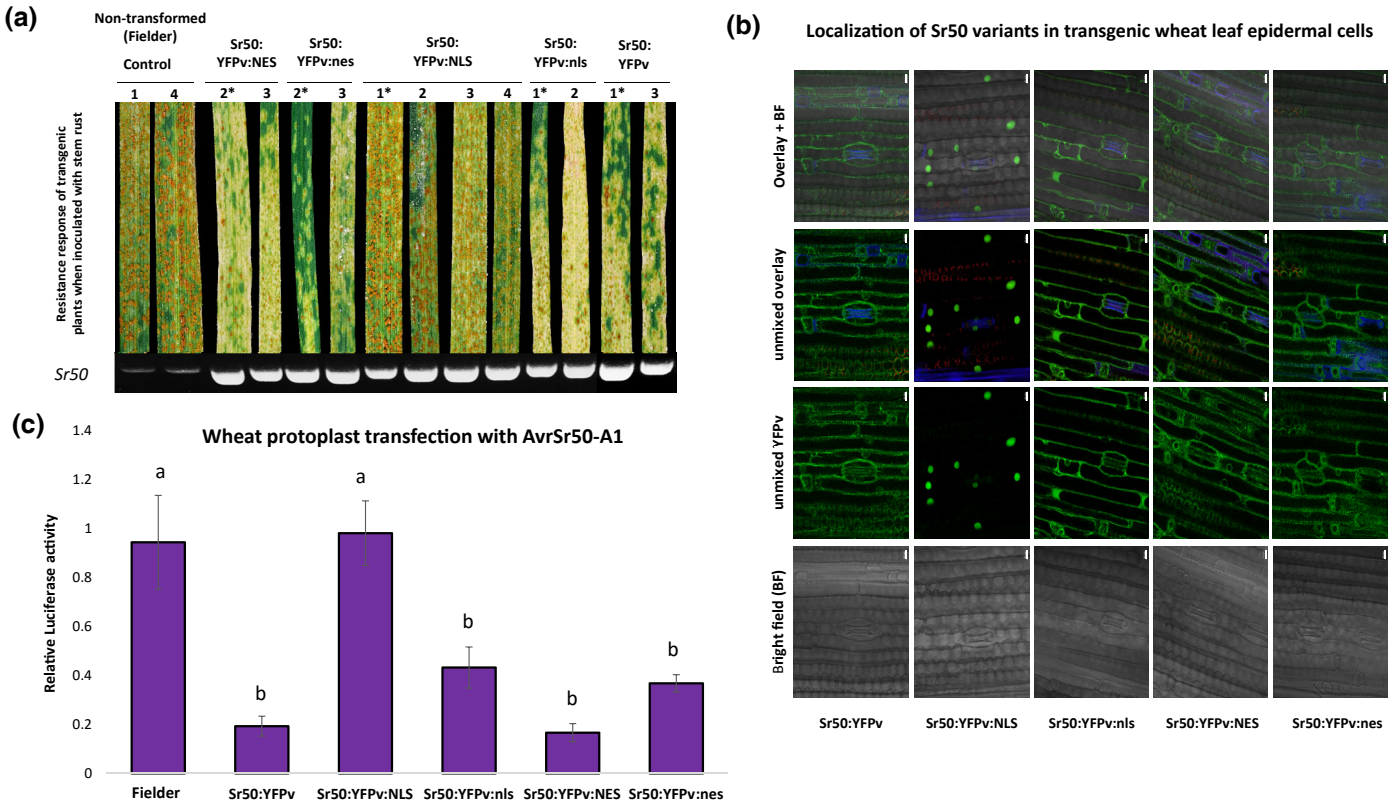


Fig. 1 Sr50 cytosolic pool confers disease resistance to stem rust. (a) Stem rust resistance phenotype of nontransformed wheat (cv Fielder) control and at least two independent T0 lines (numbered) transformed with Sr50 fused to either Yellow Fluorescence Protein-venus (YFPv):NES, YFPv:nes, YFPv:NLS, YFPv:nls or YFPv. The presence of the transgene was determined by PCR using Sr50-specific primers (lower panel). Transgenic lines used for confocal microscopy and protoplast assays are indicated by an asterisk. (b) Confocal images showing localisation of Sr50 fused to YFPv, YFPv:NLS, YFPv:nls, YFPv:NES, or YFPv:nes in transgenic wheat leaf epidermal cells. YFPv fluorescence appears in green, stomatal and chloroplastic autofluorescence emissions are shown in blue and red, respectively (Bar, 20 μm). (c) Relative luciferase activity in wheat protoplasts from transgenic lines shown in (a) were independently co-transformed with *AvrSr50* and *luciferase* constructs. Values are relative to luciferase activity of Fielder cells co-transformed with *GFP* and *luciferase* constructs. Results represent the means of three biological replicates with error bars indicating the SE. Differences between samples were assessed by analysis of variance (ANOVA). Samples marked by identical letters in the plots do not differ significantly ($P < 0.05$) in the Tukey test.

graminis isolate carrying AvrSr50, with only small pustules developing surrounded by chlorosis and necrosis (Fig. 1a; Table S5). By contrast, the nontransformed Fielder control was highly susceptible, with extensive large pustule development and no chlorosis/necrosis. Confocal microscopy showed that Sr50:YFPv was mainly present in the cytosol, as previously observed for Sr33 (Cesari *et al.*, 2016) (Fig. 1b). To determine whether Sr50 function is confined to the cytosol we generated additional transgenic lines expressing Sr50:YFPv fused to either a Nuclear Localisation Signal (NLS), a Nuclear Export signal (NES) or nonfunctional variant of these signal motifs (nls and nes respectively). T0 transgenic plants expressing Sr50:YFPv:NES, Sr50:YFPv:nes and Sr50:YFPv:nls showed resistance to *Pgt* similar to those expressing wild-type Sr50:YFPv (Fig. 1a; Table S5). However, lines expressing the nuclear-targeted Sr50:YFPv:NLS version showed a similar highly susceptible phenotype to the nontransformed Fielder control (Fig. 1a). All lines were fully susceptible to an isolate of wheat leaf rust (*Puccinia triticina*), indicating that there was no nonspecific resistance (Fig. S1a). Confocal microscopy confirmed that Sr50:YFPv:NES, Sr50:YFPv:nls and Sr50:YFPv:nes proteins were present in the cytosol but not the nucleus and were indistinguishable from each other and the untagged Sr50:YFPv, while Sr50:YFPv:NLS exclusively accumulated in the nuclei (Fig. 1b). All the fusion proteins were detected by immunoblotting of proteins extracted from the transgenic lines (Fig. S1b).

AvrSr50 was previously reported to localise in both the cytosol and nucleus of *N. benthamiana* cells after transient expression (Chen *et al.*, 2017). To further investigate whether AvrSr50 recognition by Sr50 is indeed restricted to the cytosol in wheat, we isolated protoplasts from different Sr50-tagged transgenic lines and co-transformed them with constructs expressing AvrSr50 and the luciferase protein. We then evaluated cell death induction by quantifying the reduction of luciferase activity as an indicator of cell death and therefore AvrSr50 recognition. Substantially reduced luciferase accumulation was observed in the protoplasts from transgenic lines expressing Sr50 in the cytosol compared with protoplasts expressing Sr50 in the nucleus, which were similar to the nontransgenic control line (Fig. 1c). Similar levels of luciferase activity were detected between transgenic lines when co-expressed with GFP as a control (Fig. S1c). Together, these results confirm that a cytosolic Sr50 pool is required to

