The AvrPm3-Pm3 effector-NLR interactions control both race-specific resistance and host-specificity of cereal mildews on wheat

Salim Bourras1,2,7, Lukas Kunz1,7, Minfeng Xue3,4,5,7, Coraline Rosalie Praz1, Marion Claudia Müller1, Carol Kälin1, Michael Schläfi1, Patrick Ackermann1, Simon Flückiger1, Francis Parlane1, Fabrizio Menardo1, Luisa Katharina Schaefer1, Roi Ben-David6, Stefan Roffler1, Simone Oberhaensli1, Victoria Widrig1, Stefan Lindner1, Jonatan Isaksson1, Thomas Wicker1, Dazhao Yu3,4,5 & Beat Keller1

The wheat Pm3 resistance gene against the powdery mildew pathogen occurs as an allelic series encoding functionally different immune receptors which induce resistance upon recognition of isolate-specific avirulence (AVR) effectors from the pathogen. Here, we describe the identification of five effector proteins from the mildew pathogens of wheat, rye, and the wild grass Dactylis glomerata, specifically recognized by the PM3B, PM3C and PM3D receptors. Together with the earlier identified AVRPM3A2/F2, the recognized AVR of PM3B/C (AVRPM3B2/C2), and PM3D (AVRPM3D3) belong to a large group of proteins with low sequence homology but predicted structural similarities. AvrPm3b2/c2 and AvrPm3d3 are conserved in all tested isolates of wheat and rye mildew, and non-host infection assays demonstrate that Pm3b, Pm3c, and Pm3d are also restricting the growth of rye mildew on wheat. Furthermore, divergent AVR homologues from non-adapted rye and Dactylis mildews are recognized by PM3B, PM3C, or PM3D, demonstrating their involvement in host specificity.
Fungi are highly adaptive, widespread organisms and ubiquitous pathogens of animals and plants responsible for population declines and pandemics in natural and agricultural ecosystems\textsuperscript{1–5}. In wheat, there have been two recent fungal pandemics caused by the emergence of wheat blast\textsuperscript{6} and stem rust (Ug99).\textsuperscript{7} The Sr35 resistance (R) gene encodes a nucleotide binding-leucine rich repeat (NLR) immune receptor that upon recognition of the avirulence protein AVRSr35 confers near-immunity to the avirulence protein AVRSr35.\textsuperscript{22,25,27} This case demonstrates that effector-NLR interactions conferring immunity following the simplest model of Avr–R gene-for-gene interactions\textsuperscript{3,11} provide an important leverage for achieving rapid control of emerging disease pandemics in crops.

Powdery mildews are agronomically important fungal pathogens of large numbers of wild and cultivated species. In cereals, the powdery mildew disease is caused by a single species, Blumeria graminis. It occurs in different, highly host-specific forms called formae speciales\textsuperscript{12,13}. For example, B. graminis f. sp. tritici (B. gr. tritici) exclusively grows on wheat, whereas B. graminis f. sp. secalis (B. gr. secalis) and B. graminis f. sp. dactylidis (B. gr. dactylidis) specifically infect rye and very wild grass species Dactylis glomerata, respectively.\textsuperscript{11} Powder mildews are obligate biotrophs, growing and reproducing exclusively on living host tissue. This requires suppression of host immune responses which is possibly achieved through the massive secretion of small virulence proteins called effectors.\textsuperscript{14,15} Recent studies in barley powdery mildew have established the role of some effectors as virulence factors.\textsuperscript{16–21} Effector proteins can be specifically recognised by NLR plant immune receptors whose activation confers resistance to the invading pathogen. This isolate-specific resistance is often linked to a rapid cell death response at sites of attempted infection, designated as the hypersensitive response (HR).\textsuperscript{22,23} Effectors recognised by R-proteins are called avirulence (AVR) proteins and are often found to be polymorphic in different isolates of the pathogenic fungus.\textsuperscript{15}

Allelic series of NLR-type receptors have been described in several plant species. The largest allelic series of resistance genes in wheat is formed by the Pm3 gene against powdery mildew. Up to date, 17 functional alleles (Pm3a-g, Pm3k-Pm3t) have been cloned and functionally validated\textsuperscript{24,25} and they confer complete resistance against distinctly different sets of B. gr. tritici races\textsuperscript{26}. Compared to other well studied allelic series of R genes such as the flax rust (Melampsora lini) resistance gene L or the B. gr. hordei resistance gene Mla, the Pm3 series stands out due to its very high (>97%) sequence identity on the protein level\textsuperscript{22,25,27}.

The molecular basis of the functional diversity of the closely related Pm3 alleles is largely unknown. Up to date, the only known avirulence gene for a Pm3 allele is AvrPm3\textsuperscript{82/2,2}, an effector gene recognised by the Pm3a and Pm3f alleles.\textsuperscript{28} Despite high sequence conservation among the Pm3 NLRs, no allele other than Pm3a and Pm3f can recognise AvrPm3\textsuperscript{82/2} or any of its closest gene family relatives or natural haplotype variants.\textsuperscript{28,29} Furthermore Bourras and collaborators\textsuperscript{28} provided evidence that the Pm3 resistance follows a genetically complex gene-for-gene model\textsuperscript{11}, which involves a pathogen encoded suppressor of avirulence (SvrPm3\textsuperscript{82/2}) acting on several AvrPm3-Pm3 specificities.

In this study, we describe the identification of the AvrPm3\textsuperscript{82/2} and AvrPm3\textsuperscript{82} genes and provide molecular evidence that specificity of the Pm3 NLRs is based on recognition of highly sequence diverse, but structurally similar effectors. We also demonstrate that the same effector genes are conserved in the non-adapted rye and Dactylis powdery mildews, thus demonstrating that the Pm3 NLRs, apart from their race-specific resistance function, are potent determinants of host-specificity for grass mildews. Therefore, we propose that the AvrPm3-Pm3 interactions provide a unique model system to understand how NLRs can contribute to both host and non-host resistance.

Results

Identification of AvrPm3 candidates by effector benchmarking. In a first approach to identify AvrPm3 candidate genes we established a new assay based on the hypothesis that effector proteins acting as AVR factors are likely to share structural similarities, sequence polymorphism, and similar expression patterns at the haustorial stage. At the time this assay was designed, four mildew Avrs were cloned: AvrPm3\textsuperscript{82/2}, AvrPm2 from B. gr. tritici\textsuperscript{28,29} and AvrPm1, AvrPm1\textsubscript{3} from B. gr. hordei\textsuperscript{22}. They are all encoded by typical effector proteins, ranging in size from 118 (AVRLa), to 130 aa (AVRPhS), with a predicted N terminal signal peptide, two conserved cysteine residues (except AVRLa which contains only one), and high expression at the haustorial stage.\textsuperscript{28,30,31} Therefore, we hypothesised that mildew AVRys correspond to short proteins of ca. 120 aa on average, encoded by highly expressed candidate effector genes, that can be differentiated in mildew isolates with contrasting virulence patterns based on sequence polymorphism or gene expression levels.

We scanned the genomes of the three wheat powdery mildew reference isolates, Bgt_96224, Bgt_94202 and Bgt_JIW2, which are polymorphic for several AvrPm3 specificities\textsuperscript{28,32}, and systematically classified and scored 580 effectors based on sequence polymorphism (SNPs and deletions), presence of a functional signal peptide, cysteine content, native protein size, and gene expression levels (Supplementary Note 1, Supplementary Fig. 1). This resulted in an “effector benchmarking” list with effector scores for candidate AvrPm3 genes ranging from −8 for the worst candidate (BgtE-5642) to +18 for the two best ones (BgtavrF2-9 and BgtE-20041) (Supplementary Note 1, Supplementary Data 1, Supplementary Fig. 1). AvrPm2 and AvrPm3\textsuperscript{82/2} were used as a control for assessing Avr candidate scoring, and they ranked in the top 20, with a total score of +17 and +16, respectively. We selected all effector genes scoring at least +10 (i.e. 100 candidates) for further functional validation (Supplementary Data 1).

Identification of AvrPm3 candidates using GWAS. In a second approach to identify AvrPm3\textsuperscript{82}, AvrPm3\textsuperscript{82} and AvrPm3\textsuperscript{82} candidates, we sequenced 72 additional wheat powdery mildew genomes to complete a GWAS (genome wide association study) mapping population of 100 races originating exclusively from China (Supplementary Data 2). Genetic association between sequence polymorphisms and differences in virulence/avirulence patterns on Pm3b and Pm3d was assessed using the Genome Association and Prediction Integrated Tool (GAPIT)\textsuperscript{33}. We found no significant association for AvrPm3\textsuperscript{84}. We found significantly associated SNPs for AvrPm3\textsuperscript{82} on chromosome 5\textsuperscript{54}, with a peak at position 18,860,696 (Fig. 1a, b). This region overlaps with the physical position of Locus 3, a genetic locus that has been previously described as encoding for the AvrPm3\textsuperscript{82} and AvrPm3\textsuperscript{82} specificities (Supplementary Note 2). The position of the AvrPm3\textsuperscript{82} peak was located within the genetic interval defined by the Locus 3 flanking markers M049LE and ctg118_21 in the powdery mildew consensus genetic map (Supplementary Fig. 2).

