

## 

**Citation:** Kunz L, Jigisha J, Menardo F, Sotiropoulos AG, Zbinden H, Zou S, et al. (2025) Avirulence depletion assay: Combining *R* genemediated selection with bulk sequencing for rapid avirulence gene identification in wheat powdery mildew. PLoS Pathog 21(1): e1012799. https://doi. org/10.1371/journal.ppat.1012799

Editor: Carl Mesarich, Massey University College of Sciences, NEW ZEALAND

Received: July 17, 2024

Accepted: December 4, 2024

Published: January 7, 2025

**Copyright:** © 2025 Kunz et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: PacBio HiFi reads for isolate CHVD\_042201 are available from the sequence read archive (SRA, accession: PRJNA1131794). The genome assembly of isolate CHVD\_042201 is available at the European nucleotide archive (ENA) under accession numberGCA\_964289775. Genome assembly and annotation of isolate CHVD\_042201 are also available from Zenodo: https://zenodo.org/records/ 11233413. Illumina data for the two parental isolates used for the AD assay are available as **RESEARCH ARTICLE** 

# Avirulence depletion assay: Combining *R* gene-mediated selection with bulk sequencing for rapid avirulence gene identification in wheat powdery mildew

Lukas Kunz<sup>1</sup>, Jigisha Jigisha<sup>1</sup>, Fabrizio Menardo<sup>1</sup>, Alexandros G. Sotiropoulos<sup>1,2</sup>, Helen Zbinden<sup>1</sup>, Shenghao Zou<sup>3</sup>, Dingzhong Tang<sup>3</sup>, Ralph Hückelhoven<sup>4</sup>, Beat Keller<sup>1</sup>, Marion C. Müller<sup>0</sup><sup>4</sup>\*

 Department of Plant and Microbial Biology, University of Zurich, Zurich, Switzerland, 2 Centre for Crop Health, University of Southern Queensland, Toowoomba, Queensland, Australia, 3 State Key Laboratory of Ecological Control of Fujian-Taiwan Crop Pests, Fujian Agriculture and Forestry University, Fuzhou, China,
Chair of Phytopathology, TUM School of Life Sciences, Technical University of Munich, Freising, Germany

\* marion.mueller@tum.de

### Abstract

Wheat production is threatened by multiple fungal pathogens, such as the wheat powdery mildew fungus (Blumeria graminis f. sp. tritici, Bgt). Wheat resistance breeding frequently relies on the use of resistance (R) genes that encode diverse immune receptors which detect specific avirulence (AVR) effectors and subsequently induce an immune response. While R gene cloning has accelerated recently, AVR identification in many pathogens including Bgt lags behind, preventing pathogen-informed deployment of resistance sources. Here we describe a new "avirulence depletion (AD) assay" for rapid identification of AVR genes in Bgt. This assay relies on the selection of a segregating, haploid F1 progeny population on a resistant host, followed by bulk sequencing, thereby allowing rapid avirulence candidate gene identification with high mapping resolution. In a proof-of-concept experiment we mapped the AVR component of the wheat immune receptor Pm3a to a 25 kb genomic interval in Bgt harboring a single effector, the previously described AvrPm3<sup>a2/t2</sup>. Subsequently, we applied the AD assay to map the unknown AVR effector recognized by the Pm60 immune receptor. We show that AvrPm60 is encoded by three tandemly arrayed, nearly identical effector genes that trigger an immune response upon co-expression with Pm60 and its alleles *Pm60a* and *Pm60b*. We furthermore provide evidence that *Pm60* outperforms Pm60a and Pm60b through more efficient recognition of AvrPm60 effectors, suggesting it should be prioritized for wheat breeding. Finally, we show that virulence towards Pm60 is caused by simultaneous deletion of all AvrPm60 gene paralogs and that isolates lacking AvrPm60 are especially prevalent in the US thereby limiting the potential of Pm60 in this region. The AD assay is a powerful new tool for rapid and inexpensive AVR identification in Bgt with the potential to contribute to pathogen-informed breeding decisions for the use of novel R genes and regionally tailored gene deployment.

follows: CHVD\_042201 is available from ENA (accession: PRJEB75381) and CHN\_52\_27 from SRA (accession: SRR29761601). Illumina sequencing data from the bulks generated by the AD assay are available at SRA (accession: PRJNA1125842). Scripts used to analyze the bulk sequencing data generated by the AD assay are available from https://github.com/MarionCMueller/ AD-assay.

**Funding:** This work was supported by the University Research Priority Program (URPP) 'Evolution in Action' of the University of Zurich, the University of Zurich and by grant 310030\_204165 from the Swiss National Science Foundation to BK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

#### Author summary

Breeding for disease resistant cultivars that withstand pathogen infection represents an important strategy for environmentally sustainable wheat production. The underlying disease resistance genes however vary greatly in their efficacy and durability against fast evolving fungal pathogens such as wheat powdery mildew. Pathogens evade recognition by immune receptors through mutation or loss of recognized avirulence effector proteins, resulting in local or global breakdown of resistance genes. To improve the durability of resistance genes and allow for efficient regional deployment of resistance sources, pathogen-informed strategies that rely on avirulence effector identification, characterization and monitoring are therefore crucial.

Here, we describe a new approach for rapid identification of avirulence effectors in the wheat powdery mildew pathogen: the AD assay. We show that the AD assay reduces time and costs compared to previous approaches, while simultaneously allowing candidate gene identification with high precision. We deploy the AD assay to identify the avirulence effector *AvrPm60* recognized by the wheat immune receptor *Pm60* and characterize the prerequisites of *Pm60*-mediated resistance. Our study shows how rapid avirulence effector identification can contribute to a detailed understanding of the interplay between avirulence effectors and resistance genes and can thus guide future breeding decisions.

#### Introduction

Global wheat production is threatened by numerous pathogenic organisms with fungal diseases alone resulting in 15–20% yield losses annually [1,2]. Sustainable wheat production therefore relies on extensive breeding efforts to identify new genetic resistance sources, including specific resistance (R) genes, and to introduce them into high yielding cultivars. Many Rgenes encode intracellular nucleotide-binding leucine-rich repeat (NLR) immune receptors that recognize the presence of pathogen effectors, so called avirulence (AVR) proteins, and subsequently induce an immune response. AVR recognition by NLRs often results in a hypersensitive response (HR) that includes localized cell death and thereby limits pathogen proliferation [3,4]. Evasion of NLR recognition by mutation or loss of AVR genes is often observed in fast evolving plant pathogens and thereby significantly limits the durability of R genes deployed in agricultural settings [5]. Several recent studies have highlighted the need for AVRidentification and analysis of AVR diversity within pathogen populations in order to predict Rgene durability and effective deployment in agricultural settings [6–8].

While *R* gene identification and cloning in wheat and other staple crops has significantly sped up in recent years due to technological advances and vastly improved genomic resources [9], identification of *AVR* genes lags behind for many pathogens and new, more efficient methods for *AVR* gene identification are urgently needed [7].

Wheat powdery mildew (*Blumeria graminis* f.sp. *tritici*, abbrev. *Bgt*) is an obligate biotrophic ascomycete fungus that exhibits a high level of host specificity and exclusively infects wheat [10]. Due to its short asexual life cycle it can cause explosive epidemics among susceptible wheat monocultures and result in considerable yield losses [11]. Until today, 69 *R* genes with over 100 functional alleles against *Bgt* have been genetically defined in wheat (*Pm1* to *Pm69*) with many more *R* gene candidates awaiting official definition [6,12]. However, only a fraction of all defined *Pm* genes have so far been molecularly isolated and cloned with many of them (*Pm1a*, *Pm2*, *Pm3a-t*, *Pm5e*, *Pm8*, *Pm17*, *Pm21*, *Pm41*, *Pm60* and *Pm69*) encoding classic NLR proteins [13-21]. While some of the cloned *Pm* genes provide resistance against a narrow range of *Bgt* isolates or have been largely overcome by the pathogen population during agricultural deployment, other *Pm* genes such as *Pm60* provide resistance against most *Bgt* isolates and therefore represent valuable *R* gene candidates for future breeding efforts [20,22].

*Pm60* has been identified and cloned from the wheat progenitor *Triticum urartu* and was found to encode an NLR [20]. A recent study investigating the genetic diversity of the *Pm60* locus in *T. urartu* has furthermore revealed two additional *Pm60* alleles, *Pm60a* and *Pm60b*, that provide resistance against *Bgt* [22]. Interestingly, *Pm60a* is defined by a 240 bp deletion, encompassing two entire LRR repeats. By contrast, *Pm60b* carries a 240 bp insertion and therefore two additional LRR repeats when compared with *Pm60* [22]. Based on the resistance spectrum found in *Pm60*, *Pm60a* and *Pm60b* containing lines, it was hypothesized that the three alleles might recognize similar *AVR* effector targets albeit likely with differences in recognition strength [22]. However, the identity of the corresponding *AVR* effector *AvrPm60* from *Bgt* is so far unknown, which hampered detailed investigations of *Pm60* allele specificity and predictions of their potential success for wheat breeding.

In the last decade, multiple *AVR* genes have been identified and cloned from *Bgt*: *AvrPm1a\_1*, *AvrPm1a\_2*, *AvrPm2*, *AvrPm3<sup>a2/f2</sup>*, *AvrPm3<sup>b2/c2</sup>*, *AvrPm3<sup>d3</sup>*, *AvrPm8*, *AvrPm17* were shown to be recognized by the NLRs Pm1a, Pm2, Pm3a/Pm3f, Pm3b/Pm3c, Pm3d, Pm8 and Pm17, respectively [8,13,23–27]. All *Bgt* AVR effector genes known to date encode small, secreted effector proteins with a size of approximately 120 amino acids and are predicted to exhibit an RNAse-like fold [28,29], thereby belonging to the RALPH (RNase-Like Proteins associated with Haustoria) effector superfamily, which constitutes more than half of the >800 effector genes in the *Bgt* genome [28–30].

The identification and functional validation of *AVR/R* gene pairs combined with analyses of *AVR* diversity within the worldwide *Bgt* population has proven crucial to advance the understanding of gain-of-virulence mechanisms in *Bgt* and consequently resistance gene breakdown in wheat agriculture. For instance, two recent studies on the quick breakdown of the *Pm8* and *Pm17* resistance genes, introgressed from rye, found evidence for ancient genetic variation in the corresponding *AVR* genes within the *Bgt* population, including virulence alleles that evade *R* gene recognition. Importantly, these ancient *AVR* gene variants precede the introgression of *Pm8* and *Pm17* from rye into the wheat gene pool, explaining their rapid breakdown shortly after agricultural deployment [8,27]. Such examples highlight the importance of parallel *AVR* and *R* gene identification in pathogen and host in order to predict durability of *R* genes and to allow prioritization of most promising gene candidates for wheat breeding or *R* gene stacking.

*AVR* gene identification in *Bgt* has relied on a variety of experimental strategies using genetic mapping, GWAS, effector screens and, most recently, UV mutagenesis to induce and identify gain-of-virulence mutations [24,25,31]. Genetic mapping approaches have proven particularly powerful for *AVR* gene identification in *Bgt* due to its haploid genome and a fast and simple life cycle including experimentally controllable asexual (i.e. clonal) and sexual reproduction. During the predominant asexual phase, the fungus infects wheat epidermal cells through the establishment of a haustorium for nutrient uptake and effector secretion and, within just a few days, produces large number of haploid conidiospores to repeat the infection cycle. Sexual reproduction occurs exclusively upon contact between *Bgt* isolates of opposite mating types (*MAT*) and results in the formation of sexual fruiting bodies, chasmothecia, that eventually release haploid ascospores to reinitiate the asexual phase. The experimental accessibility of both asexual and sexual reproduction systems paired with the haploid nature of the *Bgt* genome allowing phenotypic assessment of sexual progenies already in the F1 generation has made genetic mapping approaches the method of choice in many projects aiming at *AVR* 

gene identification. However, genetic mapping using biparental crosses traditionally involves time-consuming isolation of 100+ individual F1 progeny from sexually formed chasmothecia, their subsequent asexual propagation, genotyping and phenotyping thereby resulting in work-and cost-intensive projects with timelines of 1-2 years. Hence, there is a need for technological advances to speed up *AVR* identification in this important wheat pathogen.