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

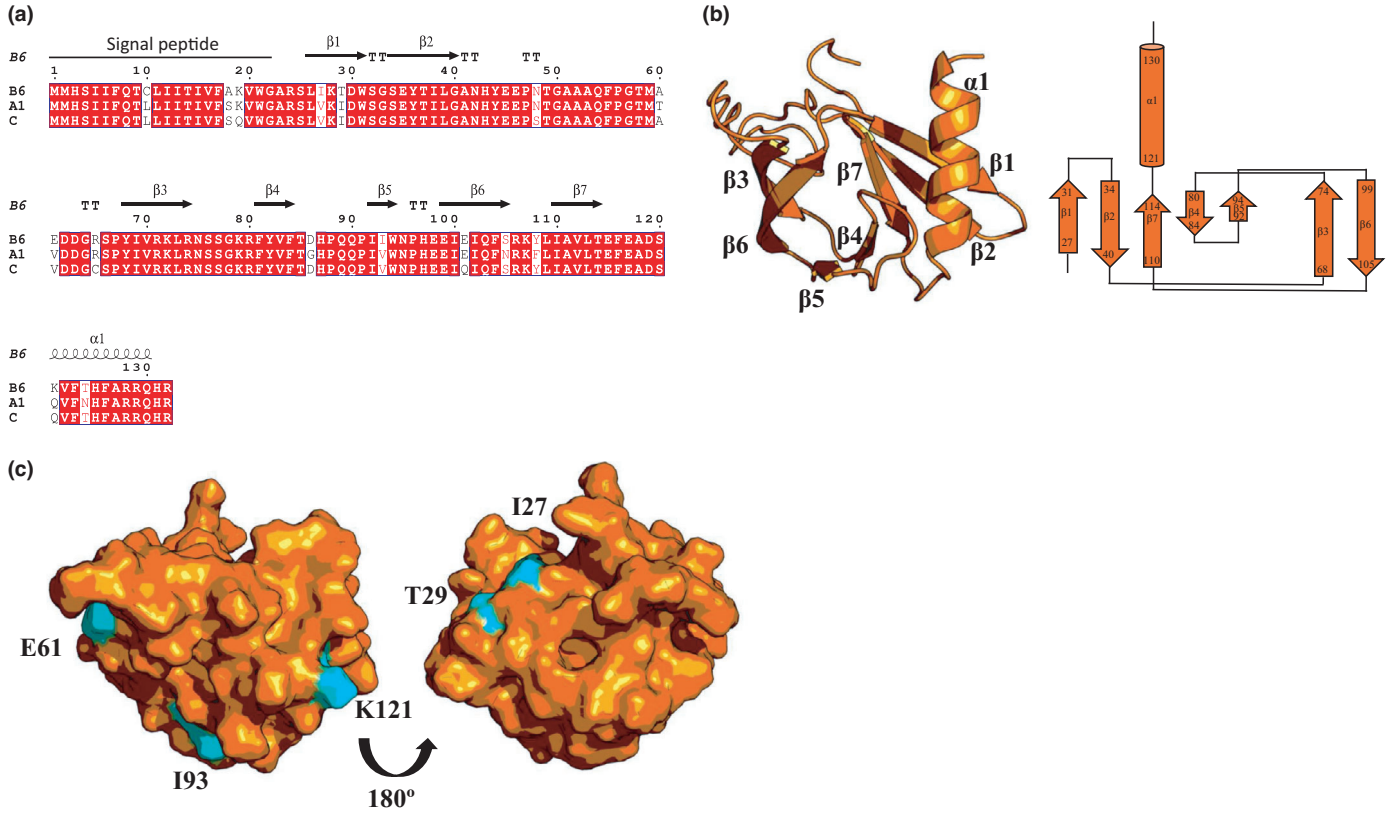


Fig. 2 The crystal structure of avrSr50-B6. (a) Multiple sequence alignment of the three AvrSr50 allelic variants, B6, A1, and C generated using ESPript (Robert & Gouet, 2014). The secondary structure elements of avrSr50-B6 are displayed on the alignment. 'T' corresponds to a β-turn. (b) Cartoon representation showing the overall structure of avrSr50-B6, which consists of seven β-strands arranged in two antiparallel β-sheets, in which the first β-sheet comprises strands β1, β2, β4, β5 and β7, and the second β-sheet consists of β3 and β6, and a C-terminal α-helix comprised of residues 121–132 (α1). Topology map of avrSr50-B6 in which the arrows represent β-strands and cylinder represents the α-helix. (c) Surface representation of avrSr50-B6 highlighting polymorphic residues (in cyan) unique to avrSr50-B6 when compared with AvrSr50-A1 and AvrSr50-C. Residue numbers are as indicated.

recognise AvrSr50 and confer resistance to *P. graminis* infection in wheat.

The crystal structure of AvrSr50 reveals a single surface-exposed residue that controls recognition by Sr50

We previously identified two divergent alleles of *AvrSr50* from isolates RKQQC and QCMJC (Chen *et al.*, 2017) that encode proteins with either 9 or 12 amino-acid differences from the originally identified avirulence allele (Fig. 2a). We designated these alleles as avrSr50-B6 (from QCMJC), AvrSr50-C (from RKQCC) and AvrSr50-A1 (wild-type form), based on a phylogenetic analysis of AvrSr50 variants performed in this study (please refer to the following sections). We previously found that AvrSr50-A1 and AvrSr50-C induced cell death when co-transformed with Sr50 in *N. benthamiana* plants, while avrSr50-B6 did not, suggesting that the unique polymorphic sites in avrSr50-B6 led to a loss of recognition (Chen *et al.*, 2017).

To advance our understanding of AvrSr50 recognition by Sr50, we expressed the three allelic variants in *E. coli* (residues 23–132) and purified the proteins to homogeneity using nickel affinity chromatography and size-exclusion chromatography. All three proteins were used for crystallisation trials, but we were only able to obtain crystals for avrSr50-B6. The structure of avrSr50-B6 was determined using a bromide ion-based SAD approach (Table S3), to a final resolution of 1.15 Å. The final model contains a single molecule in the asymmetric unit consisting of residues 23–132 of avrSr50-B6 with three nonnative residues (Gly–Pro–Met) at the N-terminus remaining following cleavage of the His_6-tag. The overall structure of avrSr50-B6 has seven β-strands arranged in a β-sandwich fold composed of two antiparallel β-sheets, with a C-terminal α-helix comprised of residues 121–132. The first β-sheet was comprised of strands β1, β2, β4, β5 and β7, and the second β-sheet consisted of β3 and β6 (Fig. 2b). The loop region between β2–β3 (residues Thr58 to Arg65) appeared to exist in multiple conformations, and while this region was included in the final model, we have reduced confidence concerning the absolute position of these residues compared with the rest of the structure (further details in Fig. S2). The five polymorphic residues unique to the nonrecognised avrSr50-B6 variant all map to the surface of the protein (Fig. 2c) and could therefore contribute to recognition specificity by affecting the interaction with the Sr50 receptor.

To evaluate the role of these residues in recognition we introduced the corresponding mutations in each of these five positions

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

individually to AvrSr50-A1 (V27I, I29T, V61E, V93I, Q121K), including one double mutation (V27I/I29T). These mutants were tested in transient expression assays in *N. benthamiana* leaves and wheat protoplasts. In wheat protoplasts, constructs expressing AvrSr50 variants were co-transformed with Sr50 and luciferase (all driven by the maize pUbq promoter), and cell death was quantified by relative luciferase expression. Co-expression of Sr50 with the wild-type AvrSr50-A1 version showed a strong reduction in luciferase activity compared with co-expression with the virulence allele avrSr50-B6, or expression of Sr50 or AvrSr50 variants alone (Fig. S3a). All the AvrSr50-A1 mutants induced a strong cell death response similar to wild-type AvrSr50-A1, except for the Q121K substitution, which showed elevated luciferase expression, although slightly less than the native B6 variant (Fig. 3a). In *N. benthamiana*, AvrSr50 variants were transiently co-expressed with Sr50 driven by either the