We used a combination of resources to thoroughly annotate this genetically complex locus (Supplementary Note 2). We found in total four candidate effector genes in the locus: Bgt-51460, BgtE-20002, Bgt-55150 and BgtE-ng2, all members of the effector gene family E018\textsuperscript{34} (Supplementary Fig. 3). Bgt-51460 is not expressed and not polymorphic between the Pm3b/c avirulent Bgt_96224, and the Pm3b/c virulent Bgt_94202 parents. BgtE-20002 is not polymorphic between the isolates Bgt_96224 and Bgt_{94202}, and it was also identified as a candidate Avr by the benchmarking approach with an overall score of +12.
An initial search for sequence variants in a few isolates from the Swiss population, led to the identification of BgtE-20002_I, a haplotype from isolate Bgt_07237 (avirulent on Pm3b, Pm3c and Pm3d) differing from the BgtE-20002 variant by 1 non-synonymous SNP (Fig. 1c). Bgt-55150 carries 2 SNPs leading to amino acid changes distinguishing the Bgt-55150 alleles from the avirulent Bgt_96224 and the virulent Bgt_94202 isolates (Supplementary Fig. 4). BgtE-ng2 is also polymorphic with one functional allele, BgtE-ng2b, encoded by the virulent isolate Bgt_94202, while the allele encoded by the avirulent Bgt_96224 isolate, BgtE-ng2a is disrupted by a transposable element insertion (Supplementary Fig. 4). Based on the GWAS results, genetic information from the powdery mildew consensus map\textsuperscript{28}, and thorough annotation of Locus\_3, we selected Bgt-51460, BgtE-20002, BgtE-20002_I, Bgt-55150a, Bgt-55150b and BgtE-ng2b for further functional validation (Supplementary Note 2, Supplementary Data 3 and 4).

Functional validation of AvrPm3\textsuperscript{b2/c2} and AvrPm3\textsuperscript{d3}. To functionally validate the AvrPm3 candidates identified by effector benchmarking and GWAS, we took advantage of our well established agrobacterium-infiltration assay in N. benthamiana\textsuperscript{28–30}, allowing transient, heterologous, co-overexpression of effector and
NLR proteins. One hundred candidates from the effector benchmarking and all six candidates from the GWAS were codon optimised for expression in *N. benthamiana* (Supplementary Data 3), synthesised without the signal peptide, and screened for recognition by *Pm3b*, *Pm3c* and *Pm3d* using our transient assays in *N. benthamiana* leaves (see Methods). We took advantage of the high sensitivity reading (HSR) technology embedded in the Fusion FX fluorescence imager (www.vilber.com) to quantitatively score the hypersensitive response 5 days after infiltration. One *Avr* candidate, *BgtE-20002*, commonly found by benchmarking and GWAS, induced strong HR when combined with *Pm3b* and weak HR with *Pm3c*, while the sequence variant *BgtE-20002_J*, induced strong HR both with *Pm3b* and *Pm3c* (Fig. 1d, Supplementary Fig. 5a, b). We observed no HR when both variants were co-infiltrated with *Pm3a*, *Pm3f*, and *Pm3e*, indicating that this effector is implicated in dual and specific recognition by the *Pm3b* and *Pm3c* alleles that have previously been shown to exhibit overlapping recognition spectra, with *Pm3c* representing the weaker allele. Therefore, we concluded that the effector is *AvrPm3b2/c2*, the *Avr* recognised by *Pm3b* and *Pm3c*. Also, because *BgtE-20002* is not polymorphic between the *Pm3b/c* avirulent isolate (Bgt 96224) and the virulent isolate (Bgt 94202), we conclude that the effector is *AvrPm3b2/c2* is genetically suppressed by the active *svrPm3a1/f1* allele encoded in Bgt 94202, in a similar way as *AvrPm3b2/c2*. A second *Avr* candidate, *BgtE-20069b*, only found in the benchmarking list, and encoded within a segmental duplication on chromosome 9 (Fig. 1e–g), induced HR in combination with *Pm3d* (Fig. 1h, Supplementary Fig. 5c). No HR was observed when *BgtE-20069b* was combined with *Pm3a*, *Pm3f*, *Pm3b*, *Pm3c* and *Pm3e*, indicating that this effector gene is *AvrPm3d3*, the *Avr* recognised by *Pm3d*. The *BgtE-20069b* copy corresponding to *AvrPm3d3* is deleted in the virulent isolate Bgt 94202. Its paralog, *BgtE-20069a*, has two non synonymous SNPs compared to the *Avr*, and it is polymorphic between Bgt 96224, and Bgt 94202 (Fig. 1g). We used the same transient assays in *N. benthamiana* to test for possible recognition of both alleles of *BgtE-20069a* by *Pm3d*, and we observed no HR (Fig. 1h). All together this data demonstrates that specificity of the *AvrPm3d3-Pm3d* interaction is based on specific recognition of one paralog of a duplicated effector gene, *BgtE-20069b*.

In a last series of functional validation assays, we attempted epitope tagging of AVRP3M3A2/F2, AVRP3M3B2/C2 and AVRP3M3D3. All effector proteins were C and N terminally fused to HA and FLAG epitope tags, and subsequently tested for (i) detectability on a western blot, and (ii) functionality in the *N. benthamiana* assay (Fig. 2, Supplementary Note 3). For AVRP3M3A2/F2, N terminal HA and FLAG fusions were detectable but not the C terminally tagged constructs. We observed the opposite for AVRP3M3B2/C2, where the C terminal HA and FLAG fusion were detectable but not the N terminal ones. For AVRP3M3D3, none of the epitope fusions was detectable on the western blot (Fig. 2a–c). We have also tested the functionality of the detectable HA-AVRP3M3A2/F2, AVRP3M3B2/C2-HA constructs in co-infiltration assays with their respective PM3 NLKs, and found that these epitope fusions did not interfere with AVR function (Fig. 2d–f, Supplementary Fig. 6). Together, these assays demonstrate that AVRP3M protein function is sensitive to epitope fusions in a tag sequence and position-dependent manner.

**AvrPm3** recognition is suppressed by *Svr* and NLR interactions. Genetic studies in F1 segregating populations of wheat powdery mildew have revealed that *SvrPm3a1/f1* is genetically suppressing several *Avr* of the PM3 genes (including *AvrPm3b2/c2* and *AvrPm3d3*), but so far functional evidence was only provided for *AvrPm3a2/f2* [28]. Here, with the cloning of two novel *AvrPm3* genes, we had a unique opportunity to functionally validate that *SvrPm3a1/f1* can act as a suppressor of *Avr* recognition for additional (probably all) *AvrPm3* genes [28]. We co-expressed *AvrPm3b2/c2*, and *AvrPm3d3* with *Pm3b* and *Pm3d*, respectively, in presence of the active *svrPm3a1/f1* allele originating from the mildew isolate Bgt 94202 [28]. HR was assessed relative to a control where the active *svrPm3a1/f1* suppressor allele was replaced by the inactive *svrPm3a1/f1* allele encoded by the mildew isolate Bgt 96224, as previously described [28]. As an additional control, we also assayed suppression of the *AvrPm3a2/f2-Pm3a* interaction using the same experimental set-up (see Methods). As previously demonstrated [28], *AvrPm3a2/f2* recognition was suppressed by *SvrPm3a1/f1* and resulted in significant reduction of HR intensity in our control assay (Fig. 3a), thus demonstrating the functionality of our improved experimental system based on codon-optimised constructs and HSR imaging. Consistent with previous genetic evidence from Bourras and colleagues [28], we found in two independent assays that recognition of *AvrPm3b2/c2* and *AvrPm3d3* was also suppressed by *SvrPm3a1/f1* and resulted in significant reduction of HR intensity (Student t-test, p-value < 0.05) (Fig. 3b, c). We also produced C and N terminal HA and FLAG epitope tag fusions of the active *svrPm3a1/f1* variant and found that all constructs were detectable on the western blot, suggesting significant differences between AVRP3M and SVRP3M proteins in terms of tolerance to epitope fusions (Fig. 3d). We have therefore tagged the inactive *svrPm3a1/f1* variant that we use as negative control for AVRP3M suppression, and tested for protein expression levels as compared to the active suppressor. We found no difference in protein abundance.
between HA-SVRPM3^A1/F1, and HA-svrPM3^A1/F1 (Fig. 3e), and we could show that the tag is also not interfering with SVRPM3^A1/F1 function (Fig. 3f, g). These data demonstrate that differences between active and inactive suppressor variants are not due to differences in protein expression. Based on this data, we could confidently test the hypothesis that SvrPm3^A1/F1 is possibly acting as a transcriptional or translational suppressor of AvrPm3 or Pm3, based on a previous observation suggesting this effector encodes a fungal ribonuclease-like protein\textsuperscript{32}. To do so, we tested protein expression levels of AVRPM3^A2/F2, AVRPM3^B2/C2 and individual PM3 NLRs in presence vs. absence of SVRPM3^A1/F1 and found no differences (Supplementary Fig. 7). These results demonstrate that SvrPm3^A1/F1 can indeed act as a suppressor of recognition of several AvrPm3 genes, and suggest this function is not based on an AVRpm3 or Pm3-specific ribonuclease activity but occurs at the level of protein–protein interactions.