In this study, we describe a new "AVR depletion assay" (AD assay) for rapid and low-cost AVR gene identification in Bgt. The assay preserves the many advantages of genetic mapping approaches but circumvents the time-consuming isolation and propagation of individual F1 progenies by combining the generation of sexual recombinant F1 populations with R gene-mediated selection and bulk sequencing. In a proof-of-concept experiment, we show that the AD assay can identify the previously described  $AvrPm3^{a2/f2}$  with high precision (i.e. identifying a single candidate effector). Furthermore, we apply the new method to identify and functionally validate the previously unknown AvrPm60, recognized by the broadly acting Pm60 resistance gene.

#### Results

# The avirulence depletion assay allows mapping of $AvrPm3^{a2/f2}$ with high resolution

In this study, we aimed at developing an "avirulence depletion assay" (AD assay) as a new approach for the identification of avirulence effectors in *Bgt*.

The AD assay is based on principles of bi-parental mapping but without the need to establish, maintain, and genotype individuals of a mapping population (Fig 1). As such, the approach relies on crossing two parental *Bgt* isolates displaying opposite virulence phenotypes on an *R* gene-containing wheat line and subsequent regeneration of a haploid sexual F1 mapping population. In contrast to classical bi-parental mapping approaches (map-based cloning, QTL mapping), in the AD pipeline, haploid conidiospores of a mixed progeny population F1-1 (i.e. the first asexually produced conidiospores of F1 cross progenies) are directly used to infect a wheat line that carries the resistance of interest thereby creating a strong and



**Fig 1. Schematic summary of the avirulence gene** (*AVR*) **depletion assay** (**AD assay**) **workflow**. A bi-parental cross between *Bgt* isolates exhibiting opposite (a)virulence phenotypes results in a haploid F1 population which subsequently is selected on a resistant wheat line resulting in *AVR* depletion (selected bulk) or a susceptible control line (non-selected bulk). Bulk NGS sequencing and bioinformatic analyses are used to define genomic regions with *AVR* depletion and subsequently identify *AVR* candidate genes. Fig created with BioRender.com.

https://doi.org/10.1371/journal.ppat.1012799.g001

directional selection pressure that depletes the F1-1 progeny population from individuals carrying the corresponding *AVR* factors. In parallel, the initial F1-1 progeny population is used to infect a susceptible wheat genotype subsequently serving as an unselected control. Following bulk harvesting and sequencing of the surviving F1-2 progenies, single nucleotide polymorphisms (SNPs) are used to assess parental genotype ratios throughout the haploid genome. For the control bulk and for genomic regions that are unaffected by *R*-gene mediated selection, F1-2 progenies are expected to inherit equal proportions of parental genotypes (1:1 ratio). In contrast, *R*-gene mediated selection is expected to severely deplete the avirulent parental genotype in genomic regions harbouring *AVR* effectors. Thus, the deviations from the 1:1 ratio in the selected bulks can be used to identify candidate *AVR* loci. In the last step of the AD assay, the identified candidate regions are inspected using genomics and transcriptomics datasets to identify candidate *AVR* genes (Fig 1).

In a first experiment, we aimed to establish a proof-of-concept dataset for the AD-assay by using the well-characterized *Bgt* avirulence gene  $AvrPm3^{a2/f2}$  located on Chr-06.  $AvrPm3^{a2/f2}$  encodes an RNAse-like effector recognized by the wheat NLR allele Pm3a [25]. To test if the AD-assay can identify  $AvrPm3^{a2/f2}$ , we crossed the Swiss isolate CHVD\_042201 (AvrPm3a, mating type (MAT) 1–2) with the Chinese isolate CHN\_52\_27 (avrPm3a, MAT 1–1) that display opposite phenotypes on the Pm3a containing near-isogenic line 'Asosan/8\*CC' (Fig 2A). From this cross, we generated a mixed population of an estimated 1500 F1 progenies on the susceptible wheat line 'Kanzler', which is devoid of *R* genes against *Bgt*. In the next step, we used conidiospores of this multiplied F1-1 progeny population to infect either 'Asosan/8\*CC', thereby creating a Pm3a-mediated selection pressure (Pm3a-selected bulk) or the susceptible wheat cultivar 'Kanzler' to create an unselected control bulk. Subsequently, we bulk-harvested conidiospores produced during the respective selection steps and subjected both bulks to DNA sequencing using Illumina paired-end reads to a coverage of approximately 120X.

The AvrPm $3^{a2/f2}$  gene is located in an effector cluster consisting of 19 members of the E008 candidate effector family [25]. Due to its complexity, the locus was not fully resolved in the previously published Bgt reference genome assembly of strain CHE 96224, where it was found to contain several sequence gaps and collapsed regions for highly similar gene copies [32]. Given the advances in long-read sequencing technologies, we saw the opportunity to generate an improved reference genome assembly that fully resolves the  $AvrPm3^{a2l/2}$  locus and potentially other collapsed or incomplete regions in the published Bgt reference genome CHE\_96224. To do so, we sequenced CHVD\_042201, which served as an avirulent parent in our AD assay, with PacBio HiFi to a coverage of 100X and assembled the genome using the hifiasm assembler (see S1 Note for details). This strategy resulted in a near-complete telomereto-telomere assembly with fully resolved centromeres and only two remaining sequence gaps in the highly expanded and repetitive rRNA encoding region on Chr-09 and a previously identified array of tandem repeats on Chr-04. Importantly, the new genome assembly also fully resolved the AvrPm3<sup>a2/f2</sup> locus and multiple other previously collapsed regions and thus represents a significant improvement in terms of genome resolution and continuity compared to previous Bgt assemblies (S1 Table and S1 Note).

To analyse the bulked DNA sequencing data generated as part of the AD assay pipeline, we mapped the Illumina reads from the *Pm3a*-selected bulk, the non-selected control bulk, and the parental isolates CHVD\_042201 and CHN\_52\_27 on the new genome assembly of CHVD\_042201 and identified 198'027 high-quality SNPs between the two parental isolates serving as genetic markers in the subsequent analysis. In the sequenced bulks, we expected regions unaffected by any selection to show a 1:1 ratio of parental genotypes at any given marker. In contrast, in regions under selection, we expected ratios to significantly deviate from a 1:1 ratio, with the genotype of the avirulent parent CHVD\_042201 being underrepresented



**Fig 2. High-resolution mapping and functional validation of the previously described**  $AvrPm3^{a^{2/f^2}}$  in a proof-of-concept study using the newly **developed** AD assay. (a) Virulence phenotypes of the parental *Bgt* isolates CHN\_52\_27 and CHVD\_042201 on the susceptible wheat line 'Kanzler' and the *Pm3a*-containing line 'Asosan/8\*CC'. (b) Statistical analysis of sequenced bulks selected on 'Kanzler' or 'Asosan/8\*CC' to detect regions with deviations from an expected 1:1 parental genotype ratio in the absence of selection. Individual datapoints represent–log10 transformed p-values of a G-test (y-axis) at individual SNPs marker positions plotted along the 11 chromosomes of *Bgt* isolate CHVD\_042201 (x-axis). A single genomic region on Chr-06 exhibits deviations from an expected 1:1 parental genotype inheritance in the *Pm3a*-selected bulk. The mapped *AvrPm3a* target interval ( $\geq$ 95% reads from virulent parent) encompassing 25 kb and the single effector gene *CHVD042201-04754* (*AvrPm3^{a2/f2}*) is depicted in a zoom-in (bottom). (c) Protein sequence alignment of AvrPm3^{a2/f2}. A found in the avirulent parental isolate CHVD\_042201 and AvrPm3^{a2/f2}. B in the virulent parental isolate CHV\_52\_27. Amino acid polymorphisms are highlighted in yellow. The signal peptide, the conserved Y/FxC motif, the C-terminal cysteine and the conserved position of a small intron, all serving as hallmarks of RNAse-like effectors (RALPH effectors) are highlighted in grey or by a small arrowhead (intron position). (d, e) *Agrobacterium*-mediated expression of AvrPm3^{a2/f2}. A, AvrPm3^{a2/f2}. B and Pm3a in *N. benthamiana*. Co-infiltrations (top) were performed with a 4 (effector): 1 (NLR) ratio. Leaves were imaged using a camera (d) or the Fusion FX imager system (e) at four days post inoculation. The assay was performed with n = 6 leaves and repeated a total of three times with similar results (total n = 18 leaves). (f) Quantification of HR intensity in the *N. benthamiana* expression assay depicted in (e). Datapoints from t

https://doi.org/10.1371/journal.ppat.1012799.g002

(i.e. depleted). As expected, we did not identify any region with a strong deviation from a 1:1 ratio for the control bulk selected on the susceptible cultivar 'Kanzler' that is devoid of any *R* genes against *Bgt* (Fig 2B). In contrast, in the *Pm3a*-selected bulk, we found a single region located on Chr-06 that showed strong deviation from a 1:1 ratio towards the genotype from

the virulent CHN 52 27 parental isolate (Fig 2B). The identified genomic segment overlapped with the previously identified E008 candidate effector cluster containing the AvrPm3<sup>a2/f2</sup> gene. Importantly, the region with the strongest depletion signal (i.e.,  $\geq$ 95% of reads from virulent parent) was constricted to 25 kb and encompassed a single gene CHVD042201-04754 (Fig 2B). The identified gene is identical to the  $AvrPm3^{a2/f2}$  variant  $AvrPm3^{a2/f2}$  -A, which was previously shown to be recognized by Pm3a [25,33]. Based on resequencing data, we determined that the CHVD042201-04754 homolog in the Pm3a-virulent isolate CHN\_52\_27 encodes for a protein variant with three amino acid changes compared to CHVD\_042201 (H36Q, G84E, E121D) (Fig 2C). This haplovariant was previously identified in isolates from China and Israel and was termed  $AvrPm3^{a2/f2}$ -B and designated as a putative  $AvrPm3^{a2/f2}$  gain-of-virulence allele [33]. To confirm this finding, we expressed the  $AvrPm3^{a2/f2}$  variants from the two parental isolates (without signal peptide) together with Pm3a in Nicotiana benthamiana using Agrobacterium-mediated transformation. Consistent with the results from the AD-assay, CHVD042201-04754 ( $AvrPm3^{a2/f2}$ -A) elicited a Pm3a-dependent hypersensitive response (HR), whereas the AvrPm3<sup>a2/f2</sup>-B from the virulent parent CHN\_52\_27 did not, thereby confirming  $AvrPm3^{a2/f2}$ -B to be a virulence allele (Fig 2D-2F). Based on these findings, we concluded that the avirulence allele depletion observed in the Pm3a-selected bulk is a direct consequence of differential recognition of AvrPm3<sup>a2/f2</sup> variants found in the two parental isolates. In conclusion, our proof-of-concept datasets showed that the AD-assay is a powerful new tool to identify avirulence effectors in Blumeria down to single gene resolution.

For this proof-of-concept study of the AD assay, we relied on the high-quality genome assembly of CHVD\_042201, which represented the avirulent parental isolate in our genetic cross. Even though the availability of a complete genome sequence of the avirulent parent is likely ideal to identify candidates genes, we wanted to test whether the AD assay succeeds in *AVR* identification also with an alternative reference genome. We therefore tested our AD pipeline using the previously published reference genome assembly of isolate CHE\_96224 (Bgt\_genome\_v3\_16) (32). Similar to the above-described analysis based on CHVD\_042201, we found no deviations from the expected 1:1 parental genotype ratio in the unselected F1-2 bulks but a single region with strong avirulence depletion signal on Chr-06 in the *Pm3a*-selected bulk (S1 Fig and S2 Table and S2 Note). The identified genomic region in the CHE\_96224 overlapped with the identified *AvrPm3*<sup>a2/f2</sup> locus in CHVD\_042201, thereby confirming that the AD approach succeeds in identifying *AvrPm3*<sup>a2/f2</sup> also with an alternative reference genome assembly (see S2 Note for details).