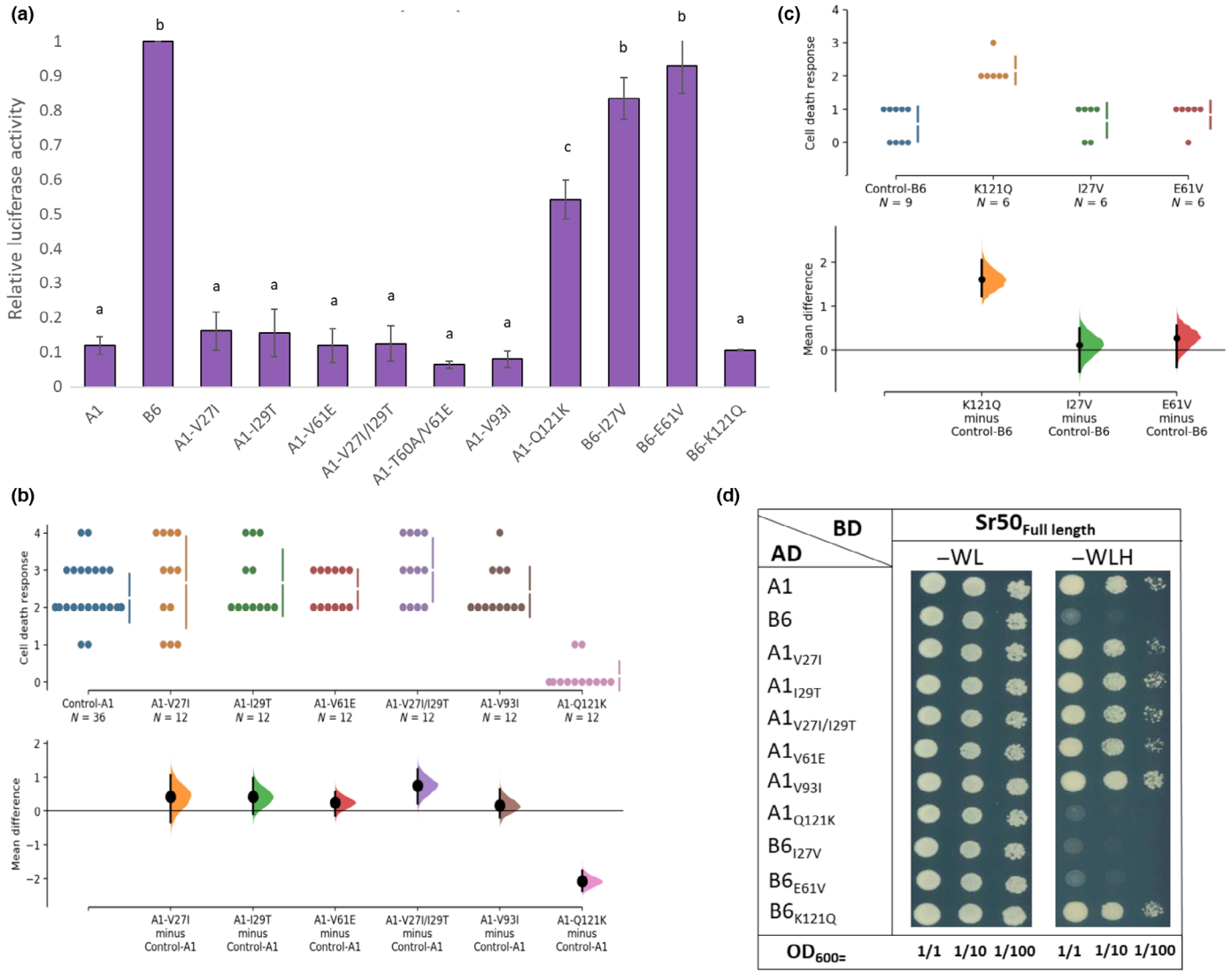
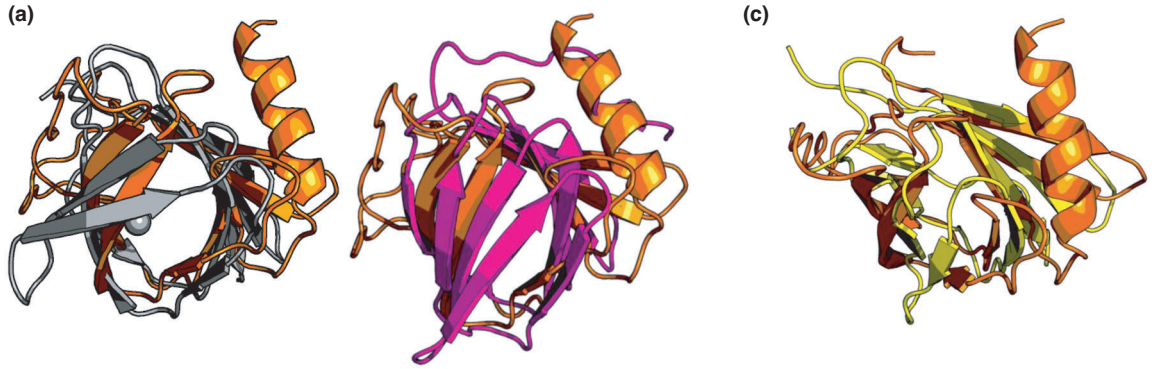


Fig. 3 Recognition of AvrSr50 mutants by Sr50. (a) Cell death induction was estimated in wheat protoplasts by measuring relative luciferase activity when *luciferase* is co-transformed with *pUbq-Sr50* and either *pUbq-AvrSr50-A1* or *pUbq-avrSr50-B6* variants carrying the indicated amino-acid substitutions. Values are relative to samples co-transformed with *avrSr50-B6* and luciferase constructs, which was set to 1. The plots represent the results for three biological replicates with error bars indicating SE. Differences between samples were assessed by analysis of variance (ANOVA). Samples marked by identical letters in the plots do not differ significantly ($P < 0.05$) in the Tukey test. (b, c) Cummings estimation plots showing cell death phenotypes observed upon co-expression of the AvrSr50-A1 (b) or avrSr50-B6 (c) variants with Sr50 (*pUbq* promoter) in *N. benthamiana* leaves scored on a scale of 0 (no symptoms) to 4 (strong cell death) (Supporting Information raw data Figs S7–S10). The upper panels show the distribution of the observed scores for each variant in which each dot represents the cell death score of an independent infiltration assay and the number of replicates per sample are indicated in the *x*-axis. The lower panels represent the mean differences for comparisons of variants against a shared control (A1 or B6). Each mean difference is depicted as a dot with 95% confidence intervals indicated by the ends of the vertical error bars and bootstrap sampling distributions indicated by the horizontal graph. (d) Yeast-two-hybrid protein interaction assay in yeast colonies co-transformed with *AvrSr50-A1* and *avrSr50-B6* variants fused to the *GAL4* activation domain (AD) and *Sr50* fused to GAL4 binding domain (BD). Yeast colonies with a OD = 600 and two subsequent dilutions of 1/10 and 1/100 were spotted on synthetic double dropout (DDO) medium (−WL, growth control) and on synthetic triple dropout TDO (−WLH, interaction selection) selective medium. Figures shown are representatives of three independent experiments and pictures were taken after 5 d of growth.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation



(b)

```
avrSr50_B6     GPMARSL I KTDWSGSEYT I LGANHYEEPNTGAAAQFPGTMAEDDGRSPY I VRKLRNSSGKRFYV
2FQP_BP2299    GMKRPGA I PTVQ I DNER--VKVTEWRFPPGGETGWHRHSMD-------YVVVPMTTGP-----L
                                                          *  *
avrSr50_B6     FTDHPQQP I I WNPHEE I E I QFSRKYL I AVLTEFEADSK-VFTHFARRQHRS
2FQP_BP2299    LLETPEGSV--TSQLTRGVSYTRPEGVEHNV I NPSDTEFVFVE I E I KAA
                                                 *                 *
```

Fig. 4 Comparative structural analysis shows AvrSr50 shares structural homology with sugar-binding proteins. (a) Structure superimposition of avrSr50-B6 (shown in orange; cartoon representation) and the top cupin superfamily member hit from the PDBeFold server: cupin domain (bp 2299) from *Bordetella pertussis* Tohama I (grey; PDB ID: 2FQP) and DALI server: HP0902: an uncharacterised protein from *Helicobacter pylori* 26 695 (shown in magenta; PDB ID: 5J4F). For the cupin domain (bp2299) the zinc metal ion bound in the structure is shown in sphere form. (b) Alignment of avrSr50-B6 and the cupin protein from *Bordetella pertussis* Tohama I (PDB ID: 2FQP). The residues involved in zinc co-ordination are indicated by a black asterisk (c) Structure superimposition of avrSr50-B6 (orange; cartoon representation) and the internal xylan binding domain (carbohydrate-binding module alone) from *Cellulomonas fimi* xylanase D (yellow; PDB ID: 1xbd).

pUbq or p35S promoters. The cell death response induced by each combination was scored on a 0–4 scale and analysed by estimation graphs (Ho *et al.*, 2019). Similar to the protoplast assays, co-expression of the AvrSr50-A1 mutants with Sr50 consistently induced cell death responses similar to the wild-type protein, with the exception of the Q121K variant that showed no recognition (Figs 3b, S4a, S5–S10). As observed previously (Chen *et al.*, 2017), expression of Sr50 from the strong p35S promoter resulted in some background cell death due to Sr50 autoactivity, but nevertheless allowed clear detection of the AvrSr50-induced response (Figs S4a, S5–S7). Conversely, the weaker activity of the pUbq promoter in *N. benthamiana* largely eliminated the background autoactivity of Sr50, but still allowed detection of the AvrSr50-induced responses (Figs 3b, S8–S10). Both constructs gave qualitatively similar results. Altogether, these results indicated that the Q121K substitution was sufficient to prevent recognition of AvrSr50 by the Sr50 receptor.

We also generated the reciprocal substitution (K121Q) in the virulence allele variant avrSr50-B6 to test if this was sufficient to restore recognition. Indeed, this single amino-acid change restored recognition of avrSr50-B6 in both *N. benthamiana* and wheat protoplast co-expression assays (Figs 3a,c, S4b, S11, S12). By contrast, the reciprocal I27V and E61V mutations in avrSr50-B6 did not restore recognition. Immunoblotting showed that all of the AvrSr50 protein variants were expressed in *N. benthamiana* (Fig. S11b). Previously, we have shown that AvrSr50-A1 binds to Sr50 in yeast, while the avrSr50-B6 allelic variant failed to interact (Chen *et al.*, 2017). Here, we tested all the AvrSr50 mutants that we generated for binding to Sr50 in yeast-two-hybrid assays and observed a complete correlation between interaction in yeast and recognition *in planta* (Fig. 4b). That is, only the Q121K mutation disrupted this interaction for the wild-type (A1) variant, while the reciprocal K121Q mutation restored the interaction for the B6 variant. All proteins were well expressed in yeast (Fig. S3b). Taken together, these results supported the idea that Sr50 recognition of AvrSr50 occurs through a direct interaction and that the Q121K polymorphism in avrSr50-B6 is the only polymorphic residue that impacts on AvrSr50 recognition.