In addition to gene suppression based on the action of SvrPm3^A1/F1, evidence from genetic studies in wheat and transient assays with auto-activated PM3 constructs\textsuperscript{26,28}, indicated that the AvrPm3-Pm3 interactions can also be suppressed by inter-allelic interactions between the PM3 NLRs.

Here, with the cloning of additional AvrPm3 genes and the availability of an improved cell death assay in \textit{N. benthamiana}, we had a unique opportunity to test for inter-allelic NLR-NLR interactions in presence of the AVR proteins. We assayed the suppression activity of Pm3a-f and the non-functional ancestral allele Pm3CS in transient co-expression assays when each allele was combined with AvrPm3^B2/C2, AvrPm3^B2/C2-I, AvrPm3^D3-I, and AvrPm3^D3 (Supplementary Figs. 8 and 9, Supplementary Note 4). We found that Pm3a and Pm3f had no NLR suppression activity and were mostly suppressed by the other Pm3 alleles. Pm3e was only active on the weaker Pm3c and Pm3f alleles, while Pm3b, Pm3c and Pm3d had reciprocal suppression capacity and acted as the strongest suppressors among the functional Pm3 NLRs. Here we conclude that inter-allelic suppression occurs upon AVR-dependent activation of the PM3 NLRs, and it is NLR sequence dependent.

**The AVRPM3 proteins belong to a group of related effectors.** AvrPm3^B2/C2 and AvrPm3^D3 are not members of the AvrPm3^B2/C2 effector gene family E008 but belong to the mildew effector family.
families E018 and E034, respectively (Supplementary Figs. 3 and 10). Sequence alignment of the AVRPM3A2/F2, AVRPM3B2/C2, and AVRPM3D3 full proteins showed that the three AVRs share low identity (14.5 to 19.9%, Supplementary Fig. 11). Sequence alignment of the three effector families revealed conservation of the signal peptide, a conserved Hydrophobic x C (HxC) motif, and a characteristic conserved pattern of alternating hydrophobic residues, (Supplementary Fig. 12). Analysis of the phylogenetic relationships between the AVRPM3 families suggests that AVRPM3A2/F2 and AVRPM3B2/C2 belong to the same superfamily which groups separately from AVRPM3D3, and all three families are phylogenetically distinct from the SVRPM3A1/F1 family (Supplementary Fig. 13).

The PM3 NLRs are highly similar on the protein level with only few amino acid differences. We therefore hypothesised that a basis for specificity could be that the sequence unrelated AVRPM3 effectors are encoding for structurally similar proteins, based on conservation of amino-acid properties and the position of structural residues such as cysteines and prolines (Supplementary Fig. 12). We applied AVRPM3A2/F2, AVRPM3B2/C2, and
AVRP3D3 to secondary structure modelling using four different prediction algorithms available at the Quick2D platform (https://toolkit.tuebingen.mpg.de/#/tools/quick2d). All modelling methods have consistently predicted the presence of one alpha helix followed by three to four beta-sheets for all three AVRs (Supplementary Fig. 14a). To further substantiate these results, we applied the AVRP3A2/F2, AVRP3B2/C2, and AVRP3D3 effector family members to three dimensional structural modelling using the RaptorX platform (raptorx.uchicago.edu37). We found no statistically robust structural model common to all three families or to all members of the same family (Supplementary Data 5). However, despite differences in the templates assigned by RaptorX to individual members, we found that several of the tertiary folds predicted for the AVRP3A2/F2 and AVRP3B2/C2 families consisted of one central helix facing three to four beta-sheets, reminiscent of the previously predicted secondary folds (Supplementary Fig. 14b, c). In particular, this putative tertiary structure was also predicted for the AVRP3A2/F2 and AVRP3B2/C2 avirulence proteins. Therefore, based on protein alignments, phylogeny, and in silico secondary and tertiary structure modelling, we suggest that AVRPM3A2/F2 and AVRPM3B2/C2 are possibly encoding for structurally similar proteins. Finally, all AvrPm3 genes are highly induced upon infection, and all encode for typical pathogen effector proteins (Supplementary Note 5, Supplementary Fig. 15). We therefore propose that the AvrPm3 genes are possibly acting as bona-fide virulence factors suppressing host immunity.

AvrPm3b2/c2 and AvrPm3d3 conservation in mildew populations. In a previous study, McNally and colleagues29 found limited sequence diversity for AvrPm3b2/c2 in worldwide mildew populations, and a high frequency of the recognised haplotype variant. These observations are contrasting with the general hypothesis arguing that avirulence genes are highly polymorphic in pathogen populations, and that active variants tend to be lost because of selective pressure to escape recognition from the host immune system29. To study the genetic diversity of AvrPm3b2/c2 and AvrPm3d3, we applied a collection of 185 powdery mildew isolates with a worldwide geographical distribution to haplotype mining as previously described by McNally and colleagues29. The full sequences of AvrPm3b2/c2 and AvrPm3d3 were recovered using genome sequencing data or PCR gene amplification and subsequent sequencing (see Methods). For AvrPm3b2/c2, we identified 10 non-synonymous mutations across the gene, including 8 in the mature protein resulting in 7 protein variants (Fig. 4a, b, Supplementary Fig. 16). Based on those haplotype sequences, we found a pN/pS (ratio of polymorphic nonsynonymous to polymorphic synonymous) sites within a species of 1.26 indicating that the AvrPm3b2/c2 gene is under diversifying selection, which is consistent with previous finding by Müller et al.34 and McNally et al.29 showing that the whole AvrPm3b2/c2 and AvrPm3d3 families are under diversifying selection. All haplotypes of AvrPm3b2/c2 were codon optimised for expression in N. benthamiana, synthesised without the signal peptide and screened for recognition by Pm3b and Pm3c. No haplotype besides the previously validated AvrPm3b2/c2-A variant from Bgt_96224 and the AvrPm3b2/c2-I from Bgt_07237, induced HR when combined with Pm3b and Pm3c in N. benthamiana (Fig. 4b). Protein expression assays demonstrated that loss of AVRPM3B2/C2 recognition is not due to absence of the protein as all haplovariants are equally detectable on a western blot (Fig. 4c, Supplementary Fig. 17). We found that the AvrPm3b2/c2-I is only present in the Swiss population and represented by only one isolate. Interestingly, the AvrPm3b2/c2-A variant was the most frequent haplotype found in all sub-populations, and in two regions, USA and Israel, this active Avr variant was more frequent than the active AvrPm3b2/c2-A (Fig. 4a). For AvrPm3d3, we identified 17 non-synonymous mutations across the whole gene, including 15 in the mature protein. However, it was not possible to distinguish between the SNPs that should be assigned to a haplotype of the BgtE-20069a or the BgtE-20069b paralogs (i.e. AvrPm3d3) which only differ by two residues (Fig. 1g). Therefore, while we could not resolve the haplotype diversity of AvrPm3d3 which shows high level of copy number variation in mildew isolates (AvrPm3b2/c2 is present in a single copy only, see Supplementary Fig. 18), these results show that at least one paralog of the AvrPm3d3 effector can be found globally in all isolates. Here we conclude that data from diversifying selection studies and global population genetics provide further evidence substantiating that the AvrPm3 effector are important virulence factors.