## The AvrPm60 effector is encoded by three tandem duplicated genes in the *Pm60*-avirulent isolate CHVD\_042201

Next, we aimed to use the AD assay to identify the so far unknown *AvrPm60*, the avirulence effector corresponding to the NLR *Pm60* [20]. The parental *Bgt* isolates CHVD\_042201 and CHN\_52\_27 displayed opposite virulence phenotypes on 'Kn199 *Pm60*' (Fig 3A), a transgenic line expressing the *Pm60* gene in the susceptible background 'Kn199' [20]. We therefore used the 'Kn199 *Pm60*' transgenic line to apply a *Pm60*-mediated selection to the CHVD\_042201 x CHN\_52\_27 F1 progeny population described above. Again, the AD-assay identified a single genomic region in which the *Pm60*-selected bulk showed a strong deviation from the 1:1 parental genotype ratio, with a depletion of the avirulent CHVD\_042201 parental genotype. Strikingly, the *AvrPm60* candidate locus partially overlapped with the above-described *AvrPm3<sup>a2/f2</sup>* locus on Chr-06 (Fig 3B). A parallel analysis of the *Pm60*-selected bulk based on the alternative CHE\_96224 reference genome assembly identified a single genomic region with strong depletion of the avirulent genotype on Chr-06 which overlapped with the



Fig 3. Mapping and functional validation of AvrPm60\_1 and AvrPm60\_2 using the AD assay. (a) Virulence phenotype of the parental Bgt isolates CHN\_52\_27 and CHVD\_042201 on the transgenic wheat line 'Kn199 Pm60'. (b) Statistical analysis of sequenced bulks selected on 'Kn199 Pm60' to detect regions with deviations from an expected 1:1 parental genotype ratio in the absence of selection. Individual datapoints represent-log10 transformed p-values of a G-test (y-axis) at individual SNPs marker positions plotted along the 11 chromosomes of Bgt isolate CHVD\_042201 (xaxis). A single genomic region on Chr-06 exhibits deviations from an expected 1:1 parental genotype inheritance in the Pm60-selected bulk. (c) Schematic zoom-in of the mapped AvrPm60 target interval (>95% reads from virulent parent) encompassing 667 kb and 16 candidate effector genes. Polymorphic candidate effectors are individually labelled, and non-synonymous SNPs indicated by a red line. The genomic region encompassing CHVD042201-04743, CHVD042201-04745 and CHVD042201-04747, which was found to be deleted in the virulent parental isolate CHN\_52\_27, is highlighted in yellow. (d) Protein sequence alignment of the identical effectors CHVD042201-04743/CHVD042201-04745 and CHVD042201-04747 (E103G). The E103G polymorphism is highlighted in yellow. The signal peptide, the conserved Y/FxC motif, the C-terminal cysteine and the conserved position of a small intron, all serving as hallmarks of RNAse-like effectors (RALPH effectors) are highlighted in grey or by a small arrowhead (intron position). (e) Co-expression of polymorphic effector candidates found within the mapped AvrPm60 locus together with Pm60 using Agrobacterium-mediated expression in N. benthamiana. Co-expression of GUS + Pm60 and AvrPm3<sup>a2/f2</sup>-A + Pm3a were used as negative and positive controls, respectively. Co-infiltrations were performed with a 4 (effector): 1 (NLR) ratio and imaged with a Fusion FX imager system 4 days post inoculation. HR intensity was quantified from three independent experiments with n = 6 leaves. Datapoints from individual experiments are indicated by symbols (cross, circle, triangle) and additionally color coded (total n = 18). Asterisks indicate statistical differences according to post-hoc pairwise Wilcoxon rank sum exact tests (\* p<0.05, \*\* p<0.01, \*\*\* p<0.001) performed after a significant Kruskal-Wallis test (p<0.05), P-values

were adjusted to account for multiple comparison using the Benjamini & Hochberg method. (f, g) *Agrobacterium*-mediated expression of AvrPm60\_1, AvrPm60\_2 and Pm60 in in *N. benthamiana*. Co-infiltrations (top) were performed with a 4 (effector): 1 (NLR) ratio. Leaves were imaged using a camera (f) or the Fusion FX imager system (g) at four days post inoculation. The assay was performed with n = 6 leaves and repeated a total of three times with similar results (total n = 18 leaves). (h) Quantification of HR intensity in the *N. benthamiana* expression assay depicted in (g). Datapoints from three independent experiments with n = 6 leaves are indicated by symbols (cross, circle, triangle) and additionally color coded (total n = 18). \* indicates statistical difference according to a Wilcoxon signed rank test (p<0.05). (i) Predicted three dimensional structures of AvrPm60\_1 (yellow) and AvrPm3<sup>a2/f2</sup> (grey) according to Alphafold 3 structural modelling. Both effector proteins are predicted to exhibit an RNAse-like structure.

https://doi.org/10.1371/journal.ppat.1012799.g003

identified region in CHVD\_042201, again highlighting the independence of the AD assay from individual reference genomes (S1 Fig and S2 Table and S2 Note).

The region with the strongest signal of avirulence allele depletion (i.e.  $\geq$ 95% of the reads originated from the virulent parent CHN\_52\_27) in the Pm60-selected bulk encompassed a region of 667 kb in CHVD\_042201 containing 16 annotated, high-quality genes, all belonging to the E008 candidate effector family (Fig 3C). We inspected the coding sequence of all genes in the locus with re-sequencing data from CHVD\_042201 and CHN\_52\_27 and found that only seven of the 16 candidate effectors, including the above-described  $AvrPm3^{a2/f^2}$  gene (CHVD042201-04754) exhibited polymorphisms between the parental isolates and therefore represented good AvrPm60 candidate genes (Fig 3C). Among the seven polymorphic genes, only four carried SNPs resulting in amino acid polymorphisms between the virulent and avirulent parents. For the remaining three genes (CHVD042201-04743, CHVD042201-04745, and CHVD042201-04747), we did not detect any alignment of sequencing reads from CHN\_52\_27, indicating that these three genes are deleted in the virulent parent (Fig 3C). Interestingly, CHVD042201-04743, CHVD042201-04745, and CHVD042201-04747 constitute tandem duplicates of the same effector gene, which are part of a 7.7kb duplicated segment with 99% sequence identity common to all three copies in the avirulent isolate CHVD\_042201. The resulting effector proteins CHVD042201-04743 and CHVD042201-04745 are identical, whereas CHVD042201-04747 differs by a single amino acid change (E103G) (Fig 3D).

To functionally validate AvrPm60, we co-expressed the seven polymorphic candidate genes found within the mapped locus together with Pm60 using transient overexpression in N. benthamiana. Both the CHVD042201-04743/CHVD042201-04745 and CHVD042201-04747 effectors triggered a Pm60-dependent HR response, whereas none of the other candidates or a GUS negative control resulted in cell death (Fig 3E-3H). Hence, we concluded that three genes CHVD042201-04743, CHVD042201-04745, and CHVD042201-04747 represent AvrPm60 by encoding two nearly identical effector proteins that we designated as AvrPm60\_1 (CHVD042201-04743/CHVD042201-04745) and AvrPm60\_2 (CHVD042201-04747). Interestingly, the AvrPm60\_2 variant carrying the amino acid polymorphism E103G elicited a slightly stronger HR response compared to AvrPm60\_1 (Fig 3E-3H). Additionally, we observed marginally higher protein accumulation of epitope tagged HA-AvrPm60\_2 compared to HA-AvrPm60\_1 in N. benthamiana, suggesting that differences in protein levels may contribute to the observed differences in recognition of the two AvrPm60 variants by Pm60 (S2 Fig). Similar observations, where higher protein abundance of AVRs correlated with stronger recognition, have been previously described for variants of the Bgt AVR effectors AvrPm3<sup>a2/f2</sup> and AvrPm17 [8,34].

#### AvrPm60\_1 and AvrPm60\_2 belong to the RNAse-like effector superfamily

The identified AvrPm60 effectors are part of the large effector family E008 with over 40 members, including the AvrPm3<sup>a2/f2</sup> effector (32). Like other members of the E008 family, the two AvrPm60 are small proteins of 121 amino acids in size, with the first 23 amino acids constituting a predicted signal peptide (Fig 3D). Similar to all previously identified AVRs in *Bgt*, the AvrPm60 proteins contain a Y/FxC motif and a conserved C-terminal cysteine, two features that were defined as hallmarks of RNAse-like effectors which comprise more than half of all effector proteins found in the *Blumeria* genus [29,35]. Indeed, structural modelling using Alphafold 3 [36] predicted that the AvrPm60 proteins exhibit an RNAse-like structure and are structurally similar to the E008 family member AvrPm3<sup>a2/f2</sup> from *Bgt* and the structurally validated RNAse-like effectors CSEP0064/BEC1054 [37] and AVRa10 [29] from the barley powdery mildew pathogen *Blumeria hordei* (Figs 3I and S3).

Multiple previously identified RNAse-like AVRs in *Bgt* were found to exhibit very high expression levels during early phases of infection [24–27]. Similarly, analysis of RNA sequencing data from five *Bgt* isolates at two days post inoculation (2dpi) showed that the *AvrPm60* genes are consistently among the top 5% of the highest expressed genes in each isolate (S4 Fig). In summary, the AvrPm60 effectors are bona-fide members of the RNAse-like effector class in *Blumeria*.

## Gain-of-virulence through deletion of *AvrPm60* genes is rare within the worldwide *Bgt* population but widespread in the US

The Pm60-virulent isolate CHN\_52\_27 used in the AD assay evades Pm60-mediated resistance due to a large-scale deletion encompassing all three AvrPm60 copies found in CHVD\_042201 (Fig 3C). We therefore aimed to investigate the frequency and distribution of this striking gain-of-virulence mechanism within the global Bgt population. To estimate the number of AvrPm60 gene copies, we used a previously described in-silico approach using publicly available resequencing data from 382 Bgt isolates [8,23,26,32,38,39]. This approach uses normalized read coverage as a proxy for the number of gene copies in each isolate. As a control we used the GAPDH gene which was previously shown to be present as a single copy gene, and  $AvrPm3^{a2/f2}$  found to occur in 1–4 copies in Bgt isolates [32]. As expected, the coverage analysis indicated that GAPDH occurs as a single copy gene in all 382 analysed isolates, whereas the analysis found signs of copy number variations for both the  $AvrPm3^{a2/f2}$  and the AvrPm60 gene in this worldwide Bgt panel (Figs <u>4A</u> and <u>S5</u>). Consistent with the literature, the coverage analysis indicates that all isolates carry at least one copy of the  $AvrPm3^{a2/f2}$ , although higher-order duplications with two or more copies are readily observed (S5 Fig), [32,33,40]. In contrast, we estimated that the majority of Bgt isolates contain two AvrPm60 copies (Fig 4A). However, we detected sizeable additional copy number variation in the Bgt diversity panel with some isolates containing a single AvrPm60 gene and isolates with three or more AvrPm60 copies. Consistent with the broad functionality of *Pm60* described in the literature [20], only a minority of isolates (13/382), including the Chinese isolate CHN\_52\_27, were devoid of any AvrPm60 copies as indicated by the absence of any sequencing coverage in our analysis (Fig 4A and Table 1). Interestingly, among the 13 isolates lacking AvrPm60 only one additional isolate originated from China (2 out of 63 Chinese isolates), while the remaining isolates all originated from the US, where 19% of investigated isolates carried the AvrPm60 deletion (Table 1). To experimentally confirm the deletion of AvrPm60 gene copies in specific isolates from China and the US, we designed AvrPm60-specific PCR primers based on conserved flanking sequences of all three AvrPm60 genes in the reference isolate CHVD\_042201. Using these primers we successfully amplified AvrPm60 from genomic DNA of CHVD\_042201 and six additional Bgt isolates with diverse geographic origin (Switzerland, UK, China, Japan, Argentina, US) for which our coverage analysis estimated between one and four AvrPm60 copies in the genome. In contrast, PCR amplification failed from CHN\_52\_27 and three isolates from the US with a