AvrSr50 shares structural homology to members of the cupin superfamily and carbohydrate-binding module moiety

The molecular function of AvrSr50 is unknown. Utilising the crystal structure of avrSr50-B6 we carried out a comparative structural analysis using the PDBeFold server (Krissinel & Henrick, 2004) and DALI server (Holm, 2020). The top hits from both servers show structural similarity to cupin superfamily members and carbohydrate hydrolases, in particular glucosidases and xylosidases (Tables S6, S7), despite sharing limited sequence identity (< 10%). In both cases, the structural similarity is limited to the β-sandwich fold and does not include the C-terminal helix. The cupin superfamily is one of the most functionally diverse families known, and includes catalytically inactive seed storage proteins, sugar-binding metal-independent epimerases and metal-dependent enzymes. Most of the known enzymatic cupin proteins, including the hit from PDBeFold (the cupin domain from *Bordetella pertussis* Tohama I (Fig. 4a; Table S6), contain a

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

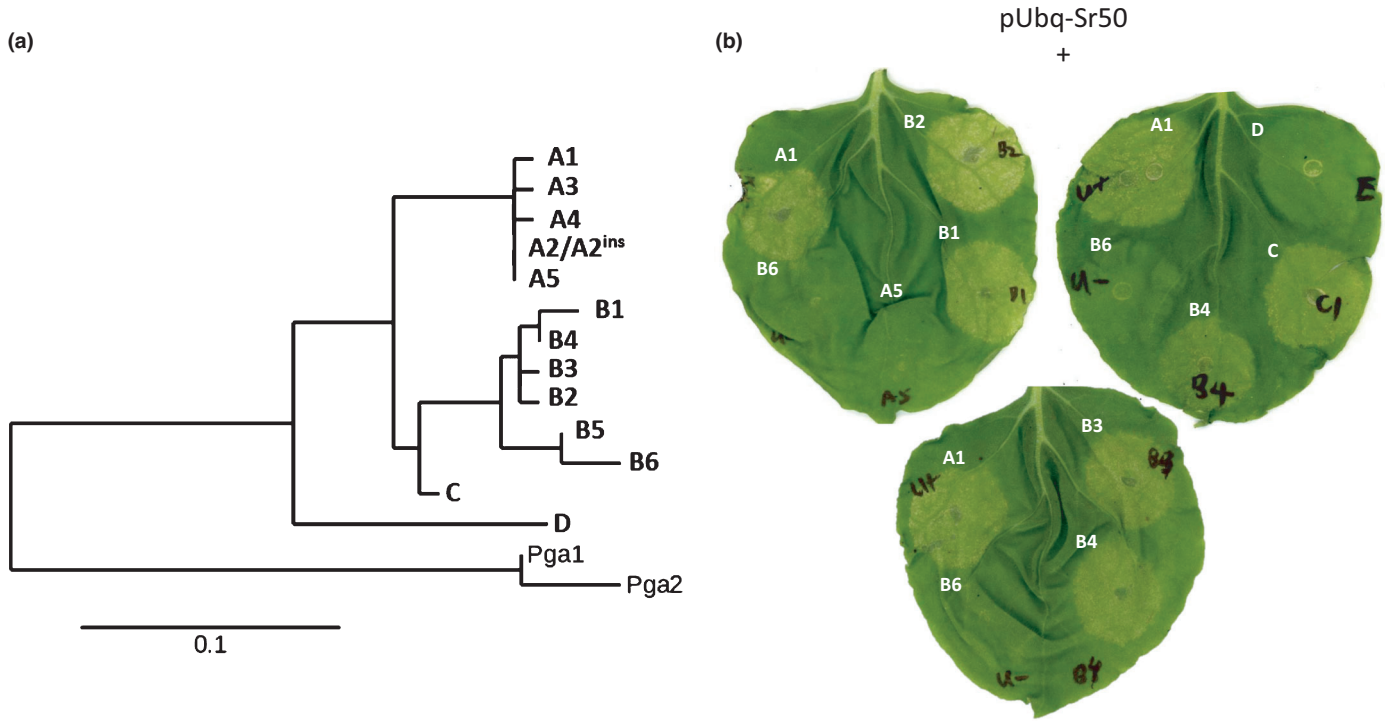


Fig. 5 Recognition of AvrSr50 natural variants from different stem rust isolates. (a) Phylogenetic tree generated from AvrSr50 variants found in 46 isolates of *Puccinia graminis* f.sp. *tritici*. AvrSr50 variants are grouped in five different clades identified as A, B, C, D and *Pga*. (b) Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *pUbq-Sr50* and *AvrSr50* variants of phylogenetic groups A, B, C and D.

transition state metal, coordinated by three histidine residues and an acidic amino acid. However, structural and sequence investigation shows that these coordinating residues are not present in AvrSr50 (Fig. 4b), suggesting it is not likely to be enzymatic. The structure of avrSr50-B6 also shares similarity to HP0902 from *Helicobacter pylori* (the top hit from the DALI server; Table S7), a nonenzymatic sugar-binding cupin, which shares structural homology to Rml-C-like cupins (Sim *et al.*, 2016) (Fig. 4a). While the substrate-binding location of HP0902 and Rml-C is conserved the residues that co-ordinate sugar binding are not, making it difficult to predict if and which sugars AvrSr50 may bind. Structural similarity between avrSr50-B6 and the carbohydrate hydrolases (Tables S6, S7) is limited to the CBM moiety, which are discrete modules in these larger enzymes that typically aid in maintaining proximity of the carbohydrate to the catalytic domain but have no catalytic activity of their own. Further investigations into these hits do not show conservation in the known residues involved in sugar co-ordination in AvrSr50, and further experimentation is required to determine if AvrSr50 is a sugar-binding protein.

Natural AvrSr50 alleles are highly polymorphic

We had previously characterised the genotype of seven *Pgt* isolates from Australia and the USA as well as Ug99 from Africa at the *AvrSr50* locus, and identified six allelic variants (Chen *et al.*, 2017). Here, we analysed publicly available genomic DNA sequence data generated from a larger set of 43 stem rust isolates from around the world (Lewis *et al.*, 2018). After mining the sequence data, we were able to reconstruct the genotypes of 38 isolates, while the other five samples did not contain sufficient sequence depth to reassemble this genomic region in full (Table S4). Most isolates (31) displayed heterozygosity for two different sequence variants, while a single allele was identified in seven isolates consistent with either homozygosity, the presence of a deletion in one haplotype, or insufficient sequence data to resolve the second allele. In total, 14 different *AvrSr50* alleles were identified among all the *Pgt* isolates, with a further two alleles identified in two isolates of *P. graminis* f. sp. *avenae* (*Pga*), which infects oat (Figs 5a, S13, S14). A phylogenetic analysis of these variants revealed two main clades in *Pgt*, designated A and B, with each containing several closely related alleles (Fig. 5a). The A group includes the originally identified *AvrSr50* avirulence allele from Pgt21-0, designated A1, as well as four additional alleles. One of these, A2, has two forms that are identical in sequence except that one version contains a 26-kbp insertion ($A2^{ins}$) within the gene and corresponds to the virulence allele detected in Pgt21-0 and Ug99 isolates (Chen *et al.*, 2017). The A5 variant is identical in sequence to A2, except that it contains a single nucleotide change in the stop codon that results in translation of a larger protein with a C-terminal extension of 149 aa. This variant occurs in the UK-01 isolate, which is a representative of the Digalu race group (pathotype TKTTF, Clade IV-B, Lewis *et al.*, 2018). The B group includes the virulence allele *avrSr50-B6* from isolate QCMJC, as well as five other closely related variants. Two additional divergent variants outside these clades, C (from RKQQC) and D (from Iranian isolate IR-01) were also observed. To determine the virulence phenotypes of these additional

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

AvrSr50 variants, we generated expression constructs and tested these for recognition by Sr50 in *N. benthamiana* plants. We observed that the AvrSr50-B1, -B2, -B3, -B4 and -B5 variants all induced cell death in *N. benthamiana* when co-expressed with pUbq-Sr50 (Figs 5b, S15, S16), and therefore are likely to represent avirulence alleles. None of these variants carried the Q121K substitution, which prevents recognition by the B6 virulence allele and appears to be a unique and recently acquired virulence mutation. The V93I and Q121K substitutions are the only polymorphisms unique to the nonrecognised B6 variant, consistent with our finding above that only the position 121 polymorphism affected recognition. Finally, we observed a lack of recognition and cell death induction of the alleles AvrSr50-A5 and AvrSr50-D. For AvrSr50-A5, this correlated with a lack of expression of the fusion protein (Fig. S16b). This suggests that the C-terminal extension due to the stop codon mutation in this allele results in destabilisation of the protein, therefore preventing recognition. The AvrSr50-D variant contains a Q121L substitution, which may explain its lack of recognition, although this variant also includes numerous other unique variants that could also affect recognition. Because these latter two variants occur in rust strains that also contain a recognised allele of AvrSr50, we cannot directly correlate the *in planta* recognition with the virulence phenotype they confer in the rust fungus.