We also assessed the individual impact on AVR recognition of the 7 amino-acid variations in the mature protein of AVRPM3B2/C2 found in our natural diversity (residue changes in the signal peptide were not tested) (Fig. 4d). For AVRPM3D3, 14 amino-acid variations were similarly assessed which very likely include polymorphic residues from its paralog BgtE-20069a, because these could not be distinguished in our haplotype mining (Fig. 4e). Every single amino acid variation was used to generate point mutants in AVRPM3B2/C2 and AVRPM3D3, and all were screened for recognition by their respective NLRs using transient assays in N. benthamiana as previously described29. We found that most of the single amino acid mutations disrupted the recognition of AVRPM3B2/C2 and AVRPM3D3 by their respective
NLRs (Fig. 4d, e), suggesting that selective pressure to escape R gene recognition is a driving force for sequence diversification of AVRPM3B2/C2 and AVRPM3D3. Protein expression analysis demonstrated that loss of AVRPM3B2/C2 recognition is not due to absence of the protein as all point mutants are equally detectable on a western blot (Supplementary Fig. 19). Interestingly, we found that most of the polymorphic positions in AVRPM3B2/C2, including 3 out of the 4 disruptive mutations, are located towards the N-terminal end within a stretch of 18 residues (aa 25–42), immediately after the predicted signal peptide.
peptide cleavage site (Fig. 4d). By contrast, most of the natural diversity in AVRPM3 D3 and its duplication BgtE-20069a, including 5 out of the 8 residues disrupting recognition, is found within a stretch of 24 residues (aa 74–97) towards the C-terminal end of the protein (Fig. 4e). This data suggests that specific domains of the AVR proteins are involved in recognition, and these seem to be located in distinctly different regions of AVRPM3B2/C2 and AVRPM3D3.

We therefore designed domain swaps covering different stretches of residues polymorphic between AVRPM3B2/C2 and AVRPM3D3 and the closest members of their effector families, BGT-51460 and BGTE-5883, respectively (Fig. 5a, b, Supplementary Fig. 20a, b, Supplementary Note 6). For both AVR proteins we found that NLR-recognition is dependent on multiple distinct protein regions that overlap with regions with higher level of polymorphism in worldwide mildew accessions (Fig. 5b, c–e, d, e).

**Fig. 5** Consequence of synthetic domain swaps on the recognition of AVRPM3B2/C2. **a–b** Consequence of synthetic domain swaps on the AVRPM3B2/C2 AVR function. **a** Protein sequence alignment of the mature protein of AVRPM3B2/C2-A and the closest family member Bgt-51460. Swapped domains are indicated and labelled a, b, c and d. Domains that do not abolish AVR recognition when introduced from Bgt-51460 into AVRPM3B2/C2 are indicated in red. **b** Schematic representation of the protein domains swapped between AVRPM3B2/C2 (grey) and Bgt-51460 (blue). **c–e** Transient expression assays in N. benthamiana indicating recognition (upper panel) and HR quantification (lower panel) for AVRPM3B2/C2 swaps #6 (**c**), swap #8 (**d**) and swap #9 (**e**) by Pm3c, and Pm3b. The number or independent leaf replicates is indicated. Mean values are indicated by the middle line in the boxplot. Individual data points are plotted along the whiskers delineating minimum and maximum values. Statistical significance was assessed with a two-sided Student’s t-test for paired data and indicated with (*p < 0.05). Complete N. benthamiana leaf pictures and raw data underlying the reported averages are provided in a Source Data File.
The Pm3 alleles can recognise AVRs from non-adapted mildews. There has been accumulating genetic evidence suggesting that Avr-R interactions are playing an important role as determinants of host-specificity in cereal mildews (Supplementary Note 7). Therefore, we aimed at investigating the role of the Pm3 alleles as possible determinants of host specificity to B. g. secalis and B. g. dactylidis, two mildew *formae speciales* only growing on rye (*Secale cereale*) or the wild grass *Dactylis glomerata*, respectively. We took advantage of the availability of the genome sequences of 5 isolates of rye mildew and 2 of *Dactylis* mildew to recover the sequences of the homologues of AvrPm3^3^, AvrPm3^2^, AvrPm3^2^, and AvrPm3^3^D. We found homologues of all three AvrPm3 genes in B. g. dactylidis (hereafter referred to as AvrPm3^2^ - Bgd, AvrPm3^2^ - Bgd, and AvrPm3^3^ - Bgd) and homologues of AvrPm3^2^ - Bgd and AvrPm3^3^ in B. g. secalis (AvrPm3^2^ - Bgs, and AvrPm3^3^ - Bgs) (Supplementary Fig. 22a-c). We found no homologue of AvrPm3^2^ in rye mildew. RNA-Seq data from two rye mildew isolates S-1391 and S-1459 show that AvrPm3^2^ and Bgs are expressed during compatible interaction between B. g. secalis and rye (Supplementary Fig. 22d, e). All AvrPm3 homologues were tested for R gene recognition using the transient assay in *N. benthamiana* (Fig. 6a-c, Supplementary Fig. 23a-g). We observed no HR when AvrPm3^2^ - Bgd was co-expressed with Pm3a or Pm3f (Fig. 6a). For AvrPm3^2^ - Bgd, co-expression of the homologue from B. g. dactylidis with Pm3b and Pm3c resulted in no HR, while interestingly the homologue from B. g. secalis was recognised by both NLRs (Fig. 6b). Similarly, both AvrPm3D homologues from rye and *Dactylis* powdery mildews were recognised by Pm3d (Fig. 6c). For AvrPm3^2^ - BGD, protein expression data indicate that loss of recognition is not due to absence of the proteins (Supplementary Fig. 23j, k). However, data for AvrPm3^2^ - BGD indicate this protein is low expressed in our assay (Supplementary Fig. 23h,i), therefore we cannot exclude that AvrPm3^2^ - BGD is also recognised by Pm3a/f if expressed at higher levels.

The recognised AvrPm3^2^C, BGS, AvrPm3^3^D, BGS and AvrPm3^3^D, BGS homologues differ by 4, 7 and 6 residues in the mature protein from their B. g. tritici encoded homologues respectively, demonstrating a higher level of sequence divergence as compared to the haplotypes of the AvrPm3 genes in B. g. tritici. Finally, we also tested possible recognition of the rye and *Dactylis* mildew AvrPm3 homologues by the rye powdery mildew resistance gene Pm8 in transient assays in *N. benthamiana* and observed no HR (Supplementary Fig. 24). Altogether, this data indicates strong conservation of AvrPm3 function and recognition specificity in grass mildews, and suggests that highly specific AvrPm3-Pm3 interactions are involved in defence against non-host mildew forms.

To further substantiate these results, we designed a micro-phenomics assay to test for wheat susceptibility to non-host mildews in presence vs. absence of the Pm3 genes. Transgenic wheat lines expressing the Pm3b and Pm3d resistance alleles were infected with two isolates from the non-adapted rye powdery mildew (Bgs_1391 and Bgs_1459). As no macroscopic growth or sporulation of rye mildew can be observed in this non-host interaction, we microscopically phenotyped these assays for micro-colony formation, i.e., the ability of mildew to establish a haustorium and few secondary hyphae (see Methods). Transgenic lines were compared to the non-transgenic 'Bobwhite' genotype, i.e., the identical genetic background of the Pm3b and Pm3d transgensics. We found a significant reduction of the ability of the non-adapted rye powdery mildew to form micro-colonies on the transgenic lines as compared to the non-transgenic control (Fig. 6d). To further extend these observations, we performed the same assay using the Pm3b and Pm3c near isogenic wheat lines (NILs) 'Chul_8xChancellor' and 'Sonora_8xChancellor', respectively. The backcrossing line 'Chancellor' corresponding to the genetic background of the Pm3b and Pm3c NILs, was used as a control. In agreement with our first assay, we also found a significant reduction of micro-colony formation on the Pm3 NILs as compared to the control (Fig. 6e), further substantiating that the Pm3b,c,d alleles are acting as determinants of host specificity for non-adapted mildews. Finally, we also cloned and assayed the suppressor activity of the SvrPm3^3^ - C homologues in the Bgs_1391 and Bgs_1459 genomes. We found these isolates encode for the previously described SvrPm3^3^ - C haplotype, which differs by only one residue from the B. g. dactylidis sequence (Supplementary Fig. 25a). RNA-Seq data indicated the rye mildew encoded SvrPm3^3^ - C is expressed at very low levels upon infection of rye (Supplementary Fig. 25b). Using transient assays in *N. benthamiana*, we could show that this variant is indeed an inactive suppressor comparable to the previously described SvrPm3^3^ - C from the Pm3 avirulent wheat mildew isolate Bgd_96224 (Supplementary Fig. 25c).

We conclude that, (i) specific recognition of divergent AvrPm3 homologues from rye and *Dactylis* mildews by the Pm3b,c,d alleles, (ii) increased non-host resistance to rye-mildew isolates encoding active AvrPm3 homologues in presence of the Pm3 alleles, and (iii) presence of an inactive SvrPm3^3^ - C suppressor variant in those same isolates, is strong evidence supporting the role of the Pm3 alleles as determinants of host-specificity in cereal mildews based on specific and conserved AvrPm3-Pm3 interactions.