**Fig 4. Functional characterization of** *AvrPm60* **copy number variation and** *Pm60* **allelic variants.** (a) Copy-number estimation of *AvrPm60* gene copies in a worldwide diversity panel of 382 publicly available sequenced *Bgt* isolates. Genomic coverage of Illumina reads aligning to *AvrPm60* genes copies were normalized to the coverage of all genes in the genome. (b) Virulence phenotypes of 11 *Bgt* isolates with varying copy numbers of *AvrPm60* genes on the transgenic wheat line 'Kn199 Pm60' and 'Kn199', serving as a susceptible control. The estimated number of *AvrPm60* gene copies according to sequencing coverage analysis are indicated next to the isolate name with normalized sequencing coverage indicated in brackets. The isolates CHVD\_042201 and CHN\_52\_27 used for initial *AvrPm60* identification are shown as comparison. (c) Schematic representation of the three tested *Pm60* alleles originating from *Triticum urartu*. Yellow boxes indicate a pair of LRR repeats that show copy number variations between the tested Pm60 alleles. (d,e) *Agrobacterium*-mediated expression of AvrPm60\_1 (e), AvrPm60\_2 (d) with Pm60, Pm60a and Pm60b in *N. benthamiana*. Co-infiltrations were performed with a 4 (effector): 1 (NLR) ratio. Leaves were imaged at 4dpi using the Fusion FX imager system. The assay was performed with n = 6 leaves and repeated a total of three times with similar results (total n = 18 leaves). Boxplots represent quantification of HR intensity in the *N. benthamiana* expression assay. Datapoints from the three independent experiments are indicated by symbols (cross, circle, triangle) and additionally color coded. Different letters next to the boxplot represent statistical differences according to post-hoc pairwise Wilcoxon rank sum exact tests performed after a significant Kruskal-Wallis test (p<0.05), P-values were adjusted to account for multiple comparison using the Benjamini & Hochberg method.

https://doi.org/10.1371/journal.ppat.1012799.g004

predicted complete *AvrPm60* deletion, thereby confirming the results of the coverage analysis (S6 Fig). We then subjected the same 11 *Bgt* isolates to virulence phenotyping on the transgenic wheat line 'Kn199 + Pm60' and the susceptible control 'Kn199'. Importantly, all isolates with at least one *AvrPm60* gene copy in the genome exhibited an avirulent phenotype on the *Pm60* transgenic line similar to CHVD\_042201, whereas all tested isolates with *AvrPm60* gene deletions exhibited full virulence on *Pm60* wheat, comparable to the virulent CHN\_52\_27 isolate (Fig 4B). Hence, we conclude that deletion of *AvrPm60* genes represents a gain-of-virulence mechanism that allows *Bgt* to overcome *Pm60*-mediated resistance and that such *AvrPm60* gene deletions are particularly prevalent in the US, likely limiting *Pm60* efficacy in this geographic area.

origin	# isolates	deletion	1 copy	2 copies	>2 copies	ambiguous
Argentina	8	-	3	5	-	-
China	63	2	2	30	24	5
Egypt	48	-	-	33	6	9
Europe	30	-	5	5	20	-
Iran	10	-	-	7	3	-
Israel	31	-	2	16	11	2
Japan	10	-	4	5	-	1
Kazakhstan	3	-	-	1	2	0
Russia	38	-	4	23	7	4
Turkey	83	-	5	63	10	5
USA	58	11	38	3	-	6
Total	382	13	63	191	83	32

https://doi.org/10.1371/journal.ppat.1012799.t001

## *Pm60* and its allelic variants *Pm60a* and *Pm60b* recognize *AvrPm60\_1* and *AvrPm60\_2* with varying efficacy

The *Pm60* resistance gene was originally isolated from *T. urartu*, the progenitor of the A genome of hexaploid wheat. Allele mining in T. urartu identified two additional functional alleles of *Pm60*, termed *Pm60a* and *Pm60b*, that provide resistance against *Bgt* [20,41]. *Pm60a* and *Pm60b* differ from *Pm60* through the deletion or tandem duplication of two LRR repeats, respectively (Fig 4C) [20]. In the literature, the Bgt recognition spectra of *Pm60a*, *Pm60b* and *Pm60* are described as largely overlapping thus prompting us to hypothesize that the three Pm60 alleles recognize the same avirulence component in Bgt. To test this hypothesis, we co-expressed the three Pm60 alleles with the AvrPm60\_1 and AvrPm60\_2 in N. benthamiana. Both AvrPm60 effector variants induced a hypersensitive response upon co-expression with Pm60a, Pm60b and Pm60 but not with a uidA (GUS) or a *Pm3a* negative control (Figs 4D, 4E and S7A). To rule out an unspecific HR response or autoactivity of Pm60a and Pm60b upon expression in N. benthamiana, we coexpressed the three *Pm60* alleles with the non-corresponding AVR effector  $AvrPm3^{a2/f2}$  and *uidA* (GUS) which did not result in any HR responses (S7A and S7B Fig). In conclusion, we show that allthree Pm60 alleles specifically recognize the AvrPm60\_1 and AvrPm60\_2 avirulence components from Bgt.

Although both *AvrPm60* variants are recognized by all three *Pm60* alleles, we observed significant differences in the strength of the hypersensitive response depending on the specific *Avr/R* combination. For the stronger *AvrPm60\_2* variant, the *Pm60a* allele showed a marked reduction in the strength of HR when compared to *Pm60*. In contrast, recognition of the weaker *AvrPm60\_1* variant resulted in reduced HR output for both *Pm60a* and *Pm60b* alleles when compared to *Pm60* (Fig 4D and 4E). Importantly, we detected similar protein levels of epitope tagged Pm60-HA, Pm60a-HA and Pm60b-HA, indicating that the difference of recognition strength between the three Pm60 alleles are not due to differential protein accumulation in *N. benthamiana* (S7C Fig). In summary, we show that the three *Pm60* alleles, originating from *T. urartu*, recognize the same avirulence effectors AvrPm60\_1 and AvrPm60\_2, albeit with differences in the strength of the elicited HR. This indicates that the copy number variation in the LRR repeats does not influence the recognition specificity of the Pm60 alleles, but it might play a role in modulating the amplitude of the HR upon effector recognition.

#### Discussion

Breeding genetically resistant cultivars is a cornerstone of sustainable agricultural production, with global efforts focusing on the identification of new resistance sources and their efficient deployment against fast evolving pathogenic organisms [42]. In wheat, more than 460 resistance gene loci against various pathogens are genetically defined but only a small fraction of the underlying R genes have been cloned to date [6,12]. Owing to technological advances and significantly improved genomic resources, the speed of R gene cloning in wheat has increased tremendously in recent years and it is expected that the majority of defined R genes will be molecularly identified within the next decade [43,44]. However, the question remains how these resistance sources can be deployed in modern agricultural systems to durably withstand continually evolving pathogen populations. In a widely acclaimed perspective paper, Hafeez and colleagues in 2021 called for a concerted, international effort to generate a wheat R gene atlas to tackle these problems [6]. In the outlined vision, R gene identification in wheat must be accompanied by AVR identification and subsequent analysis of AVR diversity in corresponding pathogen populations in order to guide future breeding decisions and allowing for efficient, regionally tailored, R gene deployment. The authors argued that for the success of such an endeavor it is crucial to improve our current limited knowledge about AVR effectors in wheat pathogens and in particular develop methods to speed up AVR gene identification in order to keep up with the improved pace of *R* gene cloning in the host [6].

In contrast to other wheat pathogens where few or no AVR effectors have been identified to date, the field of AVR gene identification in Bgt is relatively advanced. Using various methods including classic genetic mapping, GWAS, mutagenesis and effector screens, a total of 8 AVR effectors have been cloned and molecularly characterized to date [8,13,23-27,31]. Due to the haploid genome and experimentally accessible sexual reproduction of Bgt, classic genetic mapping has proven to be a powerful tool. For example, genetic mapping approaches can outperform GWAS in resolving genetically complex traits with multiple components or when components occur at low frequency in the Bgt population. Furthermore, in contrast to effector screening approaches, genetic mapping does not solely rely on HR as the primary immune outcome and is better suited to identify non-canonical AVR effectors or suppressors of R genemediated resistance such as the previously described SvrPm3 [25]. However, the requirement to isolate, phenotype and genotype individual F1 progenies represents a major obstacle in genetic mapping experiments, leading to work- and cost-intensive projects in the range of 1-2 years. The newly developed AD assay described here leverages the many strengths of genetic mapping in Bgt while eliminating the work-intensive characterization of individual F1 progenies. It thereby shortens project timelines to approximately 4 months and reduces associated sequencing efforts to a selected and an unselected F1-2 pool resulting in much lower costs. Avoiding the most work-intensive steps in mapping projects furthermore allows AVR identification to be parallelized by either selecting the same progeny population on multiple R gene containing lines, as done in this study on Pm3a and Pm60, or even by simultaneously generating multiple progeny populations arising from different genetic crosses that segregate for different AVR factors.

Our proof-of-concept experiment aiming at the mapping of  $AvrPm3^{a2/f2}$  identified a narrowly defined genomic region of 25 kb harboring a single candidate effector gene (i.e.  $AvrPm3^{a2/f2}$ ), exemplifying the mapping power and resolution of the AD assay (Fig 2B). In the case of AvrPm60 the mapped interval was however significantly bigger and encompassed a region of 667 kb with 16 effectors in total (7 polymorphic), indicating that mapping resolution varies depending on the mapped factor (Fig 3C). We hypothesize that these differences arise in part from the fact that AvrPm60 is represented by three genes spread over 130 kb of sequence.

The mapping resolution is likely also influenced by the rate of recombination within the mapped *AVR* locus, which was shown to differ throughout the *Bgt* genome [32] and the strength of selection exerted by the resistant wheat line used to generate the selected F1-2 bulks.

Interestingly, we observed near complete selection for the virulence allele (i.e.  $\geq$ 95% of sequenced reads) in *Pm3a*- and *Pm60*-selected pools after a single asexual reproduction cycle likely due to the strong resistance effect exerted by both *R* genes. We are however optimistic that the AD assay could also be used to map *AVR* effectors corresponding to *R* genes with weaker resistance effects, such as *Pm8* which was shown to result in partial resistance against isolates carrying *AvrPm8* under endogenous *Pm8* expression levels and only resulted in complete resistance in transgenic overexpression lines [27]. In such cases, the deviation from a 1:1 inheritance ratio in selected progeny pools might be less drastic, resulting in larger mapped intervals and consequentially a higher number of candidate genes. The situation could however be alleviated by extending the number of asexual cycles under selection or, if available, by the use of transgenic overexpression lines with stronger or complete resistance.

Several recent studies have shown that the genetic control of avirulence phenotypes in *Bgt* can be complex and can involve multiple genetic loci [8,13,23,25]. In the case of a rye introgression carrying *Pm17* and a genetically linked unknown second resistance gene, a QTL mapping approach using ~120 F1 progenies successfully mapped both corresponding *AVR* loci in *Bgt* [8]. However, in another study the complex avirulence landscape for several *Pm3* alleles was only partially resolved in a biparental mapping population of ~140 F1 progenies [25]. It will be interesting to see whether the newly developed AD assay, due to the high number of processed progenies (~1500 in this study), could improve mapping resolution in these complex cases that include multiple *AVR* loci. We hypothesize that the resolution achieved in the AD assay could be further improved by extending the F1 progeny population and in particular by sequencing bulks at significantly higher coverages (~120X in this study), thereby capturing a more detailed picture of the recombination landscape within the F1 progeny population.

In conclusion, we argue that the AD assay is a powerful new tool for *AVR* identification in *Bgt* that will significantly speed up *AVR* cloning in this important wheat pathogen in the future. Furthermore, we hypothesize that the AD assay could be adapted to map avirulence components in other pathosystems, where the sexual reproduction cycle of the pathogen is experimentally accessible and genetic resources of the host include clearly defined *R* genes that allow for efficient selection of progeny pools.