Discussion

Functional molecular studies of proteins involved in plant–pathogen interactions are important for a better understanding of plant disease emergencies. In many cases, these studies have shown that the breakdown of plant resistance is explained by the evolution of pathogen avirulence effector proteins that escape recognition from major resistance genes. As pathogens exhibit shorter generational cycles compared with those of plants, evolution of effector proteins often occurs at higher rates than those of resistance genes (Karasov *et al.*, 2020). A deeper understanding of how pathogens evolve to gain virulence and the genetic diversity of virulence loci is important for surveying the emergence of potential virulent isolates on natural pathogen populations. To address this, we studied the genetic variability of avirulent/virulent *AvrSr50* alleles and identified unique polymorphic sites present in the virulent effector protein variant avrSr50-B6. Using structure-guided site-directed mutagenesis in combination with binding and transient expression assays, we then investigated the role of individual amino-acid substitutions found in avrSr50-B6 to escape Sr50 recognition and gain virulence. Here, we observed that one residue substitution (Q121K) in avrSr50-B6 is sufficient to explain the lack of binding and recognition of this resistance-escaping variant. Substitution of this residue in avrSr50-B6 was sufficient to fully restore recognition as detected in wheat protoplasts, *N. benthamiana* and yeast-two-hybrid assays. The reciprocal substitution in AvrSr50-A1 abolished recognition in these assays, although not completely in wheat protoplasts, which are likely to reflect slight differences in the quantitative outputs of these assays, or possibly a weak recognition of this variant in the Fielder wheat background. Conversely, we observed that the cumulative polymorphic sites present in the avirulent AvrSr50-C and other natural variants such as AvrSr50-B1, -B2, -B3 and -B4 are not sufficient to abolish recognition. Nevertheless, it is likely that other residues in the AvrSr50 protein contributed to the binding surface interacting with Sr50. This picture would be consistent with other studies in which multiple amino acids across the effector protein surface are involved in binding and recognition by the respective resistance protein (Rehmany, 2005; Wang *et al.*, 2007; Catanzariti *et al.*, 2010; Krasileva *et al.*, 2010). An interesting example is AVR-Pik, an effector from the blast fungus *Magnaporthe oryzae*, which is recognised by the rice NLR pair, Pik1 and Pik2. Pik1 contains an integrated heavy metal associated (HMA) domain that directly interacts with AVR-Pik. Structural and biochemical studies of the effector variants and Pik-HMA domains have shown that polymorphisms in effector proteins and the HMA domain binding interface underpins binding and ultimately recognition specificity. Nevertheless, they also showed that a single polymorphism in AVR-PikC (Ala67Asp) severely reduces Pik-HMA binding (Maqbool *et al.*, 2015) which nicely illustrates both the robustness of effector recognition and the crucial role of key residues in effector binding (Kanzaki *et al.*, 2012; Maqbool *et al.*, 2015; De la Concepcion *et al.*, 2018, 2021; Maidment *et al.*, 2021). Similar observations have been reported for the interaction of the RGA5-HMA integrated domain and the *M. oryzae* effector proteins AVR-Pia and AVR1-CO39 (Ortiz *et al.*, 2017; Guo *et al.*, 2018).

To date, at least three independent examples of matching R-Avr proteins in model fungal pathosystems; AvrL567/*M. lini* L, AVR-Pik-Pik and MLAa/AVR_A from/flax, *M. oryzae*/rice and *Bgh*/barley (Dodds *et al.*, 2006; Maqbool *et al.*, 2015; Saur *et al.*, 2019), respectively, have been characterised by a high genetic diversity of the interacting proteins. In these examples, gain of virulence mainly occurred through polymorphic changes in key surface-exposed amino acids, which led to escape from recognition. Similarly, here we observed that some virulent variants of the AvrSr50 effector protein have also arisen due to changes in surface-exposed residues involved in the interaction with the resistant protein Sr50. Nevertheless, the insertion of a large DNA fragment (26 kbp) resulting in gene knockout is at the origin of the most common virulence allele of *AvrSr50* (Chen *et al.*, 2017). Gain of virulence through insertion and deletions has also recently been observed in other *Pgt* effector proteins. For example, there are two main classes of AvrSr35 virulence alleles, one containing a 400-bp MITE insertion (Salcedo *et al.*, 2017) and the other with a 57-kbp insertion detected in Ug99 (Li *et al.*, 2019). AvrSr35 variants containing a few amino-acid variations from the reference avirulence allele have also been identified, but the effect of these substitutions on recognition specificity has not yet been determined. Similarly, several *Pgt* isolates in the race 21 clonal lineage have evolved virulence for *Sr27* in the field through the deletion of the *AvrSr27* locus, which contains two adjacent copies of this *Avr* gene (Upadhyaya *et al.*, 2021). The apparent prevalence of *Pgt* gain of virulence through deletions and insertions over the accumulation of amino-acid sequence changes may be explained by the fact that *Pgt* populations are mainly clonal due to its heteroecious lifestyle. Therefore, virulent isolates are

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

often the result of spontaneous mutations that occur within asexual lineages, rather than the segregation of pre-existing allelic diversity (Figueroa *et al.*, 2020).

Sr50 naturally occurs in rye (Mago *et al.*, 2015) and is orthologous to the *Mla* resistance gene in barley, which is characterised by extensive functional diversification and encodes multiple allelic resistance proteins with different specificities of recognition (Saur *et al.*, 2019). To date, *Sr50* wheat introgression lines have not been widely deployed in commercial crops, implying that selection for the *AvrSr50* virulence allele in the agricultural context has been limited. Nevertheless, we have observed four different alleles of *AvrSr50* that can allow escape from recognition. One of these results from an insertion that disrupts the gene ($A2^{ins}$), one from the loss of the stop codon (A5) and two from amino-acid polymorphisms (B6 and D). Two of these variants (A5 and D) occur in rust isolates (UK01 and IR01 respectively) that are heterozygous and also contain a recognised allele of *AvrSr50*, so we cannot directly correlate their lack of recognition *in planta* with a virulence phenotype in the rust fungus. Evolution of the *AvrSr50* variants that we report in this study could have arisen under selection on wild relatives carrying resistance genes related to *Sr50* or through occasional infection of rye, which is also a host for *Pgt*. In addition, it cannot be excluded that *AvrSr50* alleles have evolved through the accumulation of neutral variation during long-term clonal reproduction of *Pgt* lineages in the absence of selection for Sr50 virulence. Apart from a single isolate (QCMJC) that was originally isolated from the sexual host barberry (Chen *et al.*, 2017), all of the *Pgt* isolates analysed here contain at least one presumed avirulence allele for *AvrSr50*, consistent with the lack of reported virulence for *Sr50* in *Pgt* to date. A single race, TTRTF, detected in Georgia in 2014 showed partial virulence on *Sr50*, but the genotype is not known (Olivera *et al.*, 2019). However, several isolates in this group, including members of the Ug99, Pgt21 and 'Digalu' race groups (UK-01), are heterozygous and contain a nonrecognised variant at this locus suggesting that virulence could arise in these lineages under field selection by a mutation of a single haplotype.

Identifying and understanding the mode of function of fungal effector proteins is largely hindered by their high sequence diversity, the absence of sequence similarity to proteins of known function, and the lack of conserved domains or functional motifs (Lo Presti *et al.*, 2015). The determination of the three-dimensional structure of effector proteins has been useful by providing molecular insights into biochemical functions, as well as guiding mutagenesis studies to understand how they are recognised by the plant. Here, we determined the first crystal structure of an effector protein from *Pgt*. The structure adopts a β-sandwich fold, with a C-terminal α-helix in which the residue responsible for escaping recognition from Sr50 localises. Out of interest, and given the recent developments in *ab initio* protein modelling, we predicted the structure of avrSr50-B6 using AlphaFold2 (Jumper *et al.*, 2021; Mirdita *et al.*, 2021) with multiple sequence alignment (MSA) generated using MMseqs2 (Steinegger & Söding, 2017). The top model of avrSr50-B6 from AlphaFold (Fig. S17a) has a low confidence score (pLDDT of 40) and comparison with the experimentally determined crystal structure showed a root-mean-squared deviation (RMSD) of 2.63 Å, across a small region of the protein (53 cα atoms (Fig. S17c). Visual inspection of the superimposed structures suggests that the overall structural elements are present but the topology of the predicted model is largely incorrect (Fig. S17a,b). The accuracy of AlphaFold predictions is known to be dependent on the number of sequences in the MSA (Jumper *et al.*, 2021), and on inspection the MSA used by AlphaFold for avrSr50-B6 prediction was confined to four sequences. These data highlight the potential limitations of AlphaFold for fungal effector prediction when low sequence representation exists for an effector of interest.