**Discussion**

Allelic series of resistance genes are highly informative to molecularly characterise resistance function if they are studied both on the host and the pathogen side. Genetic work in the 1990s and more recent molecular analysis has suggested that allelic series of resistance genes with distinct specificities are very promising systems to dissect the molecular basis of complex, race-specific interactions of biotrophic fungal pathogens such as rusts or
mildews with their hosts. The analysis of recognition specificity in the cereal powdery mildew and the flax rust systems based on allelic series of resistance genes has revealed additional layers of complexity to the original gene-for-gene interaction model postulated by Flor, because of the presence of suppressor/inhibitor genes and other modifiers of the interaction. In addition to the Pm3 allelic series of 17 race-specific resistance alleles to wheat mildew, there are three other well studied examples of true allelic series of functionally different resistance specificities: the downy mildew (Hyaloperonospora arabidopsidis) gene RPP13 in
**Arabidopsis thaliana**\(^{39}\), the flax rust (**Melampsora lini**) resistance gene \(^{1,24}\), and the **Mla** genes in barley (**Hordeum vulgare**)\(^{27}\). The **Pm3** alleles stand out with a particularly high level of sequence conservation (\(>97\%\) identity at the protein level)\(^{23,40}\) while the **Mla** alleles have 84.6\% sequence identity and are as divergent as the **Arabidopsis** RPP13 and the flax **L** allelic variants\(^{22,27,39}\).

Nears several allelic variants of RPP13 and the flax rust NLRs **L5**, **L6** and **L7** detect allelic AVR proteins, there is evidence that additional members of these allelic series also recognise sequence unrelated effectors\(^{33,41,42}\). This principle is further exemplified by **MLA1** and **MLA13**, which show >91\% identity on the protein level but recognize the sequence unrelated **Bgl** proteins **AVRRA1** and **AVRRA13**\(^{31}\). Consistent with these findings, the so far identified AVR proteins of the **Pm3** allelic series **AVRPm3A2/F2**, **AVRPm3B2/C2** and **AVRPm3D3** are members of three distinct **B. graminis** effector gene families with low sequence identity, and belong to large groups of phylogenetically related effectors with a possible common origin (in particular **AVRPm3A2/F2** and **AVRPm3B2/C2** (Supplementary Figs. 11 and 13). Together, phylogenetic relatedness, positional conservation of amino-acid with similar properties and structural residues across **AVRPm3** families (Supplementary Fig. 12), and possible similarities of tertiary protein folds (Supplementary Fig. 14), suggests that specificity of the **Pm3** alleles is possibly based on recognition of structurally similar effectors. This is reminiscent of the identification of diverse, but structurally conserved MAX effectors in different phytopathogenic fungi\(^{43}\), and would suggest that additional members of these allelic series also recognise sequence unrelated effectors.

Recent studies of effector functions in barley powdery mildew have established the role of effectors as important virulence factors (reviewed in ref. \(^{14}\)) that have the ability to interfere with components of the host basal metabolism and host immunity\(^{16–19}\). High frequency of active **AvrPm3** haplotypes among wheat mildew populations, and the fact that homologues can also be found in the related **formae speciales** **B. secalis** and **B. dactyliidis**, suggest that all **AvrPm3** genes are important for mildew virulence. Despite strong indication of diversifying selection on **AvrPm3** and as previously shown on **AvrPm3**\(^{32,29}\), no presence–absence polymorphisms or transposable element insertions could be detected for **AvrPm3**\(^{32,29}\) and **AvrPm3**\(^{32,29}\) (this study) in 272 and 185 **B. graminis** isolates respectively. Even in the case of **AvrPm3**\(^{33}\) with the avirulent version being absent in numerous isolates, no cases of complete loss of the effector could be identified, thus indicating that the **AvrPm3** effectors are always present together in all wheat powdery mildew isolates globally. This stand in strong contrast to **AvrPm2**, **AvrPm5** and **AvrPih** from **B. graminis**, **B. hordei** and **Magnaporthe oryzae**, respectively, where loss of avirulence was exclusively (**AvrPm2**), predominantly (**AvrPih**) or partially (**AvrPm5**) found to be dependent on presence/absence polymorphisms, gene disruptions by transposable element insertions, and segmental deletions\(^{30,31,44}\). These findings suggest there is a selective advantage for the pathogen to simultaneously encode all three effectors. We suggest that the inability to completely lose these effector genes is compensated partially by the identified diversifying selection, giving rise to new, unrecognised variants, and in parts by the presence and activity of **SrVpm3**\(^{41,42}\) that allows the wheat powdery mildew fungus to mask, and preserve effectors that would otherwise be recognised by the **Pm3** alleles. Previous studies in a few mildew isolates indicated that suppression of **AvrPm3** recognition is only effective when the active **SrVpm3**\(^{41,42}\)gene is highly expressed\(^{28,29}\). In this study, we did not determine the presence and expression levels of **SrVpm3**\(^{41,42}\) in the GWAS population. We propose that future integration of transcriptome data in genetic association and population genomics studies will provide important information on the role of **SrVpm3**\(^{41,42}\) in controlling both race-specificity and host-specificity in global cereal mildew populations. For **AVRPm3B2/C2**, domain swap studies identified two regions with higher levels of polymorphism in the mildew population that are sufficient for conferring AVR function, and two additional segments quantitatively affecting the strength of AVR recognition. McNally and colleagues\(^{29}\) also found that specific domains of **AVRPm3A2/F2** are involved in immune receptor recognition, including one region quantitatively affecting the strength of AVR recognition. We suggest these domains are possibly acting as contact regions affecting AVR-R recognition and binding affinity, which is reminiscent of the identification of 4 residues, spanning a 21 amino-acid region in the rice blast avirulence protein **AVR-Pik** and affecting recognition and binding affinity to rice Pik resistance protein\(^{15}\).

The evolutionary history of grass powdery Mildews has been reconstructed recently. Menardo and colleagues\(^{13}\) found that the radiation of the **B. graminis**-clade, consisting of **B. graminis**, **B. secalis** and **B. dactyliidis** has occurred 170,000 to 280,000 years ago and involved a host jump from wheat to Dactylis glomerata, a wild grass species, giving rise to **B. dactyliidis**. The recent radiation of the **tritici**-clade is also reflected in their respective Avr-effector complement. Of the four Avr genes that have been identified in **B. graminis**, homologues can also be found in either **B. secalis** (**AvrPm2**), **B. dactyliidis** (**AvrPm3**\(^{32,29}\)) or both closely...
related *formae specialae* (AvrPm3b2/c2, AvrPm3d3). Furthermore, all *Avr* homologues from *B. graminis* and the *AvrPm3d3* homologue from *B. dictyophila* are recognised by the corresponding *Pm3* alleles from wheat [28] (this study). The recognised variants of *AvrPm3b2/c2* from *B. graminis* and *AvrPm3d3* from *B. secalis* and *B. dictyophila* contain 4, 7 and 6 non-synonymous SNPs in the mature peptide, respectively, when compared to the *Pm3*-recognised *B. tritici* variants. This stands in strong contrast to the rather limited *AvrPm3* variation found within wheat mildew populations, where single non-synonymous mutations often result in evasion of recognition. We hypothesise that this reflects the fact that the currently active *Pm3* alleles have evolved only after wheat domestication ~11,000 years ago [28], and the *AvrPm3* homologues in closely related *formae specialae* have therefore not been targets of diversifying selection due to presence of the *Pm3* alleles in corresponding rye or *Dactylis* host species.

The presence of the *Pm3* alleles does severely restrict, often completely abolish the formation of non-sporulating micro-colonies by the non-adapted rye mildew (Fig. 6d, e). This is consistent with a previous study using a genetic cross between wheat and rye powdery mildews that has predicted over 20 years ago the presence of *AvrPm2, AvrPm3* and *AvrPm3* in the *B. secalis* genome, and suggested their role as host-specify determinant [46]. The ability of non-host mildews to establish micro-colonies on susceptible wheat cultivars might provide a basis for mating with adapted mildew isolates, thus giving raise to new hybrid forms with a broader host range such as Triticale powdery mildew [47]. Recent work by Praz and colleagues provided transcriptional evidence that the ability of Triticale mildew to grow on wheat and its new host triticale is based on downregulation (i.e. transcriptional removal) of numerous effectors mainly encoded within AVR gene families [38]. Therefore, loss of *Avr* function in grass mildew, combined with the deployment of susceptible wheat genotype (i.e. lacking *Pm* resistance genes), might provide a genetic basis for the raise of new adapted mildew forms. This hypothesis is substantiated by recent work by Inoue and colleagues [48], showing that the emergence of the wheat blast pathogen is due to widespread deployment of *rwt3* wheat (i.e. lacking the *Rwt3* resistance genes), combined with the loss of function of *PWT3* (the *Avr* of *Rwt3*) in non-adapted forms of *Pyricularia oryzae* (syn. *Magnaporthe oryzae*), resulting in a host jump on wheat [48]. These findings suggest that major *R* genes such as *Pm2* and *Pm3* must be maintained in the breeding gene pool even if they are locally defeated by virulent races from the adapted wheat mildew, because they are conferring important non-host resistance preventing the emergence of new mildew pathogens.

Based on these findings, we propose that the ability of other *formae specialae* of the *tritici*-clade to grow on wheat is restricted by the presence of many *Avr*-R gene pairs. These are simultaneously involved in race-specific resistance and in host specificity, similar to more generalised models proposed by (i) Tosa [49] who argued that the acquisition or loss of *Avr* genes is the main driving force behind the evolution of new *formae specialae*, a hypothesis recently substantiated by findings from wheat and rice blast [48,50], and (ii) Schulze-Lefert and Panstruga (2011) [51,52] suggesting that the contribution of major *R* genes to non-host resistance inversely correlates with the phylogenetic divergence time between the analysed host and non-host plants of a certain pathogen species or *formae specialae*.

**Methods**

**Fungal isolates and virulence tests.** *Blumeria graminis* f. sp. *tritici* isolates were maintained in the asexual phase on detached leaves of the susceptible bread wheat (*Triticum aestivum*) cultivar Kanzler on 0.5% food grade agar (PanReac AppliChem) plates supplemented with 4.23 mM benzimidazole [53]. Virulence tests were performed on near-isogenic lines or varieties using ‘Asosan/8’ (*Chancellor’ for *Pm3b, ‘Chul/8xChancellor’ for *Pm3d, ‘Sonora/8xChancellor’ for *Pm3e, ‘Kolibri’ for *Pm4d, ‘W130’ for *Pm5e, ‘Michigan Amber/8xChancellor’ for *Pm5f, and the *Fusarium* wheat cultivars ‘Kanzer’ and ‘Chancellor’ as a control [25]). Resistance scoring was performed 10 days after infection on at least three independent leaf segments. Scoring was performed qualitatively as follows: mildew leaf coverage (LC) of 60–100%, virulent; LC of 10–40%, intermediate; LC of 0%, avirulent [65].

For the GWAS, a population of natural isolates was established by collecting multiple isolates from infected wheat plants in the fields of six epidemiological regions in China from 2011 to 2014 described by Zeng and collaborators [44]. Single-pestid-derived isolates were purified by growing the pathogen on the susceptible wheat line ‘Chancellor’, then multiplied and stored at ~80 °C by drying spores at 23–5 °C in the presence of silica gel before freezing [31]. One hundred pure isolates with diversified geographic origin and a balanced virulence/avirulence pattern on wheat differential resistance lines were manually selected from the natural population (Supplementary Data 2).

**Blumeria graminis f. sp. secalis isolates** S-1391 and S-1459 were maintained in the asexual phase on detached leaves of the susceptible rye (*Secale cereale*) cultivar Walacite on 0.5% food grade agar (PanReac AppliChem) plates supplemented with 4.23 mM benzimidazole [45]. To assess the influence of the *Pm3b, Pm3c* and *Pm3d* alleles on early stages of infection on wheat, detached primary leaves of near-isogenic wheat lines ‘Chul/8xChancellor’ for *Pm3b*, ‘Sonora/8xChancellor’ for *Pm3c* and ‘Triticale’ for *Pm3d* were inoculated. Microscopic wheat leaf segments *Pm3b, Pm3d* and the corresponding susceptible control ‘Chancellor’ and ‘Bobwhite’ were infected with S-1391 and S-1459 as described above for *B. tritici*. Leaf segments were stained 48 hpi for reactive oxygen species in 1 mg mL−1 3’3’-dianisomindobenzaldehyde (DAB)-HCl, pH3.8 solution for 12 h [66] followed by complete de-staining of leaf pigments in ethanol:acetic acid solution (ratio 3:1) for several days. To detect fungal spores and hyphae, destained leaf segments were subsequently stained with Brilliant Blue (0.15% in EIOH absolute), washed repeatedly in H2O and samples mounted for microscopy in 50% glycerol. Using a conventional bright-field microscope, spores were scored based on the following categories: (i) microcolony formation: establishment of a haustorium and production of secondary hyphae without apparent signs of hypersensitive cell-death (HR) indicating a compatible interaction in early stages of infection, (ii) penetration of epidermal cells resulting in HR and stop of growth, or early arrest of spore growth in the absence of HR. For each isolate/cultivar combination the average of eight biological replicates (~independent leaf segments) is shown. For each leaf segment at least 30 spores were assessed. The assay was repeated with similar results. Statistical significance of observed differences was tested using a Wilcoxon rank sum test for possibly tied observation (wilcoxon.exact function, R package exactRankTests).

**Haplotyping mining.** Haplotyping diversity was assessed on a subset of the worldwide *B. graminis* f. sp. *tritici* collection previously described by McNally and colleagues [49]. For the isolates originating from the United States, Israel, and Europe the *AvrPm3b2/c2* and *AvrPm3d3* genes were PCR amplified from genomic DNA with primers listed in Supplementary Table 1, and Sanger-sequenced. All other isolates, with exception of isolates from China (Supplementary Data 1, and Sanger-sequenced. All other isolates, with exception of isolates from China (Supplementary Data 1), were mined for the same two gene pairs. These are simultaneously involved in race-specific resistance and in host specificity, similar to more generalised models proposed by (i) Tosa [49] who argued that the acquisition or loss of *Avr* genes is the main driving force behind the evolution of new *formae specialae*, a hypothesis recently substantiated by findings from wheat and rice blast [48,50], and (ii) Schulze-Lefert and Panstruga (2011) [51,52] suggesting that the contribution of major *R* genes to non-host resistance inversely correlates with the phylogenetic divergence time between the analysed host and non-host plants of a certain pathogen species or *formae specialae*.

**DNA/RNA isolation.** Fungal high molecular weight DNA was extracted using an optimised procedure described in detail by Bouras and colleagues [34,35]. In short, ~100 mg of conidia were frozen in liquid nitrogen and ground three times 30 s using stainless steel beads and a high-speed plate grinder (MM200 Mixer Mill, Retsch, Germany) at a grinding frequency of 30 s. To the ground conidia, 300 μl of 65 °C preheated 5% Sarcosyl solution was added and vigorously vortexed. After addition of 700 μl of CTAB-Buffer (0.2 M Tris(hydroxymethyl)aminomethane at pH 7.5, 50 mM EDTA, 2 M NaCl, 2% Cetyl trimethylammonium bromide (CTAB) and 0.25 M Sodium metabisulphite (Na2S2O5)) the samples were incubated at 65 °C for 30 min. After incubation, 600 μl Chloroform was added, followed by 10 min of centrifugation at 14,000 rcf, 4 °C. The supernatant was collected and concentrated with one volume of ~20 °C precooled Isopropanol, mixed by careful inversion and centrifuged for 10 min, 14,000 rcf, 4 °C. The resulting pellet was dried on ice and resuspended in standard Tris/EDTA buffer before further purification on Amicor/Ultra 0.5 ml centrifugal filters MWCO 30 KDa (Sigma-Aldrich, Steinheim, Germany).

DNA/RNA samples were extracted from infected wheat leaves of the susceptible cultivar ‘Chinese Spring’ using the SV Total RNA Isolation System (Promega) according to the manufacturer’s protocol. RNA integrity was assessed by gel electrophoresis and spectrophotometric analysis using a NanoDrop 1000 (Thermo Scientific) [38]. RACE-PCR was performed with the SMARTer RACE CDNA Amplification Kit (Clontech Laboratories Inc.) according to the manufacturer and with gene-specific primers listed in Supplementary Data 4.
Codon optimisation, gene synthesis and plasmids. Codon optimisation of effector coding sequences for expression in N. benthamiana was performed using the online codon adaptation tool in Integrated DNA Technologies (https://www.idtdna.com/GenOpton). In all cases the signal peptide, as predicted by the SignalP4 algorithm64 (http://www.cbs.dtu.dk/services/SignalP4/) was replaced by an ATG start codon. Gene synthesis including gateway-compatible attL cloning sites was performed by colleagues34 using Salmon with standard parameters66. BAC clones were sequenced using Illumina MiSeq technology (2 × 250 bp paired-end). Fungal isolates were grown in AS-medium (10 mM MES-KOH, pH 5.6; 10 mM MgCl2; 200 µM acetosyringone) to an OD of 1.2 before incubation for 3 to 4 h at 28°C with 200 rpm shaking to induce HR assessment was chosen depending on the experimental setup: 5 days with Agrobacterium infiltration for verification of AVR recognition by the corresponding R-gene, 2dpi for suppression assays involving AvrPm3 and to test for interallelic suppression. HR was quantified using high-resolution epifluorescence images obtained with the Fusion FX HRS technology, and analysed with the Image) image analysis software measuring the integrated density of fluorescence59,60. HR values were normalised to the infiltrated leaf area and to the control. Statistical significance was assessed with the Student’s t-test for paired data (p < 0.05).

Protein detection. The soluble protein fraction was extracted from A. tumefaciens infiltrated N. benthamiana leaves 2 days after infiltration. Two leaf discs (0.5 mm diameter) from 4 infiltrated leaves were pooled, ground in liquid nitrogen and resuspended in 2x modified Laemmli-Buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 2% SDS, 0.2% ammonium thiocyanate). Protein concentration was determined using bicinchoninic acid (Pierce, ThermoFisher). SDS-PAGE gels were run on a 10% gel, stained with Coomassie-blue and scanned. Predicted protein masses were calculated using Compute pI/Mw (ExPASy). Western Blotting was performed using antibody conjugated to HRP. Blotted protein was stained with Ponceau S. For HA detection, a peroxidase conjugated antibody (anti-HA-HRP, rat monoclonal, clone 3F10, Roche) was used. For HA detection, a peroxidase conjugated antibody (anti-HA-HRP, rat monoclonal, clone 3F10, Roche) was used at a dilution of 1:5000. For FLAG detection the primary antibody (anti-FLAG, mouse monoclonal, clone M2, Sigma-Aldrich) was used at a dilution of 1:10000, membranes washed with TBST and incubated with anti-mouse peroxidase anti-body (mouse-anti-HRP, goat polyclonal, Sigma-Aldrich) diluted 1:4000. Peroxidase chemiluminescence was detected using a Fusion FX Imaging System (Vilber Lourmat, Eberhardzell, Germany) and WesternBright ECL HRP substrate (Advansta).

Phylogenetic analysis and structural modelling. Protein sequences of the AVRPM3 alleles were retrieved from GSE108405 and BAC insert sizes were modelled. The resulting predicted folds were manually inspected and visualised in the Nature Research Reporting Summary linked to this article. Further information on research design is available in the Supplemental Table 2. Linearised SDM constructs including nucleotide sequence and manufacturer can be found in Supplementary Data 3. BAC sequencing. The B. t strain GV3101 using electroporation (1.44 kV, 25 μF, 2000V(300). Site-directed mutagenesis (SDM) was performed by PCR amplification on primers on overlapping primers (Supplementary Table 2). Linearised SDM constructs were phosphorylated using T4 polynucleotide kinase (New England Biolabs) and re-ligated with T4 DNA Ligase (New England Biolabs) according to the manufacturer. Epitope tagged Pm3 alleles have been published by Brunner and colleagues56. Epitope tagging of effector genes has been performed by gene-synthesis or SDM as described above (see also Supplementary Note 3). A complete list of all synthesised or mutagenised (SDM) constructs including nucleotide sequence and manufacturer can be found in Supplementary Data 3.

Germplasm and plants. Fungal isolates used for GWAS are described in Table 1. A. t. After infiltration into N. benthamiana leaves with 200 rpm shaking to induce HR. Candi- ular strains to the bam file (http://broadinstitute.github.io/picard). Bam files were merged using bartools merge command and indexed using samtools command index. SNP calling was done using freebayes49 with the following parameters: -p 1 -m 30 -q 20 -z 0.7–3 200-genome Fre- 0.05–max-meanDP 40–remove-indels–min-alles–2–min-alles–2–min-QQ 0–max-missing 0.9. VCF file was transformed into hapmap format using a custom perl script. GWAS was performed using the GAPS package for R33 using the the PCA.total model. The resulting predicted folds were manually inspected and visualised in the Quick2D modelling employs four different algorithms for prediction, namely a PSI-blast based secondary structure prediction algorithm (PSIPRED), an iterative deep neural network (DNN) based algorithm (SPIDER2), a multiple backpropagation neural network predictor (PSSPRED), and a deep convolutional neural fields based method (DEEPNF). Secondary folds depictions were extracted and adapted from the Quick2D output. Three dimensional protein modelling was performed using the RaptorX protein modelling tool (http://raptorx.uchicago.edu). Protein sequences corresponding to the predicted mature protein after cleavage of the predicted signal peptide were modelled. The resulting predicted folds were manually inspected and visualised in order to identify possible common structures. For each predicted fold, the P-value, the Global Distance Test (GDT) score, and seqEDT identiy score were recorded (Supplemental Data 5). Detailed definitions of these scores are available in the RaptorX user documentation (http://raptorx.uchicago.edu/documentation/).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article. Data availability. Gene expression data was obtained from Gene Expression Omnibus (GEO) under accession number GSE108405 [https://www.ncbi.nlm.nih.gov/geo/?term=GSE108405]. Genomic sequences of the fungal isolates used for GWAS are available under the accession number SRP062198 [https://www.ncbi.nlm.nih.gov/sra/?term=SRP062198]. Further information on research design is available in the Supplemental Table 2. Linearised SDM constructs including nucleotide sequence and manufacturer can be found in Supplementary Data 3. BAC sequencing. The B. t.
Received: 3 May 2018 Accepted: 3 May 2019
Published online: 23 May 2019

References

1. Farrer, R. A. et al. Multiple emergences of genetically diverse amphibian-infecting chytrids include a globalized hypervirulent recombinant lineage. Proc. Natl Acad. Sci. USA 108, 18732–18736 (2011).

2. Blehert, D. S. et al. Bat white-nose syndrome: an emerging fungal pathogen? Science 314, 552–553 (2004).

3. Kim, K. & Harvell, C. D. The rise and fall of a six-year coral-fungal epizootic. Science 333, 1326–1328 (2011).

4. vanEngelsdorp, D. et al. Colony collapse disorder: a descriptive study. PLoS One 4, e6481 (2009).

5. Jessup, D. A. et al. Southern Sea Otter as a Sentinel of Marine Ecosystem Evolutionary Time Scales with NGS data. Mol. Ecol. 24, 304–315 (2015).

6. Cruz, C. D. & Valient, B. Wheat blast disease: danger on the move. Trop. Plant Pathol. 32, 210–222 (2017).

7. Singh, R. P. et al. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. New Phytol. 213, 453–461 (2017).

8. Saintenac, C. et al. Identification of wheat gene Sr35 that confers resistance to Ugr99 stem rust race group. Science 341, 783–786 (2013).

9. Salcedo, A. et al. Variation in the AvrSr35 gene determines Sr35 resistance against wheat stem rust race Ug99. Science 358, 1604–1606 (2017).

10. Bourras, S., McNally, K. E., Müller, M. C., Wicker, T. & Keller, B. Avirulence genes in cereal powdery mildews: the gene-for-gene hypothesis 2.0. Front. Plant Sci. 7, 241 (2016).

11. Zhou, L.X. et al. The evolutionary history of powdery mildew lineages (Blumeria graminis) at Different Evolutionary Time Scales with NGS data. Genome Biol. Evol. 9, 446–456 (2017).

12. Troch, V. et al. Formae speciales of cereal powdery mildew: Close or distant relatives? Mol. Plant. Pathol. 15, 304–314 (2014).

13. Menardo, F. et al. The barley powdery mildew effector CSEP0081 and CSEP0254 promote fungal infection success. Plasmodiophora brassicae. Fungal Genome Biol. 69, 497–507 (2010).

14. Ahmed, A. A. et al. The barley powdery mildew effector candidate CSEP0081 and CSEP0254 promote fungal infection success. Plasmodiophora brassicae. Fungal Genome Biol. 69, 497–507 (2010).

15. Ahmed, A. et al. The barley powdery mildew effector candidate CSEP0055 with the defence protein PR17c. Mol. Plant Pathol. 13, 435–443 (2012).

16. Pennington, H. G. et al. Interactions between the Powdery Mildew Effector Bc1054 and Barley Proteins Identify Candidate Host Targets. J. Proteome Res. 15, 826–839 (2016).

17. Pliego, C. et al. Host-induced gene silencing in barley powdery mildew reveals a class of ribonuclease-like effectors. Mol. Plant Microbe Interact. 26, 633–642 (2013).

18. Whitham, E. et al. Broadly conserved fungal effector BEC1019 suppresses host cell death and enhances pathogen virulence in powdery mildew of barley (Hordeum vulgare L.). Mol. Plant-Microbe Interact. 28, 968–983 (2015).

19. Ravensdale, M., Netter, A., Thrall, P. H., Ellis, J. G. & Dodds, P. N. Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. Mol. Plant Pathol. 12, 93–102 (2011).

20. Dodds, P. N. et al. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proc. Natl Acad. Sci. USA 103, 8888–8893 (2006).

21. Yahiaoui, N., Brunnier, S. & Keller, B. Rapid generation of new powdery mildew resistance genes after wheat domestication. Plant J. 47, 85–98 (2006).

22. Bhullar, N. K., Mackay, M. & Keller, B. Genetic diversity of the Pm3 powdery mildew resistance alleles in wheat gene bank accessions as assessed by molecular markers. Diversity 2, 768–786 (2010).

23. Stinner, D. et al. Suppression among alleles encoding nucleotide-binding-leucine-rich repeat resistance proteins interferes with resistance in F1hybrid and alele-pyramided wheat plants. Plant J. 79, 893–903 (2014).

24. Seeholzer, S. et al. Diversity at the Mla powdery mildew resistance locus from cultivated barley reveals sites of positive selection. Mol. Plant-Microbe Interact. 23, 497–509 (2010).

25. Bourras, S. et al. Multiple avirulence loci and allele-specific effector recognition control the Pm3 race-specific resistance of wheat to powdery mildew. Plant Cell 27, 2991–3012 (2015).

26. McNally, K. E. et al. Distinct domains of the AVRPM3A2/F2 avirulence protein from wheat powdery mildew are involved in immune receptor recognition and putative effector function. New Phytol. 218, 681–695 (2017).

27. Praz, C. R. et al. AvrPm2 encodes an RNase-like avirulence effector which is conserved in the two different specialized forms of wheat and rye powdery mildew fungus. New Phytol. 213, 1301–1314 (2017).

28. Lu, X. et al. Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. Proc. Natl Acad. Sci. USA 113, E6486–E6495 (2016).

29. Parfange, F. et al. Genetic and molecular characterization of a locus involved in avirulence of Blumeria graminis f. sp. tritici on wheat Pm3 resistance alleles. Fungal Genet. Biol. 82, 181–192 (2015).

30. Lipka, A. E. et al. GAPIT: genome association and prediction integrated tool. Bioinformatics 28, 2397–2399 (2012).

31. Müller, M. C. et al. A chromosome-scale genome assembly reveals a highly dynamic effector repertoire of wheat powdery mildew. New Phytol. 221, 2176–2189 (2019).

32. Brunnier, S. et al. Intrageneric allele pyramiding combines different specificities of wheat Pm3 resistance alleles. Plant J. 64, 435–445 (2010).

33. Zimmermann, L. et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol. 430, 2237–2243 (2018).

34. Källberg, M. et al. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. 7, 1511–1522 (2012).

35. Praz, C. R. et al. Non-parent of origin expression of numerous effector genes indicates a role of gene regulation in host adaption of the hybrid triticale powdery mildew pathogen. Front. Plant Sci. 9, 49 (2018).

36. Rose, L. E. et al. The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in Arabidopsis thaliana. Genetics 166, 1517–1527 (2004).

37. Bhullar, N. K., Street, K., Mackay, M., Yahiaoui, N. & Keller, B. Unlocking wheat genetic resources for the molecular identification of previously undescribed fungal alleles at the Pm3 resistance locus. Proc. Natl Acad. Sci. USA 106, 9519–9524 (2009).

38. Hall, S. A. et al. Maintenance of genetic variation in plants and pathogens involves complex networks of gene-for-gene interactions. Mol. Plant Pathol. 9, 449–457 (2009).

39. Dodds, P. N. The melampsora lini Avrl567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. Plant Cell 16, 755–768 (2004).

40. de Guillon, K. et al. Structure analysis uncovers a highly diverse but structurally conserved effector family in phytopathogenic fungi. PLoS Pathog. 11, e1005228 (2015).

41. Zhang, S. et al. Function and evolution of Magnaporthe oryzae avirulence gene AvrPb responding to the rice blast resistance gene Pib. Sci. Rep. 5, 11642 (2015).

42. Dhoundial, K. et al. A structural basis for pattern recognition by an integrated HMA domain in a plant NLR immune receptor. Elife https://doi.org/10.7554/ Elife.08709 (2015).

43. Matsumura, K. & Tosa, Y. The rye mildew fungus carries avirulence genes corresponding to wheat genes for resistance to races of the wheat mildew fungus. Phytopathology 85, 753 (1995).

44. Menardo, F. et al. Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. Nat. Genet. 48, 201–205 (2016).

45. Inoue, Y. et al. Evolution of the wheat blast fungus through functional losses in a host specificity determinant. Science 357, 80–83 (2017).

46. Tosa, Y. A model for the evolution of formae speciales and races. Phytopathology 82, 728–730 (1992).

47. Liao, J. et al. Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. Elife 5, e19377 (2016).
51. Schulze-Lefert, P. & Panstruga, R. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in Plant Science* **16**, 117–125 (2011). https://doi.org/10.1016/j.tplants.2011.01.001

52. Maekawa, T., Kracher, B., Vernardi, S., Ver Loren van Themaat, E. & Schulze-Lefert, P. Conservation of NLR-triggered immunity across plant lineages. *Proc. Natl Acad. Sci. USA* **109**, 20119–20123 (2012).

53. Parlange, F. et al. A major invasion of transposable elements accounts for the large size of the Blumeria graminis Lsp. tritici genome. *Func. Integr. Genomics* **11**, 671–677 (2011).

54. Zeng, F. S. et al. Virulence and diversity of Blumeria graminis f. sp. tritici populations in China. *J. Integr. Agric.* **13**, 2424–2437 (2014).

55. Gong, S. et al. An approach for long-term preservation of Blumeria graminis f. sp. tritici. *Trop. Plant Pathol.* **40**, 127–133 (2015).

56. Brunner, S. et al. Transgenic Pm3b wheat lines show resistance to powdery mildew in the field. *Plant Biotechnol. J.* **9**, 897–910 (2011).

57. Brunner, S. et al. Transgenic Pm3 multilines of wheat show increased powdery mildew resistance in the field. *Plant Biotechnol. J.* **10**, 398–409 (2012).

58. Thordal-Christensen, H., Zhang, Z., Wei, Y. & Collinge, D. B. Subcellular localization of H2O2 in plants. H2O2 accumulation in papillae and hypersensitive response during the barley-powdery mildew interaction. *Plant J.* **11**, 1187–1194 (1997).

59. Nei, M. & Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–426 (1986).

60. Rozas, J. et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299–3302 (2017).

61. Petersen, T. N., Brunak, S., Von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).

62. Himmelbach, A. et al. A set of modular binary vectors for transformation of cereals. *Plant Physiol.* **145**, 1192–1200 (2007).

63. Frenkel, Z., Paux, E., Mester, D., Feuillet, C. & Korol, A. LTC: a novel algorithm to improve the efficiency of contig assembly for physical mapping in complex genomes. *BMC Bioinform.* **11**, 584 (2010).

64. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

65. Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

66. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

67. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

68. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

69. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at https://arxiv.org/abs/1207.3907 (2012).

Acknowledgements

We would like to acknowledge Dr. Kohji Nishimura (Shimane University, Japan) for technical support with protein expression assays. We would like to acknowledge the technical support from Gerhard Herren, Helen Zbinden, Karl Huwiler, Gabriele Buchmann and Ritta Rabbat. We would like to acknowledge the financial support of the Swiss National Science Foundation (SNSF) grants 310030_163260 and 310036B_182833, the European Research Council (ERC), and the University Research Priority Program (URPP) of the University of Zurich.

Author contributions

S.B., L.K., and B.K. wrote the paper. S.B., D.Y. and B.K. coordinated the research. S.B., L.K., M.X., D.Y. and B.K. designed the research. S.B., L.K., C.K., M.S., P.A., S.F., F.P., I.K. S., V.W., S.L. and J.I. performed lab experiments. S.B., L.K., M.X., C.R.P., M.C.M., F.P. and R.B.D. performed genetic analyses. C.R.P., M.C.M., F.M., S.R., S.O. and T.W. performed bioinformatics analyses. S.B., C.R.P. and F.M. developed the effector benchmarking approach. M.X., C.R.P., M.C.M. and F.M. developed the GWAS approach. S.B., L.K., C.K., M.S. and S.F. developed and executed the AVR screens in *N. benthamiana*.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-10274-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: *Nature Communications* thanks Yuko Tosa and other anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.