Our knowledge about AVR effectors in cereal powdery mildews has improved significantly in recent years. With a total of 8 known and functionally validated AVRs in the wheat infecting Bgt and another 6 AVRs from barley powdery mildew (Blumeria hordei), common principles of AVR effectors in this group of cereal pathogens can be defined [8,13,23-28,45,46]. Interestingly, all 14 AVRs belong to the RALPH effector superfamily of structurally similar proteins exhibiting an RNAse-like fold, but lacking RNAse activity [28,29]. Apart from the presence of a signal peptide, these proteins share several characteristics such as a N-terminal Y/FxC motif and a conserved C-terminal cysteine involved in disulfide bridge formation and hence stability of the RNAse fold [29,37]. Furthermore, RNAse-like effector genes share a single small intron at a conserved position indicating they evolved through diversification from the same ancestral gene, although they often share little amino acid identity [29,47]. While RNAse-like effectors can also be found in other phytopathogenic fungi, this group of effectors appears to be strongly expanded within the Blumeria genus where it comprises more than half of the total effector complement [29,32,35]. In this study, we found yet another three members of this effector superfamily to encode AvrPm60\_1 and AvrPm60\_2, thereby further highlighting the importance of the RNAse-like effectors in cereal powdery mildews. Interestingly, the previously

described AvrPm3<sup>a2/f2</sup> and the newly identified AvrPm60 effectors are part of the same effector family E008, whereas their corresponding NLRs belong to phylogenetically distinct NLR clades [48]. The question remains, however, why RNAse-like effectors in general, and the E008 family specifically, are especially prone to be recognized by NLRs. It has been hypothesized that the high expression levels observed during early stages of infection for many RNAse-like *AVRs*, including *AvrPm3*<sup>a2/f2</sup> and the newly identified *AvrPm60* genes (S2 Fig), makes this class of effectors predestined to be recognized by the hosts immune system [32,49].

Another characteristic observed with AVRs in Bgt is extensive copy number variation of AVR genes within the global pathogen population. For example,  $AvrPm3^{d3}$  was found to occur in up to six copies in some isolates, while other AVRs, such as  $AvrPm3^{a2/f2}$  or AvrPm17, were found to occur in one to four copies (S3 Fig) [8,24,32]. Interestingly, we found three neighboring, nearly identical AvrPm60 genes in the Pm60-avirulent isolate CHVD\_042201, with the majority of isolates within the global collection containing two or three copies (Fig 4A). We hypothesize that gain of virulence mutations affecting a single AvrPm60 gene copy might therefore not suffice to overcome Pm60 resistance. This is consistent with the observation that complete deletion of the AvrPm60 locus, as observed in the Pm60-virulent isolates (Fig 4A) and 4B). One probable cause of large genomic deletions, such as that of the AvrPm60 locus, is recombination between nearby repetitive sequences such as transposable elements. However, using short-read sequencing data for the isolates carrying AvrPm60 deletions, we were unable to pinpoint the exact mechanism behind the deletions. Ideally, long-read genome assemblies generated by third-generation technologies will allow to address this question in the future.

Interestingly, *AvrPm60* deletion is rare among Chinese *Bgt* isolates with only two out of a 63 analyzed isolates exhibiting complete absence of *AvrPm60* genes (Fig 4A and Table 1). This observation is in line with earlier findings showing that *Pm60* provides broad resistance against the Chinese *Bgt* population (20, 41). By contrast, deletion of *AvrPm60* genes is relatively common among *Bgt* isolates from the US (Fig 4A and 4B and Table 1). This finding is intriguing, given that we are not aware of any documented use of *Pm60* in agricultural production in the US. Nevertheless, we hypothesize that the frequent deletion of *AvrPm60* in this *Bgt* subpopulation could be the result of previous use of *Pm60* or *Pm60-like* resistance genes in this region, knowingly or unknowingly. Alternatively, it could be the consequence of the use of yet another resistance gene that recognizes the same effector proteins. Irrespective of underlying reasons for the frequent absence of *AvrPm60s* in the US *Bgt* population, we conclude that the resistance provided by *Pm60* is locally ineffective and therefore will likely provide only partial protection against *Bgt* infection in this geographic region.

A study by Zou and colleagues (2022) showed that the resistance provided by *Pm60a* and *Pm60b* resembles the one mediated by *Pm60* and concluded that the three alleles might have overlapping, albeit not identical, recognition spectra. In their study the authors showed that in particular *Pm60a* only provides resistance against a subset of isolates recognized by *Pm60* and *Pm60b*, which might indicate weaker or partially divergent recognition activity of *Pm60a* as compared to the other two alleles found in *T. urartu* [20,41]. This agrees with the findings in this study where we could show that AvrPm60\_1 and AvrPm60\_2 trigger HR immune responses in the presence of the Pm60, Pm60a and Pm60b NLRs albeit at varying levels (Fig 4D and 4E). For AvrPm60\_2 we observed strong HR responses upon co-expression with Pm60 and Pm60b but a weaker response with Pm60a (Fig 4D), corroborating the initial observations of Zou and colleagues [20,41]. Furthermore, co-expression of Pm60a and Pm60b with AvrPm60\_1 resulted in a significantly weaker HR response as compared to Pm60 (Fig 4E). Our findings indicate that the deletion of two LRR repeats in Pm60a results in a lower sensitivity towards AvrPm60\_1 and AvrPm60\_2 whereas the duplication of the same two LRR repeats

only influences AvrPm60\_1 recognition. In conclusion, these observations indicate that Pm60 outperforms its allelic variants Pm60a and Pm60b in their ability to recognize AvrPm60\_2 and in particular AvrPm60\_1 effectors and should therefore be prioritized for wheat breeding.

Interestingly several allele pairs of the Pm3 allelic series were shown to recognize the same AVR protein with differing recognition strength. The alleles Pm3a (strong) and Pm3f (weak) were shown to recognize AvrPm3<sup>a2/f2</sup>, Pm3b (strong) and Pm3c (weak) recognize AvrPm3<sup>b2/</sup> c<sup>2</sup> and finally, Pm3s (strong) and Pm3m (weak) a so far unknown AvrPm3<sup>m/s</sup> [24,25,50]. In the case of Pm3 diversity the underlying amino acid polymorphisms defining strong and weak alleles were identified and reside within the NB-ARC domain of the NLR [50]. In contrast, the weaker Pm60a and Pm60b alleles are defined by the lack or duplication of two LRR repeats, respectively, compared to the stronger Pm60 allele. Thus, different categories of polymorphisms within NLR proteins can influence their strength and ability to recognize AVR proteins, highlighting the importance of studying NLR diversity in order to define the most potent variants for wheat protection.

Multiple studies identified additional *Pm60* diversity in wild emmer wheat (WEW, *Triticum dicoccoides*) and defined 11 haplotypes that differed from *Pm60* found in *T. urartu* predominantly by single amino acid polymorphisms affecting the CC, the NB-ARC domain as well as the LRR repeats [51–53]. While resistance activity of some *Pm60* alleles in WEW was verified, the ability of the remaining *Pm60* diversity to recognize *Bgt* remains unknown. The identification of *AvrPm60\_1* and *AvrPm60\_2* therefore provides the opportunity to study, and potentially validate additional *Pm60* alleles from WEW as well as investigate the consequences of polymorphisms found within *Pm60* on its ability to recognize the two known *AvrPm60's* and, potentially, additional effector proteins.

Indeed, the identification and cloning of *Pm60* and *AvrPm60* genes in combination with investigations into their natural diversity opens new avenues of research and provides the opportunity to study the consequences of individual polymorphisms. These investigations will be crucial to estimate the value of *Pm60* for breeding and deployment and could, as exemplified by the high frequency of *AvrPm60* deletion observed in the US *Bgt* population, inform and guide regional deployment of resistance alleles. These considerations exemplify how the *R/AVR* atlas envisioned by Hafeez and colleagues could improve wheat resistance breeding in the future and how new technologies for rapid and cost-efficient *AVR* identification, such as the AD assay described in this study, can pave the way for pathogen-informed resistance gene deployment.

#### Methods

#### Bgt isolates, sexual crosses and selection of F1-1 populations

The *Bgt* isolate CHVD\_042201 (mating type MAT 1–2) was collected from a powdery mildew infected wheat field in Begnins, canton of Vaud, Switzerland in spring 2022 and subsequently single spore isolated twice in order to ensure a genetically uniform culture. The *Bgt* isolate CHN\_52\_27 (MAT 1–1) has previously been described in [26,54]. All other *Bgt* isolates used in this study have previously been described in [39]. *Bgt* isolates were maintained clonally on the susceptible wheat cultivar 'Kanzler'. For conidiospore production, infected leaf segments were placed on food grade agar plates (0.5%, PanReac AppliChem) supplied with 4.23mM benzimidazole and incubated at 20°C as described previously [55]. The sexual cross between CHVD\_042201 and CHN\_52\_27 was performed as previously described [56] by co-infecting the susceptible wheat cultivar 'Kanzler'. Leaf segments harboring chasmothecia were harvested and dried at room temperature for several weeks. For ascospore ejection, dried chasmothecia were exposed to high humidity using Whatman filter paper soaked in sterilized water for up to

10 days and arising F1 progeny collected and grown on the susceptible wheat cultivar 'Kanzler'. Resulting F1-1 conidiospores were subsequently subjected to bulk selection on wheat cultivars 'Kanzler' (no selection), 'Asosan/8\*CC' (Pm3a selection) or 'Kn199 + Pm60' transgenic plants (Pm60 selection). Selected F1-2 conidiospore bulks were harvested after 10 days and fungal DNA extracted as described below.

#### Plant material, virulence scoring

The *Pm3a* containing line 'Asosan/8\*CC' has been previously described in [25]. The 'Kn199 + *Pm60*' transgenic line, expressing *Pm60* under endogenous promoter and terminator sequences in the susceptible 'Kn199' background, has been previously described in [20].

To determine virulence phenotypes, leaf segments of the cultivars 'Kanzler', 'Asosan/8\*CC' or 'Kn199+*Pm60*' were placed on agar plates as described above, infected with the indicated *Bgt* isolates and disease phenotypes imaged at 8–9 days post infection. Virulence scoring was performed on at least six biological replicates for each tested interaction. Representative images were chosen for the depiction of virulence phenotypes throughout the manuscript.

#### DNA sequencing and genome assembly

Fungal DNA was extracted from conidiospores using a previously described CTAB/phenolchloroform extraction procedure [25]. For the CHVD\_042201 genome assembly, 5µg of high molecular weight DNA was used for library preparation and PacBio HiFi sequencing was performed on the PacBio Sequel Ile platform using a 30h movie at the Functional Genomics Center Zurich (FCGZ). Resulting PacBio HiFi raw data is available at the sequence read archive (SRA, accession: PRJNA1131794).

PacBio HiFi reads were assembled using HiCanu [57], HiFlye [58] and hifiasm [59] as described in <u>S1 Note</u>. HiCanu assembly was performed with the options genomeSize = 141m -pacbio-hifi. HiFlye assembly was performed with the options -g 141m –pacbio-hifi. Hifiasm assembly was performed with the options -t16 -l0 -f0 --hg-size 141m. Subsampling of PacBio HiFi reads was performed with seqtk sample command (https://github.com/lh3/seqtk).

Whole genome alignments of assemblies against Bgt\_genome\_v3\_16 was performed using the mummer suite (v4.0.0) [60] using the command nucmer. Subsequent plots were rendered with the mummerplot command using the following specification: --filter --color --png. Subsequently, plots were produced with gnuplot.

Blast searches of PacBio HiFi reads against the mitochondrial sequence of CHE\_96224 (Genebank: MT880591.1) were performed using the BLAST+ suite (v2.12.0) [61] with the following specifications: -qcov\_hsp\_perc 50 and all reads that aligned to the mitochondrial genome were retained. Subsequently, these reads were used to perform an assembly using hifiasm with the -l0 option.

The final assembly was polished with Illumina reads from isolate CHVD\_042201 that were mapped against the assembly with the method described in [27]. Subsequent polishing was performed with Pilon (v1.24) [62] with the following specifications: --fix bases --changes. The polished genome assembly of isolate CHVD\_042201 is available as Bgt\_CHVD\_042201\_genome\_v1 from the European Nucleotide Archive (ENA) under the accession number GCA\_964289775and on Zenodo (https://zenodo.org/records/11233413).

#### Annotation

Annotation of the Bgt\_CHVD\_042201\_genome\_v1 was performed using the MAKER2 software (v2.31.11) [63], available from the European Galaxy server (https://usegalaxy.eu/). Repeat masking of the genome was achieved using the TREP database (trep-db\_nr\_Rel-19.fasta and trep-db\_proteins\_Rel-19.fasta) available at https://trep-db.uzh.ch/. We used the prot2genome option of MAKER2 to create a homology-based draft annotation of Bgt\_CHVD042201\_geno-me\_v1 in two rounds. The first round was performed using the predicted proteome of CHE\_96224 (v4\_23, https://zenodo.org/records/7018501), and the second round using the proteome of *Bh* strain DH14 (https://github.com/lambros-f/blumeria\_2017/tree/master/ annotation\_genome\_dh14). Genes predicted in the second round were only included if they did not overlap with any genes predicted in the first round. The annotation of of Bgt\_CHVD042201\_genome\_v1 is available on Zenodo: (https://zenodo.org/records/11233413)

#### Avirulence depletion assay

DNA from the *Bgt* isolates CHVD\_042201 and CHN\_52\_27 was sequenced at the Functional Genomics Center Zurich (FGCZ). Sequencing libraries were generated using the Illumina Trueseq Nano protocol and sequencing was performed on the Illumina Novaseq 6000 platform. DNA from unselected, *Pm3a-* or *Pm60-*selected F1-2 bulks was sequenced to a coverage of ~120X with our commercial partner Novogene UK on a NovaSeq X Plus platform. Mapping of Illumina reads was performed as described previously [27]. For the analysis of the bulk-sequencing data, we established a pipeline specifically tailored towards the haploid genomic structure of *Bgt*. All steps of the pipeline were executed using a custom Python script, available on Github: https://github.com/MarionCMueller/AD-assay.

In detail, the pipeline first identified high-quality single nucleotide polymorphisms (SNPs) in the two parental isolates, CHVD\_042201 and CHN\_52\_27 as follows: Illumina mapping files were simultaneously used to perform SNP calling with FreeBayes (v1.3.6) [64], using the following options: --haplotype-length 0, --min-alternate-count 20, --min-alternate-fraction 0, --pooled-continuous, and --limit-coverage 400. The resulting polymorphic sites were further filtered to retain only those SNP positions where both parents had at least 10 reads and exhibited opposite genotypes. Genotypes were only accepted if 95% of the reads supported the genotype. Subsequently, the pipeline identified SNPs in the alignment files (BAM files) of the unselected and selected bulk only at the polymorphic positions between the parental isolates using FreeBayes with the options: --haplotype-length 0, --min-alternate-count 1, --min-alternate-fraction 0, --pooled-continuous, and --report-monomorphic.

The subsequent statistical analysis was conducted in R using the functions provided in the BSA\_Blumeria\_functions.R object available on GitHub at https://github.com/ MarionCMueller/AD-assay. To ensure the removal of sites exhibiting copy number variation in one of the parental isolates, the read coverage at each marker position was analysed. Positions with sequencing coverage either above or below twice the standard deviation of the coverage of all sites were excluded. Next, the G.test() function of the R package RVAideMemoire was utilized to assess the deviation of the parental SNP ratio from an expected 1:1 distribution for unselected marker positions. Finally, resulting p-values were averaged over 10 SNPs using the runner() command. Scripts used for analysis are available as an R Markdown object from Github (https://github.com/MarionCMueller/AvrPm60).

#### Cloning

For expression of fungal effectors in *N. benthamiana*, the predicted signal peptide (SignalP4.0 [65]) was removed and replaced by a start codon. The remaining coding sequences of all effector candidates were codon-optimized for expression in *N. benthamiana* based on the codon-optimization tool of Integrated DNA technologies (https://eu.idtdna.com). Optimized sequences were gene synthesized with gateway compatible flanking attL sites with our commercial partners (BioCat GmbH https://www.biocat.com; Thermo Fisher Scientific https://

www.thermofisher.com). For HA epitope tagging of AvrPm60\_1 and AvrPm60\_2, N-terminal HA epitope tags were introduced by site-directed mutagenesis using Phusion HF Polymerase (New England Biolabs) and the primers listed in S3 Table. Linearized PCR fragments were subsequentially phosphorylated and ligated using Polynucleotide Kinase and T4 Ligase (New England Biolabs) according to the manufacturer. The resulting gateway compatible entry clones were subsequentially mobilized into the binary expression vector pIPKb004 [66] using Invitrogen LR clonase II according to the manufacturer. The construct pIPKb004--HA-AvrPm3<sup>a2/f2</sup>, used as a control for effector protein expression, has been previously described [24].

For expression of Pm3a in N. benthamiana we made use of a pIPKb004-Pm3a-HA construct that has been previously described [24,25]. For the expression of *Pm60*, *Pm60a* and Pm60b in N. benthamiana we first amplified the Pm60 coding sequence from pEarlyGate-Pm60 described in [20] using KAPA HiFi Polymerase (KAPA Biosystems) with the primers listed in S3 Table and cloned the resulting PCR amplicon into the gateway compatible entry vector pDONR221 using BP clonase (Invitrogen), resulting in pDONR221-Pm60. The polymorphic regions defining the Pm60a and Pm60b alleles were gene synthesized with our commercial partner Thermo Fisher Scientific (https://www.thermofisher.com) and introduced into the Pm60 coding sequence using PCR amplification with KAPA HiFi Polymerase (KAPA Biosystems) and the primers listed in S3 Table applying the In-Fusion cloning method (Takara Bio) according to the manufacturer, resulting in pDONR221-Pm60a and pDONR221-Pm60b. To produce HA-epitope tagged Pm60, Pm60a and Pm60b, C-terminal HA epitope tags were introduced by site-directed mutagenesis using Phusion HF Polymerase (New England Biolabs) and the primers listed in S3 Table. Linearized PCR fragments were phosphorylated and ligated as described above, resulting in pDONR221-Pm60-HA, pDONR221-Pm60a-HA and pDONR221-Pm60b-HA. The entry clones of Pm60, Pm60-HA, Pm60a, Pm60a-HA, Pm60b and Pm60b-HA were mobilized into pIPKb004 as described above. The sequences of all DNA fragments produced by gene-synthesis can be found in S4 Table. All constructs in the binary expression vector pIPKb004 were transformed into A. tumefaciens strain GV3101 using freeze-thaw transformation [67].

## Co-expression of *AVR* candidates and *R* genes for HR quantification in *N*. *benthamiana*

*Agrobacterium*-mediated transient expression of effector candidates and resistance genes in *N. benthamiana* was achieved with the detailed protocol described in [24]. For co-expression of *AVR* candidates and corresponding *R* genes, *Agrobacteria* OD1.2 were mixed in a 4:1 ratio (AVR:R) prior to infiltration. HR imaging and quantification was performed 4–5 days after *Agrobacterium* infiltration using a Fusion FX imaging system (Vilber Lourmat <u>https://www.vilber.com/</u>) and the Fiji software [68] as described previously in [24]. Wilcoxon signed rank tests were performed in R using the function pairwise.wilcox.test() with parameter paired = TRUE. Kruskal-Wallis tests were performed in R using the kruskal.test() function followed by post-hoc analysis using pairwise Wilcoxon rank sum exact tests with function pairwise.wilcox.test(). To account for multiple comparison in post-hoc testing, p-values were adjusted using the Benjamini & Hochberg method (p.adjust.method ="BH").

#### Western blotting

To verify efficient protein production in *N. benthamiana*, individual constructs were expressed as described above and harvested two days after infiltration. The tissue was ground in liquid nitrogen and resuspended in 2x modified Laemmli-Buffer (100 mM Tris-HCl pH 6.8, 200 mM

DTT, 2% SDS, 20% glycerol, 0.04% bromphenol blue), boiled for 5 minutes and centrifuged for 10min at 10'000g. Total protein extracts were separated on SDS polyacrylamide gels (4–20% gradient gel, homemade), and subsequently blotted to a nitrocellulose membrane (Amersham Protran 0.2  $\mu$ m NC) using semi-dry transfer. Equal loading of protein samples was assessed by staining for total protein with Ponceau S. Detection of HA epitope tagged proteins was achieved using anti-HA-HRP antibody (monoclonal, clone 3F10, Roche) at a dilution of 1:3000. Peroxidase chemiluminescence was imaged using WesternBright ECL HRP substrate (Advansta) and a Fusion FX imaging System (Vilber Lourmat, Eberhardzell, Germany). Protein expression in *N. benthamiana*, protein extraction and Western blotting was performed three times with similar results for each Western blot analysis.

#### Gene expression analysis

To quantify gene expression, we used previously published dataset of *Bgt* isolates CHE\_96226, CHE\_94202, GBR\_JIW2, ISR\_7 and CHN\_17\_40 [8,27,69]. Accession numbers of the RNAseq libraries are listed in <u>S5 Table</u>. RNAseq reads were pseudoaligned to the CHVD\_042201 CDS using the salmon software v1.4.0 [70]. First, CHVD\_042201 CDS (Bgt\_CHVD042201\_CDS\_v1\_1.fasta) was indexed using the command salmon index. Then, single or paired end reads were quantified with the command salmon quant -l A. Subsequently, raw read counts were converted to RPKM values using the edgeR package 3.40.2 [71] using the rpkm() command. Plots were generated with ggplot2 v3.4.3 using a custom R script in RStudio v2023.03.0+386 [72,73]. An R Markdown script detailing all code used in this analysis to conduct read count, quantification and plot generation is available on Github (https:// github.com/MarionCMueller/AvrPm60).

#### Copy number variation

Analysis of copy number variation was performed based on previously published *Bgt* diversity data available described in [23,39]. Sequencing reads were aligned to the *Bgt* reference genome Bgt\_genome\_v3\_16 using the method described in [39]. Subsequently read coverage for each gene was extracted and normalised to the average coverage of all genes in the genome with a previously published script available on Github (https://gist.github.com/caldetas/24576da33d1ff91057ecabb1c5a3b6af). Genes exhibiting a normalized coverage below 0.1 were considered deleted. A table with normalized coverage values for all isolates is available here: (https://github.com/MarionCMueller/AvrPm60). Data was visualized and further analysed with a custom R script available on Github (https://github.com/MarionCMueller/AvrPm60).

#### AlphaFold 3 modelling

For AlphaFold 3 modelling of AVRPM3A2/f2 and AVRPM60, the signal peptides were removed based on predictions from SignalP 5.0 (https://services.healthtech.dtu.dk/services/SignalP-5.0/). The structures were then predicted using the AlphaFold 3 server (https://alphafoldserver.com/) [36]. Visualization of the structures was performed using ChimeraX software [74].

#### **Supporting information**

S1 Note. PacBio HiFi-based telomere-to-telomere assembly of Swiss *Bgt* isolate CHVD\_042201. (DOCX)

## **S2** Note. Performing the AD assay based on alternative reference genomes. (DOCX)

**S1 Fig. The AD assay performed on reference genome Bgt\_genome\_v3\_16.** The AD assay based on the CHVD\_042201 x CHN\_52\_27 cross was performed using the reference genome of Bgt\_genome\_v3\_16 (CHE\_96224). Cultivars used for selection are indicated in the respective plots. The Y-axis indicates–log10(p-values) of G-test statistics used to test for a deviation from 1:1 parental genotype ratio expected in the absence of selection. G-test values were averaged over 10 neighbouring SNPs plotted along the 11 chromosomes of *Bgt* isolate CHE\_96224 (x-axis).

(TIFF)

S2 Fig. *Agrobacterium*-mediated expression of HA-AvrPm60\_2 results in slightly higher protein levels compared to HA-AvrPm60\_1 in *N. benthamiana*. Detection of HA-AvrPm60\_1 and HA-AvrPm60\_2 by anti-HA western blotting (top panel). HA-AvrPm3a2/f2 served as a positive control, protein extracts from uninfiltrated *N. benthamiana* leaf areas served as a negative control. Total protein Ponceau S staining is shown as a loading control (bottom panel). Black arrows indicate 10 kDa and 15 kDa protein size markers. Protein expression, protein extraction and Western blot analysis were performed three times with similar results.



**S3 Fig. AvrPm60 variants are predicted to adopt an RNAase-like fold.** The predicted threedimensional structures of AvrPm60\_1 and AvrPm60\_2, according to Alphafold 3 structural modelling, are shown. The predicted AvrPm60\_1 structure, shown in Fig 3E, is highlighted in yellow. The experimentally validated structures of AVRA10 (8OXK) [29] and BEC1054 (6FMB) [37] are shown as comparison. (TIFF)

S4 Fig. *AvrPm3*<sup>*a*2/*f*2</sup> and *AvrPm60* are among the highest expressed genes in five *Bgt* isolates. Distribution of all expressed (rpkm >0) genes in the *Bgt* isolates CHE\_96224, CHE\_94202, GBR\_JIW2, ISR\_7 and CHN\_17\_40 are displayed as log2(rpkm) values. The coding sequence of isolate CHVD\_042201 was used as a reference for quantification. All RNA sequencing data were obtained at 2 days post infection on the susceptible wheat cultivar 'Chinese Spring'. The solid line and the dashed line indicate the median and the 95% quartile of all expressed genes, respectively. Expression of *AvrPm3*<sup>*a*2/*f*2</sup> (*CHVD042201-04754*) and *AvrPm60* are indicated by a yellow and red dot respectively. *AvrPm60* expression represents the sum of the rpkm values of all three *AvrPm60* genes in the reference isolate CHVD\_042201 (*CHVD042201-04743*, *CHVD042201-04745* and *CHVD042201-04747*). (TIFF)

**S5 Fig. Distribution of normalized genomic coverage of the** *GAPDH* **and** *AvrPm3*<sup>*a2/f2*</sup> **genes.** For each gene, the genomic coverage of sequencing reads was normalised to the coverage of all genes in the genome. (TIFF)

**S6 Fig. PCR verification of the** *AvrPm60* **gene deletion in** *Bgt* **isolates originating from China and the US.** *AvrPm60* specific primers were designed in conserved regions flanking the three *AvrPm60* genes (CHVD042201-04743, CHVD042201-04745 and CHVD042201-04747) in CHVD\_042201 with a predicted amplicon size of 1179bp. Fungal *GAPDH* served as a positive control. The estimated number of *AvrPm60* gene copies according to sequencing coverage analysis are indicated next to the isolate name with normalized sequencing coverage indicated

in brackets. The isolates CHVD\_042201 and CHN\_52\_27 used for initial *AvrPm60* identification are shown as comparison.

(TIFF)

S7 Fig. The HR response of Pm60, Pm60a and Pm60b is specific towards AvrPm60 effectors and does not derive from NLR autoactivity. (a) Agrobacterium-mediated co-expression of Pm60, Pm60a and Pm60b with the non-corresponding AVR effector AvrPm3<sup>a2/f2</sup>, and coexpression of AvrPm60\_1 and AvrPm60\_2 with the non-corresponding NLR Pm3a does not result in a detectable HR response. Co-expression of AvrPm60\_2 with Pm60 served as a positive control. Asterisks above the boxplots indicate statistical differences compared to the positive control according to post-hoc pairwise Wilcoxon rank sum exact tests (\* p < 0.05, \*\* p < 0.05, 0.01, \*\*\* p< 0.001) performed after a significant Kruskal-Wallis test (p<0.05). P-values were adjusted to account for multiple comparison using the Benjamini & Hochberg method. (b) Pm60, Pm60a and Pm60b do not exhibit autoactivity upon co-expression with GUS. Co-expressions of AvrPm60\_2 with Pm60a or Pm60b served as positive controls. Coinfiltrations in (a) and (b) were performed with a 4(effector): 1 (NLR) ratio. Leaves were imaged at 4dpi using the Fusion FX imager system. The assay was performed with n = 6 leaves and repeated a total of three times with similar results (total n = 18 leaves). Boxplots represent quantification of HR intensity in the N. benthamiana expression assay. Datapoints from the three independent experiments are indicated by symbols (cross, circle, triangle) and additionally color coded. Asterisks indicate statistical differences according to post-hoc pairwise Wilcoxon rank sum tests (\* p<0.05, \*\* p< 0.01, \*\*\* p< 0.001) performed after a significant Kruskal-Wallis test (p < 0.05). P-values were adjusted to account for multiple comparison using the Benjamini & Hochberg method. (c) Detection of Pm60-HA, Pm60a-HA and Pm60b-HA expressed in N.benthamiana by anti-HA western blotting (top panel). Pm3a-HA served as a positive control, protein extracts from uninfiltrated N. benthamiana leaf areas served as a negative control. Total protein Ponceau S staining is shown as a loading control (bottom panel). Black arrows indicate 130 kDa and 250 kDa protein size markers. Protein expression, protein extraction and Western blot analysis were performed three times with similar results.

(TIFF)

S1 Table. Genome stastistics of Bgt\_CHVD042201\_genome\_v1. (DOCX)

**S2** Table. Results of the AD assay using different reference genomes. (DOCX)

**S3 Table.** List of primers used in this study. (DOCX)

**S4** Table. Sequences of DNA fragments produced by gene synthesis. (DOCX)

**S5 Table. RNA sequencing datasets used in this study.** (DOCX)

#### **Author Contributions**

Conceptualization: Lukas Kunz, Beat Keller, Marion C. Müller.

**Data curation:** Jigisha Jigisha, Fabrizio Menardo, Alexandros G. Sotiropoulos, Marion C. Müller.

Formal analysis: Lukas Kunz, Alexandros G. Sotiropoulos, Marion C. Müller.

Funding acquisition: Ralph Hückelhoven, Beat Keller.

Investigation: Lukas Kunz, Helen Zbinden.

Methodology: Lukas Kunz, Marion C. Müller.

**Resources:** Jigisha Jigisha, Fabrizio Menardo, Alexandros G. Sotiropoulos, Helen Zbinden, Shenghao Zou, Dingzhong Tang.

- Software: Marion C. Müller.
- Supervision: Fabrizio Menardo, Dingzhong Tang, Ralph Hückelhoven, Beat Keller, Marion C. Müller.

Visualization: Lukas Kunz, Marion C. Müller.

Writing – original draft: Lukas Kunz, Marion C. Müller.

Writing – review & editing: Lukas Kunz, Jigisha Jigisha, Fabrizio Menardo, Alexandros G. Sotiropoulos, Shenghao Zou, Dingzhong Tang, Ralph Hückelhoven, Beat Keller, Marion C. Müller.

#### References

- 1. Figueroa M, Hammond-Kosack KE, Solomon PS. A review of wheat diseases-a field perspective. Molecular Plant Pathology. 2018; 19(6):1523–36. https://doi.org/10.1111/mpp.12618 PMID: 29045052
- Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. The global burden of pathogens and pests on major food crops. Nature Ecology & Evolution. 2019; 3(3):430–9.
- Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nature Reviews Genetics. 2010; 11(8):539–48. https://doi.org/10.1038/nrg2812 PMID: 20585331
- 4. Ngou BPM, Ding PT, Jones JDG. Thirty years of resistance: Zig-zag through the plant immune system. Plant Cell. 2022; 34(5):1447–78. https://doi.org/10.1093/plcell/koac041 PMID: 35167697
- Brown JKM. Durable Resistance of Crops to Disease: A Darwinian Perspective. In: VanAlfen NK, editor. Annual Review of Phytopathology, Vol 53. Annual Review of Phytopathology. 532015. p. 513–39.
- Hafeez AN, Arora S, Ghosh S, Gilbert D, Bowden RL, Wulff BBH. Creation and judicious application of a wheat resistance gene atlas. Molecular Plant. 2021; 14(7):1053–70. https://doi.org/10.1016/j.molp. 2021.05.014 PMID: 33991673
- Minter F, Saunders DG. Safeguarding wheat yields from cereal fungal invaders in the postgenomic era. Current Opinion in Microbiology. 2023;73. https://doi.org/10.1016/j.mib.2023.102310 PMID: 37018996
- Müller MC, Kunz L, Schudel S, Lawson AW, Kammerecker S, Isaksson J, et al. Ancient variation of the AvrPm17 gene in powdery mildew limits the effectiveness of the introgressed rye Pm17 resistance gene in wheat. Proceedings of the National Academy of Sciences of the United States of America. 2022; 119(30):e2108808119–e. https://doi.org/10.1073/pnas.2108808119 PMID: 35857869
- Running KLD, Faris JD. Rapid Cloning of Disease Resistance Genes in Wheat. In: Appels R, Eversole K, Feuillet C, Gallagher D, editors. The Wheat Genome. Cham: Springer International Publishing; 2024. p. 187–212.
- Kusch S, Qian JZ, Loos A, Kümmel F, Spanu PD, Panstruga R. Long-term and rapid evolution in powdery mildew fungi. Mol Ecol. 2024; 33(10). https://doi.org/10.1111/mec.16909 PMID: 36862075
- Beest DET, Paveley ND, Shaw MW, van den Bosch F. Disease-weather relationships for powdery mildew and yellow rust on winter wheat. Phytopathology. 2008; 98(5):609–17. <u>https://doi.org/10.1094/</u> PHYTO-98-5-0609 PMID: 18943230
- 12. McIntosh RA, Dubcovsky J, Rogers WJ, Xia XC, Raupp WJ. Catalogue of gene symbols for wheat: 2019 Supplement Annul Wheat Newsletter. 2019; 65:98–109.
- Hewitt T, Müller MC, Molnár I, Mascher M, Holušová K, Šimková H, et al. A highly differentiated region of wheat chromosome 7AL encodes a *Pm1*a immune receptor that recognises its corresponding *AvrPm1a* effector from *Blumeria graminis*. New Phytologist. 2020; 229(5):2812–26.

- Sánchez-Martín J, Steuernagel B, Ghosh S, Herren G, Hurni S, Adamski N, et al. Rapid gene isolation in barley and wheat by mutant chromosome sequencing. Genome Biology. 2016; 17:221. https://doi. org/10.1186/s13059-016-1082-1 PMID: 27795210
- Yahiaoui N, Srichumpa P, Dudler R, Keller B. Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. Plant Journal. 2004; 37(4):528– 38.
- Xie J, Guo G, Wang Y, Hu T, Wang L, Li J, et al. A rare single nucleotide variant in Pm5e confers powdery mildew resistance in common wheat. New Phytologist. 2020; 228(3):1011–26. https://doi.org/10. 1111/nph.16762 PMID: 32569398
- Hurni S, Brunner S, Buchmann G, Herren G, Jordan T, Krukowski P, et al. Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. Plant Journal. 2013; 76(6):957–69.
- Singh SP, Hurni S, Ruinelli M, Brunner S, Sánchez-Martín J, Krukowski P, et al. Evolutionary divergence of the rye *Pm17* and *Pm8* resistance genes reveals ancient diversity. Plant Molecular Biology. 2018; 98(3):249–60.
- Li M, Dong L, Li B, Wang Z, Xie J, Qiu D, et al. A CNL protein in wild emmer wheat confers powdery mildew resistance. New Phytologist. 2020; 228(3):1027–37. <u>https://doi.org/10.1111/nph.16761</u> PMID: 32583535
- Zou SH, Wang H, Li YW, Kong ZS, Tang DZ. The NB-LRR gene Pm60 confers powdery mildew resistance in wheat. New Phytologist. 2018; 218(1):298–309. <u>https://doi.org/10.1111/nph.14964</u> PMID: 29281751
- Li YH, Wei ZZ, Sela H, Govta L, Klymiuk V, Roychowdhury R, et al. Dissection of a rapidly evolving wheat resistance gene cluster by long-read genome sequencing accelerated the cloning of Pm69. Plant Communications. 2024; 5(1). https://doi.org/10.1016/j.xplc.2023.100646 PMID: 37415333
- Zhao FK, Li YH, Yang BJ, Yuan HB, Jin C, Zhou LX, et al. Powdery mildew disease resistance and marker-assisted screening at the Pm60 locus in wild diploid wheat Triticum urartu. Crop Journal. 2020; 8(2):252–9.
- Kloppe T, Whetten RB, Kim SB, Powell OR, Lück S, Douchkov D, et al. Two pathogen loci determine Blumeria graminis f. sp. tritici virulence to wheat resistance gene Pm1a. New Phytologist. 2023; 238 (4):1546–61. https://doi.org/10.1111/nph.18809 PMID: 36772855
- Bourras S, Kunz L, Xue M, Praz CR, Muller MC, Kalin C, et al. The AvrPm3-Pm3 effector-NLR interactions control both race-specific resistance and host-specificity of cereal mildews on wheat. Nat Commun. 2019; 10(1):2292. https://doi.org/10.1038/s41467-019-10274-1 PMID: 31123263
- Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, Praz CR, et al. Multiple avirulence loci and allele-specific effector recognition control the Pm3 race-specific resistance of wheat to powdery mildew. Plant Cell. 2015; 27(10):2991–3012. https://doi.org/10.1105/tpc.15.00171 PMID: 26452600
- **26.** Praz CR, Bourras S, Zeng FS, Sanchez-Martin J, Menardo F, Xue MF, 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 Phytologist. 2017; 213(3):1301–14.
- Kunz L, Sotiropoulos AG, Graf J, Razavi M, Keller B, Müller MC. The broad use of the Pm8 resistance gene in wheat resulted in hypermutation of the AvrPm8 gene in the powdery mildew pathogen. BMC Biology. 2023; 21(1):29. https://doi.org/10.1186/s12915-023-01513-5 PMID: 36755285
- 28. Bauer S, Yu D, Lawson AW, Saur IML, Frantzeskakis L, Kracher B, et al. The leucine-rich repeats in allelic barley MLA immune receptors define specificity towards sequence-unrelated powdery mildew avirulence effectors with a predicted common RNase-like fold. PLoS Pathog. 2021; 17(2):e1009223. https://doi.org/10.1371/journal.ppat.1009223 PMID: 33534797
- 29. Cao Y, Kümmel F, Logemann E, Gebauer JM, Lawson AW, Yu DL, et al. Structural polymorphisms within a common powdery mildew effector scaffold as a driver of coevolution with cereal immune receptors. Proceedings of the National Academy of Sciences of the United States of America. 2023; 120(32). https://doi.org/10.1073/pnas.2307604120 PMID: 37523523
- Spanu PD. Cereal immunity against powdery mildews targets RNase-Like Proteins associated with Haustoria (RALPH) effectors evolved from a common ancestral gene. New Phytologist. 2017; 213 (3):969–71. https://doi.org/10.1111/nph.14386 PMID: 28079937
- Bernasconi Z, Stirnemann U, Heuberger M, Sotiropoulos AG, Graf J, Wicker T, et al. Mutagenesis of Wheat Powdery Mildew Reveals a Single Gene Controlling Both NLR and Tandem Kinase-Mediated Immunity. Molecular Plant-Microbe Interactions. 2024; 37(3):264–76. https://doi.org/10.1094/MPMI-09-23-0136-FI PMID: 37934013
- Müller MC, Praz CR, Sotiropoulos AG, Menardo F, Kunz L, Schudel S, et al. A chromosome-scale genome assembly reveals a highly dynamic effector repertoire of wheat powdery mildew. New Phytologist. 2019; 221(4):2176–89. https://doi.org/10.1111/nph.15529 PMID: 30388298

- McNally KE, Menardo F, Luthi L, Praz CR, Muller MC, Kunz L, et al. Distinct domains of the AVRPM3 (A2/F2) avirulence protein from wheat powdery mildew are involved in immune receptor recognition and putative effector function. New Phytologist. 2018; 218(2):681–95. <u>https://doi.org/10.1111/nph.15026</u> PMID: 29453934
- Lindner S, Keller B, Singh SP, Hasenkamp Z, Jung E, Muller MC, et al. Single residues in the LRR domain of the wheat PM3A immune receptor can control the strength and the spectrum of the immune response. Plant Journal. 2020; 104:200–14. https://doi.org/10.1111/tpj.14917 PMID: 32645755
- Seong K, Krasileva KV. Prediction of effector protein structures from fungal phytopathogens enables evolutionary analyses. Nat Microbiol. 2023; 8(1):174–87. https://doi.org/10.1038/s41564-022-01287-6 PMID: 36604508
- Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024; 630(8016):493–500. https://doi.org/10.1038/ s41586-024-07487-w PMID: 38718835
- Pennington HG, Jones R, Kwon S, Bonciani G, Thieron H, Chandler T, et al. The fungal ribonucleaselike effector protein CSEP0064/BEC1054 represses plant immunity and interferes with degradation of host ribosomal RNA. PLoS Pathog. 2019; 15(3):e1007620. <u>https://doi.org/10.1371/journal.ppat.</u> 1007620 PMID: 30856238
- Menardo F, Praz CR, Wyder S, Ben-David R, Bourras S, Matsumae H, et al. Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. Nature Genetics. 2016; 48 (2):201–5. https://doi.org/10.1038/ng.3485 PMID: 26752267
- Sotiropoulos AG, Arango-Isaza E, Ban T, Barbieri C, Bourras S, Cowger C, et al. Global genomic analyses of wheat powdery mildew reveal association of pathogen spread with historical human migration and trade. Nat Commun. 2022; 13(1):4315. https://doi.org/10.1038/s41467-022-31975-0 PMID: 35882860
- Müller MC, Kunz L, Graf J, Schudel S, Keller B. Host adaptation through hybridization: Genome analysis of triticale powdery mildew reveals unique combination of lineage-specific effectors. bioRxiv. 2021:2021.05.06.442769. https://doi.org/10.1094/MPMI-05-21-0111-SC PMID: 34503345
- Zou SH, Shi WQ, Ji JH, Wang HM, Tang YS, Yu DZ, et al. Diversity and similarity of wheat powdery mildew resistance among three allelic functional genes at the Pm60 locus. Plant Journal. 2022; 110 (6):1781–90. https://doi.org/10.1111/tpj.15771 PMID: 35411560
- Huang L, Li YH, Chen SS, Periyannan S, Fahima T. Editorial: Advances in crop resistance breeding using modern genomic tools. Frontiers in Plant Science. 2023; 14. <u>https://doi.org/10.3389/fpls.2023.</u> 1143689 PMID: 36875621
- Sánchez-Martín J, Keller B. NLR immune receptors and diverse types of non-NLR proteins control race-specific resistance in Triticeae. Current Opinion in Plant Biology. 2021; 62:102053. <u>https://doi.org/ 10.1016/j.pbi.2021.102053</u> PMID: 34052730
- Wulff BB, Krattinger S. The long road to engineering durable disease resistance in wheat. Current Opinion in Biotechnology. 2022; 73:270–5. https://doi.org/10.1016/j.copbio.2021.09.002 PMID: 34563932
- 45. Lu X, Kracher B, Saur IML, Bauer S, Ellwood SR, Wise R, et al. Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. Proceedings of the National Academy of Sciences of the United States of America. 2016; 113(42):E6486–E95. <u>https://doi.org/10.1073/pnas.1612947113 PMID: 27702901</u>
- 46. Saur I, Bauer S, Kracher B, Lu XL, Franzeskakis L, Muller MC, et al. Multiple pairs of allelic MLA immune receptor-powdery mildew AVR(A) effectors argue for a direct recognition mechanism. Elife. 2019; 8:e44471. https://doi.org/10.7554/eLife.44471 PMID: 30777147
- Pedersen C, van Themaat EVL, McGuffin LJ, Abbott JC, Burgis TA, Barton G, et al. Structure and evolution of barley powdery mildew effector candidates. BMC Genomics. 2012; 13:694. <u>https://doi.org/10.1186/1471-2164-13-694 PMID: 23231440</u>
- Avni R, Lux T, Minz-Dub A, Millet E, Sela H, Distelfeld A, et al. Genome sequences of three Aegilops species of the section Sitopsis reveal phylogenetic relationships and provide resources for wheat improvement. Plant Journal. 2022; 110(1):179–92. https://doi.org/10.1111/tpj.15664 PMID: 34997796
- **49.** Bourras S, Praz CR, Spanu PD, Keller B. Cereal powdery mildew effectors: a complex toolbox for an obligate pathogen. Current Opinion Microbiology. 2018; 15(46):26–33. https://doi.org/10.1016/j.mib. 2018.01.018 PMID: 29455142
- Stirnweis D, Milani SD, Jordan T, Keller B, Brunner S. Substitutions of Two Amino Acids in the Nucleotide-Binding Site Domain of a Resistance Protein Enhance the Hypersensitive Response and Enlarge the PM3F Resistance Spectrum in Wheat. Molecular Plant-Microbe Interactions. 2014; 27(3):265–76. https://doi.org/10.1094/MPMI-10-13-0297-FI PMID: 24329172
- 51. Li YH, Wei ZZ, Fatiukha A, Jaiwar S, Wang HC, Hasan S, et al. TdPm60 identified in wild emmer wheat is an ortholog of Pm60 and constitutes a strong candidate for PmG16 powdery mildew resistance.

Theoretical and Applied Genetics. 2021; 134(9):2777–93. https://doi.org/10.1007/s00122-021-03858-3 PMID: 34104998

- Wu QH, Chen YX, Li BB, Li J, Zhang PP, Xie JZ, et al. Functional characterization of powdery mildew resistance gene MIIW172, a new Pm60 allele and its allelic variation in wild emmer wheat. Journal of Genetics and Genomics. 2022; 49(8):787–95. <u>https://doi.org/10.1016/j.jgg.2022.01.010</u> PMID: 35167980
- Wu QH, Zhao F, Chen YX, Zhang PP, Zhang HZ, Guo GH, et al. Bulked segregant CGT-Seq-facilitated map-based cloning of a powdery mildew resistance gene originating from wild emmer wheat (Triticum dicoccoides). Plant Biotechnology Journal. 2021; 19(7):1288–90. https://doi.org/10.1111/pbi.13609 PMID: 33949081
- Zeng F-s, Yang L-j, Gong S-j, Zhang X-j, Wang H, Xiang L-b, et al. Virulence and diversity of *Blumeria* graminis f. sp. tritici populations in China. J Integr Agr. 2014; 13(11):2424–37.
- 55. Parlange F, Oberhaensli S, Breen J, Platzer M, Taudien S, Simkova H, et al. A major invasion of transposable elements accounts for the large size of the Blumeria graminis f.sp. tritici genome. Funct Integr Genomics. 2011; 11(4):671–7. https://doi.org/10.1007/s10142-011-0240-5 PMID: 21809124
- 56. Parlange F, Roffler S, Menardo F, Ben-David R, Bourras S, McNally KE, 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. 2015; 82:181–92.
- Nurk S, Walenz BP, Rhie A, Vollger MR, Logsdon GA, Grothe R, et al. HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. Genome Research. 2020; 30(9):1291–305. https://doi.org/10.1101/gr.263566.120 PMID: 32801147
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology. 2019; 37(5):540-+. https://doi.org/10.1038/s41587-019-0072-8 PMID: 30936562
- Cheng HY, Concepcion GT, Feng XW, Zhang HW, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nature Methods. 2021; 18(2):170–+. https://doi.org/10.1038/ s41592-020-01056-5 PMID: 33526886
- Marcais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: A fast and versatile genome alignment system. PLoS Comput Biol. 2018; 14(1):e1005944. <u>https://doi.org/10.1371/journal.pcbi.1005944</u> PMID: 29373581
- 61. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST plus: architecture and applications. Bmc Bioinformatics. 2009; 10:421.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. Plos One. 2014; 9 (11):e112963. https://doi.org/10.1371/journal.pone.0112963 PMID: 25409509
- Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, et al. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Research. 2008; 18(1):188–96. https://doi.org/10.1101/gr.6743907 PMID: 18025269
- Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012: 1207.3907v2.
- Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods. 2011; 8(10):785–6. <u>https://doi.org/10.1038/nmeth.1701</u> PMID: 21959131
- **66.** Himmelbach A, Zierold U, Hensel G, al. e. A set of modular binary vectors for transformation of cereals. Plant Physiology. 2007; 145(4):1192–200.
- Weigel D, Glazebrook J. Transformation of agrobacterium using the freeze-thaw method. CSH protocols. 2006; 2006(7). https://doi.org/10.1101/pdb.prot4666 PMID: 22484682
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature Methods. 2012; 9(7):676–82. <u>https://doi.org/10.1038/nmeth.2019</u> PMID: 22743772
- **69.** Praz CR, Menardo F, Robinson MD, Muller MC, Wicker T, Bourras S, 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. 2018; 9:49. https://doi.org/10.3389/fpls. 2018.00049 PMID: 29441081
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods. 2017; 14(4):417–9. https://doi.org/10.1038/nmeth.4197 PMID: 28263959

- 71. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26(1):139–40. https://doi.org/10.1093/bioinformatics/btp616 PMID: 19910308
- 72. RStudio-Team. RStudio: integrated development environment for R. http://www.rstudio.com/; 2018.
- 73. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 2009:1–212.
- 74. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, et al. UCSF ChimeraX: Tools for structure building and analysis. Protein Science. 2023; 32(11). <u>https://doi.org/10.1002/pro.4792</u> PMID: 37774136