An emerging theme amongst fungal effector proteins is the conservation of structural folds among some effector families without detectable sequence similarity and sometimes originating from different taxa (for recent reviews please refer to Franceschetti *et al.*, 2017; Mukhi *et al.*, 2020). Based on structural similarity searches using DALI and PDBeFold servers, we did not identify any fungal effectors with high structural similarity to AvrSr50. Despite this, visual inspection of the structures of AvrSr50 and the RNase-like effector BEC1054 from *Blumeria graminis* (PDB: 6FMB) (Pennington *et al.*, 2019) do appear similar. We inspected this further and show via pairwise structural alignment (PDBeFold) both the overall structure and topology are distinct (Fig. S18). These analyses demonstrate that the fold of AvrSr50 is novel amongst fungal effectors, perhaps suggestive of a new effector structural family.

Our structural homology searches revealed similarity of avrSr50-B6 to cupin superfamily members and the CBM of carbohydrate-active enzymes, such as glucosidases and xylosidases. To date, carbohydrate-binding activity in several effector proteins have been shown experimentally. For example, Avr4 from *Cladosporium fulvum* (Hurlburt *et al.*, 2018) and SnTox1 from *Parastagonospora nodorum* (Liu *et al.*, 2016), are able to bind chitin and protect the invading fungi from the hydrolytic activity of plant chitinases. Other effectors are known to modulate chitin-induced immunity by sequestering chitin to dampen the host defence response, such as the lysin motif (LysM) effectors including Ecp6 (Sánchez-Vallet *et al.*, 2013), MpChi from *Moniliophthora perniciosa* (Fiorin *et al.*, 2018) and Tal6 from *Trichoderma atroviride* (Romero-Contreras *et al.*, 2019). It will be interesting to investigate further this structural identity between AvrSr50 and carbohydrate-binding modules by determining the ability of AvrSr50 to bind to different classes of carbohydrates, and the biological relevance of this.

Conclusions

The elucidation of the molecular mechanisms underlying the co-evolution between effector and resistance proteins is important to inform both surveillance of natural pathogen populations and deployment of resistance genes in the field. A complementary knowledge of genetic diversity of interacting proteins coupled to functional analyses is a steppingstone towards the development of durable resistance in crops. In this study, we used the natural

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

genetic variability and structural studies of AvrSr50 to investigate how virulence variants of this protein arise that escape from Sr50 recognition. Functional analyses showed that one residue substitution, Q121K, in the coding sequence was sufficient for AvrSr50 to escape recognition and based on the structure this residue was shown to be surface exposed. These results suggest that single changes in avirulent effector proteins can result on the overcoming of resistance based on major resistance genes. Although *Sr50* provides resistance to almost all known lineages of *Pgt*, several important lineages are heterozygous for avirulence at this locus and may be predisposed to mutate to overcome resistance in *Sr50* wheat lines. We also determined the structure of AvrSr50, which is the first structure of a stem rust effector, and show that it has a novel fold among fungal effectors but shares structural homology to cupin superfamily members and carbohydrate-binding modules.

Acknowledgements

Work described here was supported in part by funding from the 2Blades foundation. SW was supported by an Australian Research Council Future fellowship (FT200100135). JC was supported by a Chinese Scholarship Council (CSC) postgraduate fellowship. IMLS was supported by the Horizon 2020 Framework Programme (742263) and the Daimler and Benz Foundation. We thank Mark Youles in the TSL Synbio team, Adam Bentham and Mark Banfield for providing the modified pOPIN vector. The authors acknowledge the use of the ANU Crystallization and X-ray facility. The authors also acknowledge use of the Australian Synchrotron MX facility and thank the staff for their support. This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. The authors declare no competing interests.

Author contributions

PND and SJW conceptualised the project, acquired funding and supervised the work. DO, JC, MAO, DJE RM, CC, IMLS, SC and NMU acquired experimental data and conducted data analysis. DO supervised the work, drafted the manuscript and all authors contributed to writing, reviewing and editing. DO, JC and MAO contributed equally to this work.

ORCID

Stella Cesari https://orcid.org/0000-0001-8558-0371
Chunhong Chen https://orcid.org/0000-0001-7813-2415
Jian Chen https://orcid.org/0000-0001-8670-6984
Peter N. Dodds https://orcid.org/0000-0003-0620-5923
Daniel J. Ericsson https://orcid.org/0000-0001-5101-9244
Rohit Mago https://orcid.org/0000-0001-5286-073X
Diana Ortiz https://orcid.org/0000-0001-6120-5788
Megan A. Outram https://orcid.org/0000-0003-4510-3575
Isabel M.L. Saur https://orcid.org/0000-0002-5610-1260
Narayana M. Upadhyaya https://orcid.org/0000-0002-3052-0416
Simon J. Williams https://orcid.org/0000-0003-4781-6261

Data availability

The structural data that support the findings of this study are openly available in the Protein Data Bank at http://www.wwpdb.org/ with accession no. 7MQQ. Pgt isolate sequence data used in this study are available in the public domain from the European Nucleotide Archive, accession no. PRJEB22223.

References

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D: Biological Crystallography* **68**: 352–367.

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**: 420–423.

Aragão D, Aishima J, Cherukuvada H, Clarken R, Clift M, Cowieson NP, Ericsson DJ, Gee CL, Macedo S, Mudie N *et al.* 2018. MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian synchrotron. *Journal of Synchrotron Radiation* **25**: 885–891.

Arndell T, Sharma N, Langridge P, Baumann U, Watson-Haigh NS, Whitford R. 2019. GRNA validation for wheat genome editing with the CRISPR-Cas9 system. *BMC Biotechnology* **19**: 1–12.

Arora S, Steuernagel B, Gaurav K, Chandramohan S, Long Y, Matny O, Johnson R, Enk J, Periyannan S, Singh N *et al.* 2019. Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature Biotechnology* **37**: 139–143.

Bentham AR, Youles M, Mendel MN, Varden FA, De la Concepcion JC, Banfield MJ. 2021. pOPIN-GG: a resource for modular assembly in protein expression vectors. *bioRxiv*. doi: 10.1101/2021.08.10.455798.

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S *et al.* 2002. The protein data bank. *Acta Crystallographica Section D: Biological Crystallography* **58**: 899–907.

Bernoux M, Ve T, Williams S, Warren C, Hatters D, Valkov E, Zhang X, Ellis JG, Kobe B, Dodds PN. 2011. Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host and Microbe* **9**: 200–211.

Catanzariti A, Dodds PN, Ve T, Kobe B, Ellis JG, Staskawicz BJ. 2010. The AvrM effector from flax rust has a structured C-terminal domain and interacts directly with the M resistance protein. *Molecular Plant–Microbe Interactions* **23**: 49–57.

Cesari S, Moore J, Chen C, Webb D, Periyannan S, Mago R, Bernoux M, Lagudah ES, Dodds PN. 2016. Cytosolic activation of cell death and stem rust resistance by cereal MLA-family CC-NLR proteins. *Proceedings of the National Academy of Sciences, USA* **113**: 10204–10209.

Chen J, Upadhyaya NM, Ortiz D, Sperschneider J, Li F, Bouton C, Breen S, Dong C, Xu B, Zhang X *et al.* 2017. Loss of AvrSr50 by somatic exchange in stem rust leads to virulence for Sr50 resistance in wheat. *Science* **358**: 1607–1610.

Chen S, Rouse MN, Zhang W, Zhang X, Guo Y, Briggs J, Dubcovsky J. 2020. Wheat gene Sr60 encodes a protein with two putative kinase domains that confers resistance to stem rust. *New Phytologist* **225**: 948–959.

Chen S, Zhang W, Bolus S, Rouse MN, Dubcovsky J. 2018. Identification and characterization of wheat stem rust resistance gene Sr21 effective against the Ug99 race group at high temperature. *PLoS Genetics* **14**: 1–21.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

Cowieson NP, Aragao D, Clift M, Ericsson DJ, Gee C, Harrop SJ, Mudie N, Panjikar S, Price JR, Riboldi-Tunnicliffe A *et al.* 2015. MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian synchrotron. *Journal of Synchrotron Radiation* **22**: 187–190.

Davis IW, Murray LW, Richardson JS, Richardson DC. 2004. MolProbity: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research* **32**: 615–619.

De la Concepcion JC, Franceschetti M, Maqbool A, Saitoh H, Terauchi R, Kamoun S, Banfield MJ. 2018. Polymorphic residues in rice NLRs expand binding and response to effectors of the blast pathogen. *Nature Plants* **4**: 576–585.

De la Concepcion JC, Maidment JHR, Longya A, Xiao G, Franceschetti M, Banfieldi MJ. 2021. The allelic rice immune receptor Pikh confers extended resistance to strains of the blast fungus through a single polymorphism in the effector binding interface. *PLoS Pathogens* **17**: 1–23.

Dodds PN, Lawrence GJ, Catanzariti A-M, Teh T, Wang C-IA, Ayliffe MA, Kobe B, Ellis JG. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences, USA* **103**: 8888–8893.

Ellis JG, Lagudah ES, Spielmeyer W, Dodds PN. 2014. The past, present and future of breeding rust resistant wheat. *Frontiers in Plant Science* **5**: 1–13.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of coot. *Acta Crystallographica Section D: Biological Crystallography* **66**: 486–501.

Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography* **69**: 1204–1214.

Figueroa M, Dodds PN, Henningsen EC. 2020. Evolution of virulence in rust fungi — multiple solutions to one problem. *Current Opinion in Plant Biology* **56**: 20–27.

Fiorin GL, Sanchéz-Vallet A, Thomazella DPDT, do Prado PFV, do Nascimento LC, Figueira AVDO, Thomma BPHJ, Pereira GAG, Teixeira PJPL. 2018. Suppression of plant immunity by fungal chitinase-like effectors. *Current Biology* **28**: 3023–3030.e5.

Franceschetti M, Maqbool A, Jiménez-Dalmaroni MJ, Pennington HG, Kamoun S, Banfield MJ. 2017. Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiology and Molecular Biology Reviews* **81**: 1–17.

Gao Y, Wang W, Zhang T, Gong Z, Zhao H, Han GZ. 2018. Out of water: the origin and early diversification of plant R-genes. *Plant Physiology* **177**: 82–89.

de Guillen K, Lorrain C, Tsan P, Barthe P, Petre B, Saveleva N, Rouhier N, Duplessis S, Padilla A, Hecker A. 2019. Structural genomics applied to the rust fungus *Melampsora larici-populina* reveals two candidate effector proteins adopting cystine knot and NTF2-like protein folds. *Scientific Reports* **9**: 1–12.

Guo L, Cesari S, de Guillen K, Chalvon V, Mammri L, Ma M, Meusnier I, Bonnot F, Padilla A, Peng Y-L *et al.* 2018. Specific recognition of two MAX effectors by integrated HMA domains in plant immune receptors involves distinct binding surfaces. *Proceedings of the National Academy of Sciences, USA* **115**: 11637–11642.

Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A. 2019. Moving beyond P values: data analysis with estimation graphics. *Nature Methods* **16**: 565–566.

Holm L. 2020. DALI and the persistence of protein shape. *Protein Science* **29**: 128–140.

van der Hoorn RAL, Kamoun S 2008. From guard to decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**: 2009–2017.

Hurlburt NK, Chen LH, Stergiopoulos I, Fisher AJ. 2018. Structure of the Cladosporium fulvum Avr4 effector in complex with (GlcNAc)6reveals the ligand-binding mechanism and uncouples its intrinsic function from recognition by the Cf-4 resistance protein. *PLoS Pathogens* **14**: 1–22.

Iverson SV, Haddock TL, Beal J, Densmore DM. 2016. CIDAR MoClo: improved MoClo assembly standard and new E. coli part library enable rapid combinatorial design for synthetic and traditional biology. *ACS Synthetic Biology* **5**: 99–103.

Jacob F, Takaki M, Vernaldi S. 2013. Evolution and conservation of plant NLR functions. *Frontiers in Immunology* **4**: 1–16.

Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* **444**: 323–329.

Jones JDG, Vance RE, Dangl JL. 2016. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**: 1–8.

Jubic LM, Saile S, Furzer OJ, El Kasmi F, Dangl JL. 2019. Help wanted: helper NLRs and plant immune responses. *Current Opinion in Plant Biology* **50**: 82–94.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al.* 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583–589.

Kabsch W. 2010. XDS. *Acta Crystallographica Section D Biological Crystallography* **66**: 125–132.

Kanzaki H, Yoshida K, Saitoh H, Fujisaki K, Hirabuchi A, Alaux L, Fournier E, Tharreau D, Terauchi R. 2012. Arms race co-evolution of *Magnaporthe oryzae* AVR-Pik and rice Pik genes driven by their physical interactions. *The Plant Journal* **72**: 894–907.

Karasov TL, Shirsekar G, Schwab R, Weigel D. 2020. What natural variation can teach us about resistance durability. *Current Opinion in Plant Biology* **56**: 89–98.

Krasileva KV, Dahlbeck D, Staskawicz BJ. 2010. Activation of an arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell* **22**: 2444–2458.

Krissinel E, Henrick K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D: Biological Crystallography* **60**: 2256–2268.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**: 1547–1549.

Lewis CM, Persoons A, Bebber DP, Kigathi RN, Maintz J, Findlay K, Bueno-Sancho V, Corredor-Moreno P, Harrington SA, Kangara N *et al.* 2018. Potential for re-emergence of wheat stem rust in the United Kingdom. *Communications Biology* **1**: 1–9.

Li F, Upadhyaya NM, Sperschneider J, Matny O, Nguyen-Phuc H, Mago R, Raley C, Miller ME, Silverstein KAT, Henningsen E *et al.* 2019. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nature Communications* **10**: 1–15.

Liu Z, Gao Y, Kim YM, Faris JD, Shelver WL, de Wit PJGM, Xu SS, Friesen TL. 2016. SnTox1, a *Parastagonospora nodorum* necrotrophic effector, is a dual-function protein that facilitates infection while protecting from wheat-produced chitinases. *New Phytologist* **211**: 1052–1064.

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. 2015. Fungal effectors and plant susceptibility. *Annual Review of Plant Biology* **66**: 513–545.

Mago R, Zhang P, Vautrin S, Šimková H, Bansal U, Luo M-C, Rouse M, Karaoglu H, Periyannan S, Kolmer J *et al.* 2015. The wheat Sr50 gene reveals rich diversity at a cereal disease resistance locus. *Nature Plants* **1**: 1–3.

Maidment JHR, Franceschetti M, Maqbool A, Saitoh H, Jantasuriyarat C, Kamoun S, Terauchi R, Banfield MJ. 2021. Multiple variants of the fungal effector AVR-Pik bind the HMA domain of the rice protein OsHIPP19, providing a foundation to engineer plant defense. *Journal of Biological Chemistry* **296**: 1–13.

Maqbool A, Saitoh H, Franceschetti M, Stevenson C, Uemura A, Kanzaki H, Kamoun S, Terauchi R, Banfield MJ. 2015. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor *eLife* **4**: 1–24.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *Journal of Applied Crystallography* **40**: 658–674.

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2021. ColabFold - Making protein folding accessible to all. *BioRXiv* doi: 10.1101/2021.08.15.456425.

Mukhi N, Gorenkin D, Banfield MJ. 2020. Exploring folds, evolution and host interactions: understanding effector structure/function in disease and immunity. *New Phytologist* **227**: 326–333.

Ngou BPM, Ahn HK, Ding P, Jones JDG. 2021. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature* **592**: 110–115.

Okuyama Y, Kanzaki H, Abe A, Yoshida K, Tamiru M, Saitoh H, Fujibe T, Matsumura H, Shenton M, Galam DC *et al.* 2011. A multifaceted genomics

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. *The Plant Journal* **66**: 467–479.

Olivera PD, Sikharulidze Z, Dumbadze R, Szabo LJ, Newcomb M, Natsarishvili K, Rouse MN, Luster DG, Jin Y. 2019. 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.

Ortiz D, De Guillen K, Cesari S, Chalvon V, Gracy J, Padilla A, Kroj T. 2017. Recognition of the *Magnaporthe oryzae* effector AVR-Pia by the decoy domain of the rice NLR immune receptor RGA5. *Plant Cell* **29**: 156–168.

Ortiz D, Dodds PN. 2018. Plant NLR origins traced back to green algae. *Trends in Plant Science* **23**: 651–654.

Pennington HG, Jones R, Kwon S, Bonciani G, Thieron H, Chandler T, Luong P, Morgan SN, Przydacz M, Bozkurt T *et al.* 2019. The fungal ribonuclease-like effector protein CSEP0064/BEC1054 represses plant immunity and interferes with degradation of host ribosomal RNA. *PLoS Pathogens* **15**: e1007620.

Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, Deal K, Luo M, Kong X, Bariana H *et al.* 2013. The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science* **341**: 786–788.

Rehmany AP. 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two arabidopsis lines. *Plant Cell* **17**: 1839–1850.

Richardson T, Thistleton J, Higgins TJ, Howitt C, Ayliffe M. 2014. Efficient agrobacterium transformation of elite wheat germplasm without selection. *Plant Cell, Tissue and Organ Culture* **119**: 647–659.

Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**: W320–W324.

Romero-Contreras YJ, Ramírez-Valdespino CA, Guzmán-Guzmán P, Macías-Segoviano JI, Villagómez-Castro JC, Olmedo-Monfil V. 2019. Tal6 from trichoderma atroviride is a LysM effector involved in mycoparasitism and plant association. *Frontiers in Microbiology* **10**: 1–15.

Saintenac C, Zhang W, Salcedo A, Rouse MN, Trick HN, Akhunov E, Dubcovsky J. 2013. Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science* **341**: 783–786.

Salcedo A, Rutter W, Wang S, Akhunova A, Bolus S, Chao S, Anderson N, De Soto MF, Rouse M, Szabo L *et al.* 2017. Variation in the AvrSr35 gene determines Sr35 resistance against wheat stem rust race Ug99. *Science* **358**: 1604–1606.

Sánchez-Vallet A, Saleem-Batcha R, Kombrink A, Hansen G, Valkenburg DJ, Thomma BPHJ, Mesters JR. 2013. Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. *eLife* **2**: 1–16.

Saur IML, Bauer S, Kracher B, Lu X, Franzeskakis L, Müller MC, Sabelleck B, Kümmel F, Panstruga R, Maekawa T *et al.* 2019. Multiple pairs of allelic MLA immune receptor-powdery mildew AVR a effectors argue for a direct recognition mechanism. *eLife* **8**: 1–31.

Saur IML, Hückelhoven R. 2021. Recognition and defence of plant-infecting fungal pathogens. *Journal of Plant Physiology* **256**: 1–14.

Saur IML, Panstruga R, Schulze-Lefert P. 2021. NOD-like receptor-mediated plant immunity: from structure to cell death. *Nature Reviews Immunology* **21**: 305–318.

Shao ZQ, Xue JY, Wang Q, Wang B, Chen JQ. 2019. Revisiting the origin of plant NBS-LRR genes. *Trends in Plant Science* **24**: 9–12.

Sim DW, Kim JH, Kim HY, Jang JH, Lee WC, Kim EH, Park PJ, Lee KH, Won HS. 2016. Structural identification of the lipopolysaccharide-binding capability of a cupin-family protein from *Helicobacter pylori*. *FEBS Letters* **590**: 2997–3004.

Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Bhavani S, Njau P, Herrera-Foessel S, Singh PK, Singh S, Govindan V. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat. *Production* **49**: 465–481.

Skubák P, Pannu NS. 2013. Automatic protein structure solution from weak X-ray data. *Nature Communications* **4**: 1–6.

Steinegger M, Söding J. 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* **35**: 1026–1028.

Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, Hatta A, Ayliffe M, Bariana H, Jones JDG *et al.* 2016. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology* **34**: 652–655.

Tang D, Wang G, Zhou JM. 2017. Receptor kinases in plant-pathogen interactions: more than pattern recognition. *Plant Cell* **29**: 618–637.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2007. Iterative model building, structure refinement and density modification with the PHENIX autobuild wizard. *Acta Crystallographica Section D: Biological Crystallography* **64**: 61–69.

Upadhyaya NM, Mago R, Panwar V, Hewitt T, Luo M, Chen J, Sperschneider J, Nguyen-Phuc H, Wang A, Ortiz D *et al.* 2021. Genomics accelerated isolation of a new stem rust avirulence gene – wheat resistance gene pair. *Nature Plants* **7**: 1220–1228.

Ve T, Williams SJ, Catanzariti A-M, Rafiqi M, Rahman M, Ellis JG, Hardham AR, Jones DA, Anderson PA, Dodds PN *et al.* 2013. Structures of the flax-rust effector AvrM reveal insights into the molecular basis of plant-cell entry and effector-triggered immunity. *Proceedings of the National Academy of Sciences, USA* **110**: 17594–17599.

Wang C-I A, Gunčar G, Forwood JK, Teh T, Catanzariti A-M, Lawrence GJ, Loughlin FE, Mackay JP, Schirra HJ, Anderson PA *et al.* 2007. Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *Plant Cell* **19**: 2898–2912.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* 2011. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**: 235–242.

Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Yoshida K, Tosa Y, Chuma I, Takano Y, Win J *et al.* 2009. Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*. *Plant Cell* **21**: 1573–1591.

Zhang J, Hewitt TC, Boshoff WHP, Dundas I, Upadhyaya N, Li J, Patpour M, Chandramohan S, Pretorius ZA, Hovmøller M *et al.* 2021. A recombined Sr26 and Sr61 disease resistance gene stack in wheat encodes unrelated NLR genes. *Nature Communications* **12**: 1–8.

Zhang W, Chen S, Abate Z, Nirmala J, Rouse MN, Dubcovsky J. 2017. Identification and characterization of Sr13, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. *Proceedings of the National Academy of Sciences, USA* **114**: E9483–E9492.

Zhang X, Farah N, Rolston L, Ericsson DJ, Catanzariti AM, Bernoux M, Ve T, Bendak K, Chen C, Mackay JP *et al.* 2018. Crystal structure of the *Melampsora lini* effector AvrP reveals insights into a possible nuclear function and recognition by the flax disease resistance protein P. *Molecular Plant Pathology* **19**: 1196–1209.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1. T1 progeny of transgenic Fielder lines show susceptibility to leaf rust indicating the absence of nonspecific autoactive resistance response.

Fig. S2. The loop region encompassing residues Thr58 to Arg65 is likely to occupy multiple conformation states.

Fig. S3. AvrSr50-A1 is specifically recognised by Sr50 and induces cell death in wheat protoplasts.

Fig. S4. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *AvrSr50-A1* and *avrSr50-B6* mutant variants with *p35S-Sr50*.

Fig. S5. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing $AvrSr50\text{-}A1_{V27I}$ and $AvrSr50\text{-}A1_{I29T}$ with *p35S-Sr50*.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation

Fig. S6. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *AvrSr50-A1*$_{V61E}$ and *AvrSr50-A1*$_{V27I/I29T}$ with *p35S-Sr50*.

Fig. S7. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *AvrSr50-A1*$_{V93IE}$ and *AvrSr50-A1*$_{Q121K}$ with *p35S-Sr50*.

Fig. S8. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *AvrSr50-A1*$_{V27I}$ and *AvrSr50-A1*$_{I29T}$ with *pUbq-Sr50*.

Fig. S9. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *AvrSr50-A1*$_{V27I/I29T}$ and *AvrSr50-A1*$_{V61E}$ with *pUbq-Sr50*.

Fig. S10. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *AvrSr50-A1*$_{V93I}$ and *AvrSr50-A1*$_{Q121K}$ with *pUbq-Sr50*.

Fig. S11. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *avrSr50-B6*$_{K121Q,}$ *avrSr50-B6*$_{E61V}$ and *avrSr50-B6* $_{I27V}$ with *pUbq-Sr50*.

Fig. S12. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing a*vrSr50-B6*$_{K121Q}$, *avrSr50-B6*$_{E61V}$ and *avrSr50-B6*$_{I27V}$ with *p35S-Sr50*.

Fig. S13. Protein alignment of AvrSr50 natural variants from 46 different rust isolates

Fig. S14. Polymorphic residues of AvrSr50 natural variants.

Fig. S15. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing naturally occurring *AvrSr50 variants* with *pUbq-Sr50*.

Fig. S16. Cell death response of *Nicotiana benthamiana* leaves transiently co-expressing *p35S-Sr50* with naturally occurring *AvrSr50* variants.

Fig. S17. Structural alignment of the AlphaFold predicted and experimentally determined structure of avrSr50-B6.

Fig. S18. Structural alignment of the crystal structure of AvrSr50-B6 and BEC1054.

Table S1 List of plasmids used in this study.

Table S2 List of primers used in this study.

Table S3 X-ray data collection, structure solution and refinement statistics for AvrSr50-B6.

Table S4 Segregation of stem rust resistance in transgenic wheat lines.

Table S5 Data of rust isolates analysed in this study.

Table S6 The top hits from the PDBeFold showing structural similarity to carbohydrate-binding domains and cupin superfamily members.

Table S7 The top hits from the DALI server showing structural similarity to carbohydrate-binding domains.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation