

Article

# Discovery of Resistance Genes in Rye by Targeted Long-Read Sequencing and Association Genetics

Nikolaj M. Vendelbo <sup>1,2,\*</sup> , Khalid Mahmood <sup>1</sup> , Burkhard Steuernagel <sup>3</sup>, Brande B. H. Wulff <sup>3,4,5</sup> , Pernille Sarup <sup>1</sup> , Mogens S. Hovmøller <sup>2</sup>, Annemarie Fejer Justesen <sup>2</sup>, Peter S. Kristensen <sup>1</sup> , Jihad Orabi <sup>1</sup>  and Ahmed Jahoor <sup>1,6</sup>

- <sup>1</sup> Department of Molecular Breeding, Nordic Seed A/S, 8300 Odder, Denmark; khma@nordicseed.com (K.M.); pesa@nordicseed.com (P.S.); pskr@nordicseed.com (P.S.K.); jior@nordicseed.com (J.O.); ahja@nordicseed.com (A.J.)
- <sup>2</sup> Department of Agroecology, Faculty of Technology, Aarhus University, 4200 Slagelse, Denmark; mogens.hovmoller@agro.au.dk (M.S.H.); annemariefejer.justesen@agro.au.dk (A.F.J.)
- <sup>3</sup> John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK; burkhard.steuernagel@jic.ac.uk (B.S.); brande.wulff@kaust.edu.sa (B.B.H.W.)
- <sup>4</sup> Plant Science Program, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia
- <sup>5</sup> Center for Desert Agriculture, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia
- <sup>6</sup> Department of Plant Breeding, The Swedish University of Agricultural Sciences, SE-234 22 Alnarp, Sweden
- \* Correspondence: nive@nordicseed.com; Tel.: +45-2648-6644

**Abstract:** The majority of released rye cultivars are susceptible to leaf rust because of a low level of resistance in the predominant hybrid rye-breeding gene pools Petkus and Carsten. To discover new sources of leaf rust resistance, we phenotyped a diverse panel of inbred lines from the less prevalent Gülzow germplasm using six distinct isolates of *Puccinia recondita* f. sp. *secalis* and found that 55 out of 92 lines were resistant to all isolates. By performing a genome-wide association study using 261,406 informative SNP markers, we identified five resistance-associated QTLs on chromosome arms 1RS, 1RL, 2RL, 5RL and 7RS. To identify candidate *Puccinia recondita* (*Pr*) resistance genes in these QTLs, we sequenced the rye nucleotide-binding leucine-rich repeat (NLR) intracellular immune receptor complement using a Triticeae NLR bait-library and PacBio® long-read single-molecule high-fidelity (HiFi) sequencing. Trait-genotype correlations across 10 resistant and 10 susceptible lines identified four candidate NLR-encoding *Pr* genes. One of these physically co-localized with molecular markers delimiting *Pr3* on chromosome arm 1RS and the top-most resistance-associated QTL in the panel.

**Keywords:** *Secale cereale* L.; leaf rust; brown rust; *Puccinia recondita* f. sp. *secalis*; nucleotide-binding leucine-rich repeat (NLR); resistance gene enrichment sequencing (RenSeq); single-molecule high-fidelity sequencing (HiFi); genome-wide association study (GWAS); *k*-mer



**Citation:** Vendelbo, N.M.; Mahmood, K.; Steuernagel, B.; Wulff, B.B.H.; Sarup, P.; Hovmøller, M.S.; Justesen, A.F.; Kristensen, P.S.; Orabi, J.; Jahoor, A. Discovery of Resistance Genes in Rye by Targeted Long-Read Sequencing and Association Genetics. *Cells* **2022**, *11*, 1273. <https://doi.org/10.3390/cells11081273>

Academic Editor: Laurens Pauwels

Received: 14 March 2022

Accepted: 6 April 2022

Published: 9 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Leaf rust is one of the most devastating diseases in hybrid rye (*Secale cereale* L.) as it can reduce grain yield by up to 27% in the field [1]. Because the predominant heterotic gene pools used for the breeding of rye hybrids, ‘Petkus’ and ‘Carsten’, suffer from low inherent resistance, most newly released cultivars are highly susceptible to leaf rust (Supplementary Table S1) [2,3]. In rye, leaf rust is caused by the heteroecious macrocyclic fungal basidiomycete, *Puccinia recondita* f. sp. *secalis* (*Prs*), which is prolific across the rye growing regions in Europe [4]. Unlike wheat leaf rust, stripe rust and stem rust, the alternate plant host of *Prs*, small bugloss (*Anchusa arvensis* L.), is widespread in Northern European flora and a common weed in agricultural fields [5–8]. The evolutionary plasticity generated by potential frequent sexual recombination and the capacity of spores to

migrate long distances accentuate the risk of novel, aggressive *Prs* pathotypes emerging and spreading, which in turn constitutes a serious threat to hybrid rye production [4,9–11].

Host resistance represents a sustainable and more environmentally conscious alternative to chemical control strategies [12]. Currently, six major leaf rust resistance (*R*) genes have been identified in rye: *Pr3* (on the short arm of chromosome 1, 1RS), *Prs4* (on the long arm of chromosome 1, 1RL), *Pr5* (1RL), *Pr1* (6RL), *PrNOS1* (7RS) and *Pr2* (7RL) [13–15]. In addition, three major leaf rust *R* genes have been described in translocation lines between rye and wheat (*Triticum aestivum* L.), with rye as the resistance donor and designated according to the wheat gene nomenclature as *Leaf rust resistance 26* (*Lr26*, from translocation 1BL-1RS), *Lr25* (4BS.4BL—2RL) and *Lr45* (2AS—2RS.2RL) [16,17]. The majority of *R* genes encode intracellular nucleotide-binding leucine-rich repeat (NLR) immune receptor proteins [18]. In the grasses, canonical NLR genes are comprised of three domains, (i) a C-terminal leucine-rich repeat (LRR) domain for pathogen effector recognition and NLR autoinhibition [19,20]; (ii) a conserved nucleotide-binding (NB) domain in the middle of the protein that is involved in regulation of NLR activation [21]; and (iii), an N-terminal coiled-coil (CC) domain mediating the downstream immune-signaling cascade [22]. In the reference genome of the German inbred grain-type rye line Lo7, 1167 NLR genes have been reported, with this number reaching 1447 in the Chinese forage-type Weining [23,24].

The breeding of novel, resistant cultivars has been accelerated by the adoption of genomic and molecular breeding techniques. The use of marker-assisted selection (MAS) allows the pyramiding of *R* genes and intermediate-effect quantitative trait loci (QTLs) for enhancing the durability of resistance [25–27]. When coupled with speed-breeding approaches, MAS enables the rapid development of novel cultivars by the introgression of exotic *R* genes into an elite background [28,29]. Importantly, mining for novel genetic variation for resistance against a given pathogen is a prerequisite for the continuous development of novel resistant cultivars. Next-generation mapping resources such as nested association mapping (NAM) and multiparent advanced generation intercross (MAGIC) populations, in combination with high-density single-nucleotide polymorphism (SNP) genotyping, have provided powerful techniques for the identification of *R* genes and QTLs [30,31]. Historical recombination events naturally present within a diverse panel for a crop species can also be exploited for mining *R* genes and QTLs through genome-wide association studies (GWAS) and SNP genotype data [32]. GWAS identify a genomic region controlling a trait of interest and, predominantly, a nonfunctional marker associated with the trait that can be readily implemented in MAS [33]. However, as the markers often reside outside the actual genes of interests, how informative they are can be compromised by crossover events. Furthermore, the identification of the causal gene is often not possible in cases when the *R* gene is absent from the reference genome or exhibits considerable divergence from the reference sequence [34].

The limitations of GWAS can be overcome by performing trait association on sub-sequences (*k*-mers) derived from *R* gene enrichment sequencing (AgRenSeq). To obtain trait-genotype associations in cis, the associated *k*-mers are mapped to the contigs of a de novo assembly of a resistant accession [35]. AgRenSeq makes it possible to mine for novel *R* genes in undomesticated wild relatives of crop species without the need for a high-quality reference genome. Instead of identifying a resistance-associated genomic region, RenSeq identifies candidate NLR genes conferring resistance, from which functional markers can be developed for MAS and as a resource for genetic engineering [36]. However, the de novo assembly of NLR genes from short read data is a complex task due to their high copy number and sequence similarity between paralogs [37]. This limitation was previously alleviated by coupling RenSeq with long-read PacBio® single-molecule real-time (SMRT) sequencing [38,39]. Since then, improvements to the PacBio® HiFi sequel II system have demonstrated 99.9% accuracy in the recall rate of single-nucleotide variants (SNVs) and 96% accuracy for insertions and deletions (InDels) [40].

Currently, RenSeq has been deployed in *Arabidopsis* (*Arabidopsis thaliana*), strawberry (*Fragaria × ananassa*), potato (*Solanum tuberosum* L.), tomato (*Solanum lycopersicum* L.)

hexaploid bread wheat, an introgression line between wheat and mosquito grass (*Dasypyrum villosum* L.) and the wheat wild relatives Tausch's goatgrass (*Aegilops tauschii*) and goat grass (*Ae. peregrina*) [35,38,39,41–44].

The objective of this study was to determine the efficacy of SMRT AgRenSeq for the mining of leaf rust *R* genes in rye using a panel of restorer lines recently demonstrated to exhibit a high level of resistance in the field [15].

## 2. Materials and Methods

### 2.1. Collection of Danish *Puccinia Recondita* f. sp. *Secalis* Samples, Their Multiplication and Single Pustule Isolation

With no molecular markers or differential sets available for the identification of races in *Puccinia recondita* f. sp. *secalis* (*Prs*), a national sampling of the pathogen was conducted at six geographically distinct sites in Denmark from 2018 to 2019 (Supplementary Table S2, Supplementary Figure S1). Multiplication of field samples and isolation of single-pustule isolates (SPI) were performed as described by Vendelbo et al. [15].

### 2.2. Plant Materials and DNA Extraction

In total, 92 inbred rye (*Secale cereale* L.) restorer lines belonging to the elite hybrid rye-breeding germplasm at Nordic Seed A/S (Dyngby, Denmark) were selected for the study. Four seeds per line were sown per well in 104-well trays containing a fine-grained sphagnum substrate. The seedlings were propagated in the greenhouse facilities at Nordic Seed A/S under 16 h of daylight at 18–24 °C and 8 h of darkness at 14–16 °C. The lowest sections of two coleoptiles and primary leaves were excised after 7 days (equivalent to 75 mg of plant material) and placed in a 96-well Micro-Dilution Tube System (STARLAB International GmbH) containing two 4 mm glass beads per 1.2 mL tube. Plant tissue samples were stored at –20 °C for 2 days prior to freeze-drying for an additional 2 d. DNA extraction was performed using an adapted SDS-based method according to USDA [45] according to Pallotta et al. [46]. The DNA concentration and 260/280 nm absorption ratio were measured using an Epoch™ microplate spectrophotometer (Biotek® Instruments, Winooski, V.T., USA). Fragmentation of genomic DNA was assessed by size separation on a 1.2% (*w/v*) agarose gel.

### 2.3. Phenotyping for Pathogen Resistance

All lines were phenotyped for their resistance response against six geographically distinct *Prs* SPIs to investigate the level of resistance to leaf rust in the panel. Phenotyping was conducted using a 10-step (0–9) infection-type (IT) scale adapted by McNeal et al. [47] and Hovmoller et al. [48]. Seedlings with a mean IT of 0–2 were categorized as ‘resistant’, IT of 3–4 as ‘partially resistant’, IT of 5–6 as ‘partially susceptible’ and IT of 7–9 as ‘susceptible’. Large-scale multiplication of trial inoculum was done in spore-proof greenhouse cabins at Nordic Seed A/S (Dyngby, Denmark) in accordance with the protocol described by Vendelbo et al. [15]. In the trial, eight seeds per line were sown in a 28-hole tray containing a coarse-grained sphagnum substrate. The trial followed a partially randomized design with two repetitions for each of the two trial replicates, hence producing four scorings per line per SPI evaluated.

At 14 days after sowing, each tray was inoculated with 30 mg of SPI spores solubilized in 4 mL 3M™ Novex™ 7100 (Sigma-Aldrich®, St. Louis, MO, USA) engineering fluid according to the method described by Thach et al. [49]. After incubation for 14 d, the lines were phenotyped for their infection-type response by scoring both the first and second leaf of each of the eight seedlings per pot.

### 2.4. Molecular Marker Resource and SNP Genotyping

For each line, 200 ng of high molecular-weight genomic DNA with a 260/280 nm absorbance ratio of  $\geq 1.8$  were sent for SNP genotyping at Eurofins Genomics Europe Genotyping (Aarhus, Denmark). Genotyping was accomplished using a 600 K SNP array

with 600,843 SNP markers on an Affymetrix GeneTitan™ Scanner platform (Thermo Fisher Scientific, Waltham, MA, USA) [50]. Mapping of the SNP markers of the Lo7 rye reference genome and characterization of its performance in the assayed germplasm was recently investigated by [51]. The marker map is available at <https://doi.org/10.5281/ZENODO.5086235> (accessed on 2 January 2022).

### 2.5. Data Analysis

Bioinformatics analysis was performed with the R Studio (v. 1.3.959) interface with R statistical software (v. 4.0.1) with various predesigned packages [52,53]. Prior to analysis of the SNP genotyping data, markers were filtered for marker allele frequencies of at least 5%, missing individual scores below 20% and missing marker scores below 10% to identify informative markers. Computationally demanding tasks were run at the national, high-performance computing facility GenomeDK at the Aarhus Genome Data Center, Aarhus University. Custom R scripts for all visual outputs have been provided as supporting information.

### 2.6. Genome-Wide Association Study

The exploration of SNPs associated with leaf rust resistance relied on a genome-wide association study (GWAS) using the genomic-association and prediction-integration-tool (GAPIT v. 3) package in R [54]. The phenotypic input for GWAS included all individual recordings for the six SPI *Prs* scorings, averaged across the two replicates.

### 2.7. In Silico Test of the Bait Library

Before using the 600 K Triticeae nucleotide-binding leucine-rich repeat (NLR) bait-library (Tv\_1) designed by Steuernagel et al. [42], an in silico test was run to evaluate its capture success with rye. Baits were mapped to the rye reference-genome Lo7 NLR repertoire using the Basic Local Alignment Search Tool for nucleotide (BLASTN v. 2.9.0+) function at the National Center for Biotechnological Information (NCBI) [55]. Bait hits were then filtered for a minimum of 80% sequence identity over the full 120-nt bait length, according to the findings made by Jupe et al. [37], who reported that ~80% sequence identification between bait and NLR sequence is sufficient for enrichment. The Tv\_1 bait library is available at <https://github.com/steuernb/MutantHunter/> (accessed on 7 August 2021). The bait library was assessed using custom scripts and inherent functions in R for NLR sequence capture success, unique baits per NLR sequence, unique NLR sequences per bait, distribution of bait alignment (in bp) and alignment identity (in %).

### 2.8. Phylogenetic Analysis and Pairwise Selection of Restorer Lines

To reduce the false association from noncausative background, 10 pairs of lines were selected for Single-Molecule Real Time R gene enrichment and Sequencing (SMRT RenSeq), each comprising two phylogenetically closely related lines but exhibiting a divergent resistance profile. Phylogenetic analysis was performed with a neighbor-joining clustering of lines using measures of their Euclidean genetic distance using the ape (v. 5.3) R package [56]. The tree was constructed after 10,000 bootstrapping iterations with weak nodes showing less than 80% recurrence collapsed into multifurcations. The tree was drawn with iTOL (v. 6.1.1) (<http://itol.embl.de/> (accessed on 7 August 2021)), allowing a color-based visualization of the resistance response presented by each line against the six *Prs* SPIs as concentric circles [57].

### 2.9. Single-Molecule Real-Time High-Fidelity Resistance Gene Enrichment Sequencing

The construction of PacBio® RenSeq libraries was outsourced to Arbor Biosciences (M.I., USA). A minimum of 5 µg per sample of genomic DNA demonstrating a UV 260/280 absorbance ratio between 1.7 and 1.9 and of at least 10 kb modal length was provided. Genomic DNA libraries were enriched for NLR sequences using the Tv\_1 60 K Triticeae NLR bait-library [42]. The enriched target DNA was sequenced on a PacBio® Sequell II long-read

SMRT platform according to PacBio® methods without modifications to generate 1.5 Gb of high-fidelity (HiFi) circular consensus sequence (CCS) per line. Prior to sequencing, DNA libraries were multiplexed to allow pooling.

#### 2.10. De Novo Assembly and NLR Annotation

The CCSs were assembled with HiCanu (v. 2.0; -pacbio-hifi, trimReadsCoverage = 2, errorRate = 0.01, genomeSize = 8.8 m, minOverlapLength = 500, minReadLength = 1000) [58]. The expected genome size of the rye NLR repertoire was estimated using NLR data from the recent rye reference genome Lo7 [23] with the addition of a 40% flanking region in accordance with the study by Van de Weyer et al. [41]. Contigs associated with NLR genes were annotated using NLR-Parser (v. 3) and NLR-Annotator (<https://github.com/steuernb/NLR-Annotator> (accessed on 7 August 2021)) [59,60]. NLR contigs were mapped to the rye reference genome Lo7 using NCBI BLASTN (v. 2.9.0+) with a significance threshold set to  $1 \times 10^{-5}$ , selecting the physical position of the top hit [55]. Contig positions were kept if they met the criteria of  $\geq 90\%$  alignment length and  $\geq 85\%$  alignment identity.

#### 2.11. In Vitro Test of the Bait Library

To investigate Tv\_1 bait-library performance in vitro, the individual restorer line SMRT RenSeq data were subjected to a similar analysis to that described for the in silico test. In addition, to assess the bait distribution amongst lines, baits were classified into three categories: (i) 'core', aligning in more than 12 lines; (ii) 'common', aligning in 5–12 lines; and (iii) 'rare', aligning in fewer than 5 lines. The distribution of bait alignment, bait alignment identity and bait category distribution were visualized using the ggplot2 (v. 3.3.3) R package [61]. The distributions of baits per NLR contig and NLR contigs per bait were visualized using the YaRrr R (v. 0.1.5) package. The R scripts for the in silico and in vitro assessment of the 60 K bait library on the panel are provided as supporting information.

#### 2.12. K-Mer Presence/Absence Matrix

Raw CCS reads were processed into  $k$ -mers ( $k = 51$  nt) using Jellyfish (v. 2.2.10), discarding rare  $k$ -mers with a count  $\leq 10$  per line [62]. A collective  $k$ -mer presence/absence matrix was generated by pooling  $k$ -mers from all lines in a binary format with 1 ('present') and 0 ('absent'). The matrix was filtered for rare  $k$ -mers present in  $\leq 3$  lines.  $k$ -mer processing of the CCS reads was performed following the Java source code published at <https://github.com/steuernb/AgRenSeq> (accessed on 7 August 2021).

#### 2.13. Association Genetics RenSeq Analysis

The phenotypic data were converted using the following formula developed for phenotypic scores in Stackman's IT scale (−2 to 2).

$$\text{Stackman's IT score} = 2 - \left( \frac{4}{9} * \text{Severity Score}_{0-9} \right) \quad (1)$$

Identification of candidate leaf rust resistance genes amongst the NLR-annotated contigs was performed by AgRenSeq analysis following the Java pipeline provided by Arora et al. [35] at <https://github.com/steuernb/AgRenSeq> (accessed on 7 August 2021) and the general linear regression model (AgRenSeq-GLM) using the Python module provided at [https://github.com/kgaurav1208/AgRenSeq\\_GLM](https://github.com/kgaurav1208/AgRenSeq_GLM) (accessed on 19 August 2021). This modulated analysis utilizes a linear regression model for each of the  $k$ -mers to account for population structure amongst lines and to likelihood-test for nested models to output a  $p$ -value for each of the  $k$ -mers' association with resistance. A standard Bonferroni-corrected threshold of  $\alpha = 0.025$  was used as the significance threshold.

#### 2.14. Characterization of the Candidate Leaf Rust Resistance Gene *Pr3*

To improve NLR annotation, a manual reference assembly was generated for all leaf-rust-resistance-associated contigs identified by AgRenSeq to expand the contig size. For each contig, raw CCS mapping to the contig was performed using NCBI BLASTN (v. 2.9.0+) function, selecting CCS aligning over at least 2000 bp and with 95% sequence identity [55]. CCS reads were trimmed for adaptor sequences and assembled using the respective leaf-rust-resistance-associated contig as reference and manually inspected to generate a consensus sequence in Geneious Prime (v. 2020.2.3) [63]. The gene structure of the manually assembled consensus NLR contigs was predicted using the AUGUSTUS (3.4.0) program [64]. The prediction of protein domain structures was performed using InterPro-Scan [65]. Prediction of LRR motifs was performed using LRRpredictor (v. 1.0) [66]. Coding sequences were mapped to the Lo7 and Weining reference genomes using NCBI BLASTN (v. 2.9.0+) [23,24,55].

To distinguish unique NLR genes from homologs within the panel of leaf-rust-resistance-associated contigs identified by AgRenSeq, a phylogenetic analysis was undertaken on the basis of protein sequence similarity using the pipeline developed by Toparslan et al. [67] in R. Multiple sequence alignment of protein sequences was performed using 'Clustal Omega' in the *msa* (v. 1.20) R package. A neighbor-joining tree was constructed using Nei's standard genetic distance and visualized using the *ggtree* (v. 2.2.4) R package [68]. To validate the tree, 10,000 bootstrapping iterations were run. As reference, sequences of known NLR genes conferring rust resistance in cereals were included in the analysis, obtained from the UniProt and NCBI online databases [55,69].

Identification of single-nucleotide variants (SNVs) and insertions/deletions (InDels) between candidate gene variants was performed by multiple sequence alignment using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) method in Geneious Prime (v. 2020.2.3) with  $\leq 10$  iterations. For comparative analysis, NLR genes at mapping positions in the Lo7 and Weining reference genomes were extracted from the NLR annotation file provided by [51]. Annotation files are available at <https://doi.org/10.5281/zenodo.5085854> (accessed on 10 November 2021). If present in the reference genomes ( $\geq 95\%$  sequence similarity), coding sequences and de novo, predicted protein sequences were extracted. To investigate whether the candidate *Pr* gene was positioned near designated *Pr* genes on chromosome arm 1RS, flanking co-segregating markers were extracted and mapped to the Lo7 and Weining reference genomes by BLASTN [14,70].

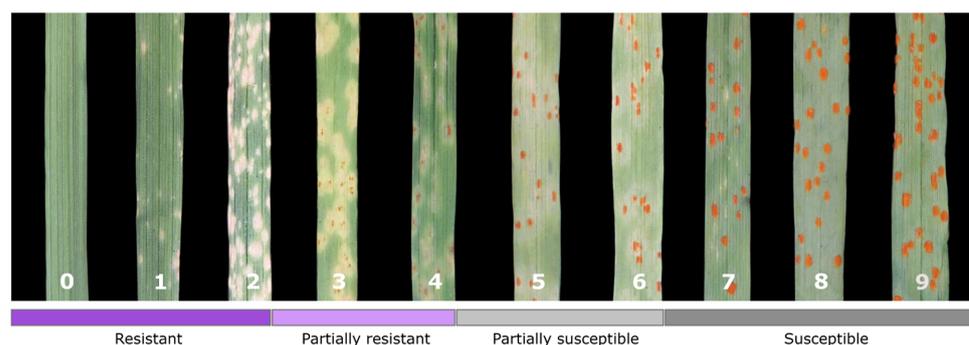
#### 2.15. Graphical Editing

Graphs and Figures were saved from R in .svg format and manually curated using the Inkscape (v. 1.1) program (<https://inkscape.org/> (accessed on 7 August 2021)).

### 3. Results

#### 3.1. Phenotyping of Rye Breeding Lines for Resistance to Leaf Rust

We tested 92 inbred rye restorer breeding lines for leaf rust resistance using six *Puccinia recondita* f. sp. *secalis* (*Prs*) single-pustule isolates (SPIs) of distinct geographical origin, which uncovered a high level of resistance in the germplasm (Supplementary Table S2, Supplementary Figure S1). We categorized 51 lines as consistently resistant (IT 0–2), 11 as partially resistant (IT 3–4), 11 as partially susceptible (IT 5–6) and the remaining 19 as consistently susceptible (IT 7–9) (Supplementary Table S3). Four lines showed an SPI-specific resistance response. Figure 1 illustrates the 10 infection types as a resource for future studies of leaf rust resistance in rye.



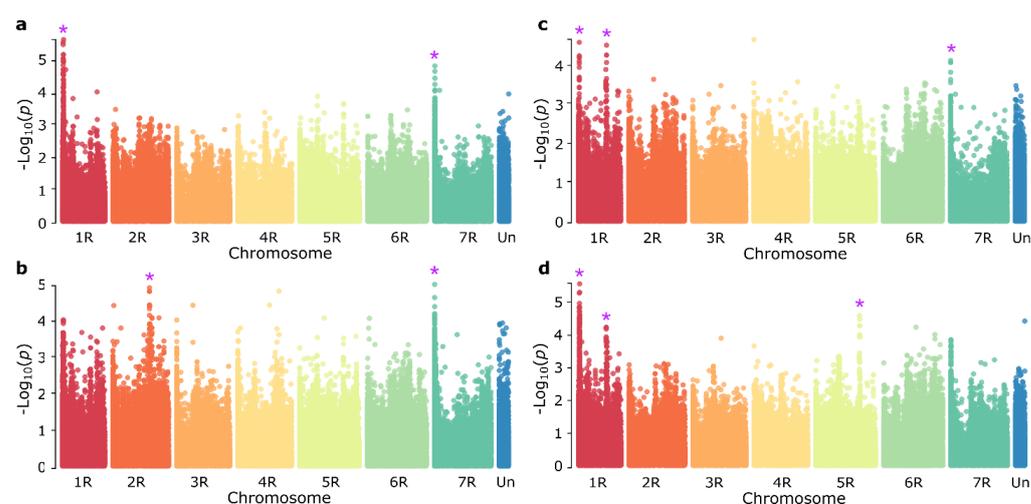
**Figure 1.** Infection-type (IT) response scale for leaf rust in rye (*Secale cereale* L.) caused by the fungal pathogen *Puccinia recondita* f. sp. *secalis* according to Hovmøller et al. (2017) and McNeal et al. (1971). In terms of virulence/avirulence, IT 0–6 are considered ‘avirulent’ and 7–9 as ‘virulent’.

### 3.2. Genome-Wide Association Study

Quality filtering of markers from the 600 K SNP array allowed the identification of 261,406 informative markers (Data File S1). To explore the genetic basis underlying leaf rust resistance in the panel, we performed a genome-wide association study (GWAS) across the six *Prs* SPIs (Supplementary Figure S2, Supplementary Table S4). We identified five genomic positions associated with leaf rust resistance against one or more SPI that mapped to chromosome arms 1RS, 1RL, 2RL, 5RL and 7RS (Table 1, Figure 2). However, none of the markers exhibited a  $p$ -value above the Bonferroni significance threshold of  $-\log_{10}(p) = 6.72$ . Each marker explained between 9.3 and 13.1% of the phenotypic variance (Table 1).

**Table 1.** Genetic markers associated with leaf rust resistance identified from 92 hybrid rye restorer breeding lines by association genetics using 261,406 SNP array markers.

Chromosome	Position (Mb)	Marker ID	$-\log_{10}(p)$	Phenotypic Variance Explained (%)
1RS	115.55	AX-99251803	6.11	13.1
1RL	625.54	AX-99805135	4.50	9.3
2RL	818.90	AX-99478491	4.90	9.6
5RL	770.21	AX-99776626	4.49	9.3
7RS	26.93	AX-99684185	4.83	12.1



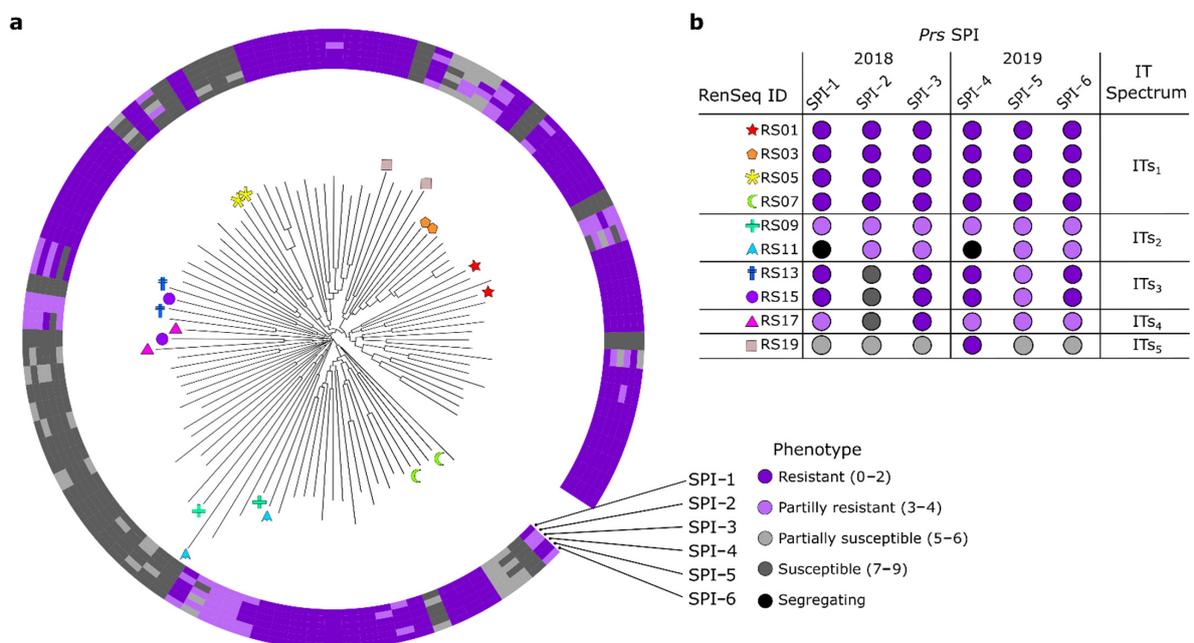
**Figure 2.** Manhattan plots for genome-wide association study (GWAS) for leaf rust resistance in 92 inbred rye lines. GWAS was performed using 261,406 informative SNP markers mapped to the Lo7 reference genome. The entire germplasm collection was phenotyped for resistance against six distinct *Puccinia recondita* f. sp. *secalis* single-pustule isolates (SPIs) in a greenhouse in two replicate trials. (a) SPI-5, (b) SPI-3, (c) SPI-1 and (d) SPI-2. Five genomic regions associated with resistance are marked by magenta asterisks. The Bonferroni-adjusted significance threshold based on informative markers was 6.72.

### 3.3. In Silico Test of 600 K Triticeae NLR Bait-Library for Rye

Prior to RenSeq analysis, we assessed the performance of the 600 K Triticeae-specific NLR bait-library Tv\_1 in rye by performing an in silico analysis of its capture success rate against the NLR repertoire of the Lo7 reference genome. Using a set threshold, we determined that 29,785 baits align, capturing 1125 out of the 1167 annotated NLR sequences in the Lo7 reference (Supplementary Table S7). Each bait mapped, on average, to  $5 \pm 6$  (standard deviation, SD) unique NLR sequences, with a mean of  $140 \pm 164$  (SD) unique baits mapped per NLR sequence (Supplementary Table S5).

### 3.4. Phylogenetic Analysis and Pairwise Selection of Restorer Lines

To identify *Pr* genes in the panel, we reasoned that selecting pairs of lines that are closely related but differ only in their resistance profile would maximize the discriminatory power of informative polymorphisms without having to resort to sequencing the entire panel (Figure 3a). The corresponding shortlisted resistant lines exhibited five different infection-type (ITs) spectra: resistant (IT<sub>s1</sub>), partially resistant (IT<sub>s2</sub>) and three different SPI-specific spectra (IT<sub>s3-5</sub>) (Figure 3b).



**Figure 3.** Phylogenetic analysis and type of leaf rust infection response in 92 inbred rye lines. (a) Circular neighbor-joining dendrogram indicating the infection types (IT) against six geographically distinct Danish *Puccinia recondita* f. sp. *secalis* single-pustule isolates, shown as concentric circles. Ten pairs, comprising two closely related lines with divergent ITs, selected for SMRT RenSeq are marked by stars. (b) IT spectrum of the 10 resistant lines. The 10 susceptible lines are not depicted. They all had IT scores of 7 to 9 (Supplementary Table S3).

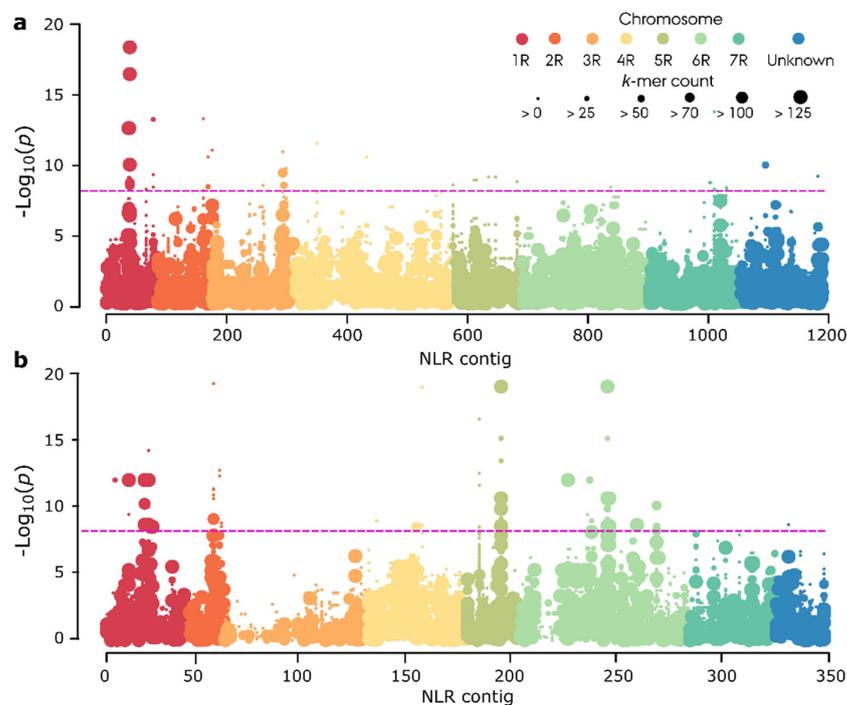
### 3.5. SMRT RenSeq, De Novo Assembly and NLR Annotation

To increase the contiguity and accuracy of the assembled NLRs, we sequenced the selected lines using PacBio<sup>®</sup> HiFi technology. SMRT RenSeq of these 20 selected restorer lines yielded between 0.27 to 2.24 Gb of circular consensus sequence (CCS) data per line with a mean of 1.40 Gb (Supplementary Table S6). De novo assembly using HiCanu produced 482 to 2122 contigs per line, with a mean of 1330. NLR annotation of the assemblies led to the identification of 288 to 1102 NLR contigs per line, with a mean of 646 (Supplementary Table S6), consisting of 151 to 573 partially (incomplete) annotated NLR contigs and 137 to 529 complete (full-length) NLR contigs per line.

Following SMRT RenSeq, we conducted an analysis across the 20 restorer lines to test the in vitro capture efficacy of the 600 K Triticeae NLR bait-library. In total, 35,199 baits aligned to a minimum of one line. Of these baits, 23,195 were shared with Lo7. On average,  $15,796 \pm 2910$  (SD) baits aligned to the genomic sequence from each line, of which 60.9% were considered core, as they aligned to NLR sequences present in more than 12 lines; 31.1% were common baits that aligned to NLR sequences present in 5–12 lines, and 8.0% were rare baits that aligned to NLR sequences present in fewer than 5 lines (Supplementary Figure S3e). Across lines, 75.9% of the aligned baits were shared with Lo7. Baits were found to align with an average of  $113 \text{ bp} \pm 14.9$  (SD), with a mean alignment identity of  $91.9\% \pm 4.3$  (SD) (Supplementary Figure S3a,b). Approximately 0.5% of the baits aligned with less than 80% alignment identity. Each bait mapped to an average of  $4 \pm 7$  SD unique NLR contigs, with a mean of  $63 \pm 82$  (SD) unique baits per NLR contig (Supplementary Figure S3c,d; Supplementary Table S5).

### 3.6. *K*-Mer-Based Association Genetics RenSeq (AgRenSeq) Analysis

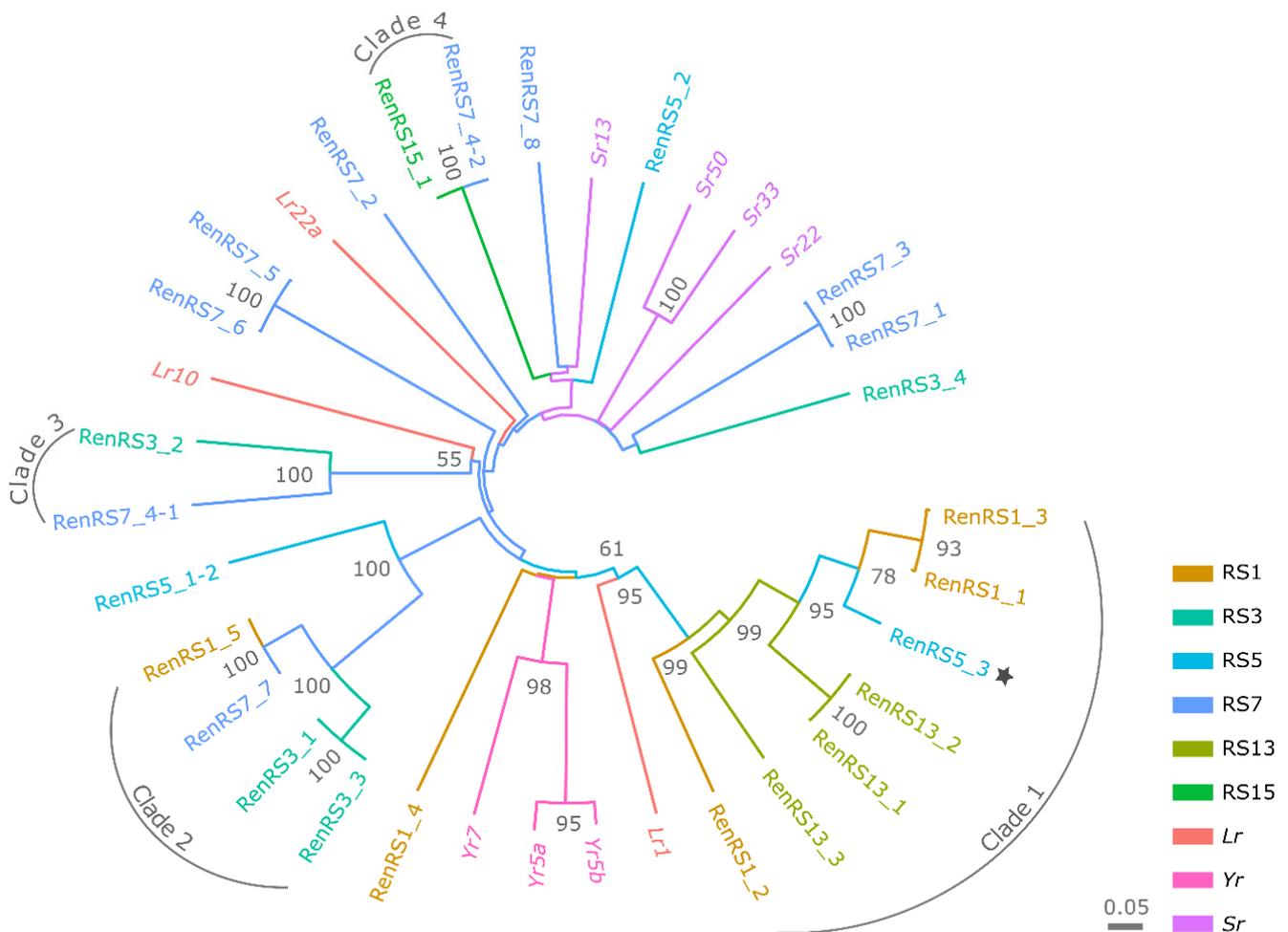
To identify candidate NLR gene(s) for leaf rust resistance in the restorer panel, we performed association genetics on the *k*-mers (sub-reads) derived from the raw RenSeq data. To correct for false associations due to population structure, we applied a general linear model (GLM). The NLRs from a resistant line were anchored to the Lo7 reference genome to provide an ordered template for AgRenSeq (Figure 4). AgRenSeq identified three peaks, the most promising of which mapped to chromosome arm 1RS ( $-\text{Log}_{10}(p) = 17.9$ ) with a *k*-mer count of 932. The other two peaks pointed to chromosome arm 5RL ( $-\text{Log}_{10}(p) = 18.7$ ), with a *k*-mer count of 231, and to chromosome arm 6RL ( $-\text{Log}_{10}(p) = 18.7$ ), with a *k*-mer count of 160, as being potentially involved in resistance against leaf rust (Supplementary Figure S6). In total, we identified 25 NLR contigs harboring resistant-specific *k*-mers were identified across the shortlisted resistant lines (Figure 4a,b; Supplementary Table S8a).



**Figure 4.** Identification of candidate leaf rust resistance NLR genes by AgRenSeq in 20 inbred rye lines. Each circle column on the x-axis represents an NLR contig from the SMRT RenSeq assembly of a resistant line anchored to the Lo7 rye reference genome and each dot on the y-axis represents a *k*-mer assigned a statistical association to resistance. (a) Resistant line RS13, phenotyped using isolate SPI-4. (b) Resistant line RS03, phenotyped using isolate SPI-6. The purple dashed line represents the Bonferroni-adjusted significance threshold based on number of *k*-mers.

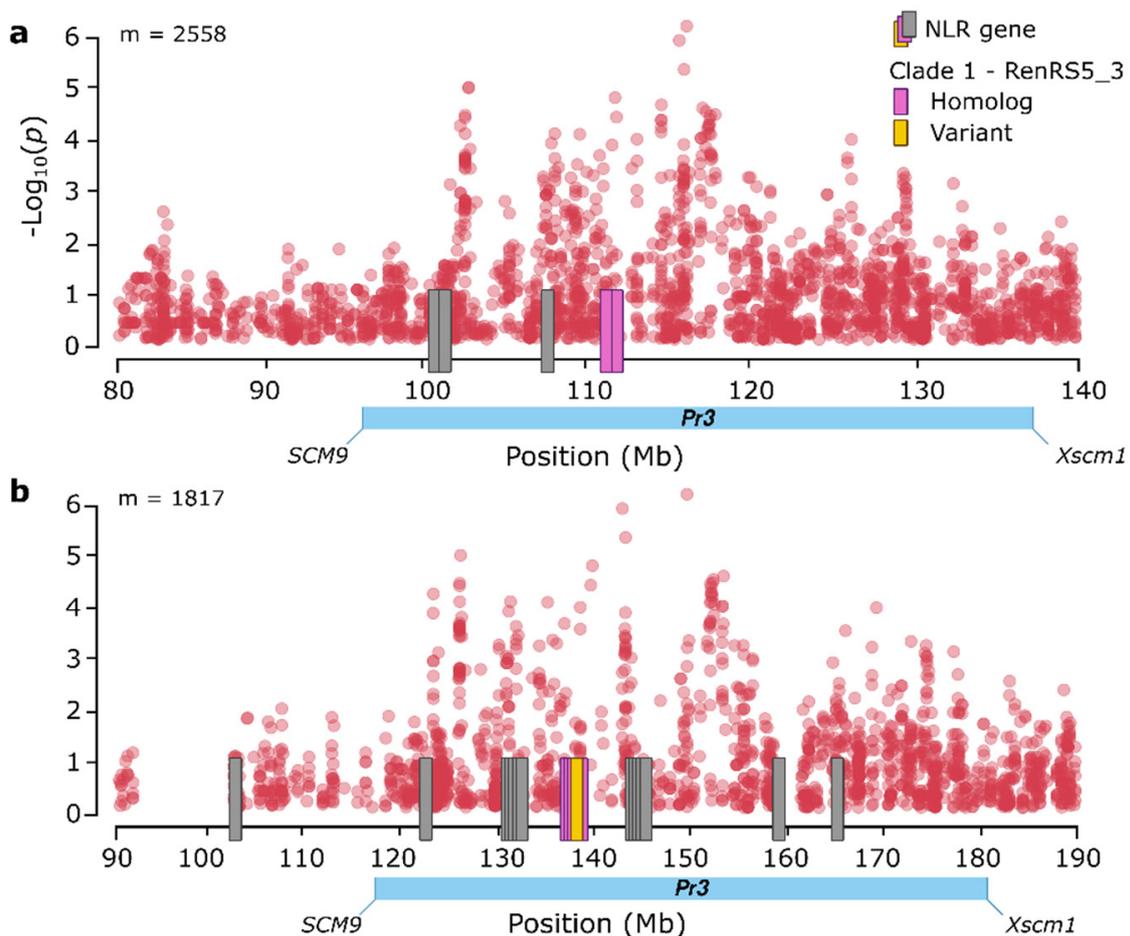
### 3.7. Characterization of Candidate Genes Conferring Resistance to Leaf Rust Resistance

We sought to improve the quality of the NLR annotation by expanding contig size. To this end, we generated a manually curated reference-based local assembly for each of the 25 candidate leaf rust resistance NLR contigs obtained by AgRenSeq. Accordingly, we selected raw CCS reads mapping to each contig in the relevant resistant line harboring the candidate NLR contig and trimmed and assembled them. This manual assembly step increased the contig size by  $2 \text{ kb} \pm 1.5 \text{ kb}$  (SD), resulting in a mean contig size of  $7.6 \text{ kb} \pm 1.8 \text{ kb}$  (SD) (Supplementary Table S8). Out of the 25 contigs, 16 contigs were considered to be complete (full-length) NLR genes, with the remaining 9 being partial (incomplete). On average, the full-length NLRs encoded an NLR protein of  $1147 \text{ amino acids (aa)} \pm 297$  (SD). To distinguish unique NLR genes from homologous contigs within the panel, we performed a phylogenetic analysis based on the sequence alignment of the NLR proteins (Figure 5), leading to the identification of four clades comprising 15 NLR proteins and consisting of candidate Pr proteins shared between at least two resistant assemblies. Out of the 10 NLRs forming basal splits in the phylogeny, 5 belonged to the resistant line RS7.



**Figure 5.** Phylogenetic relationship of candidate leaf rust resistance (*R*) NLR genes in 20 inbred rye lines. The tree was constructed using protein sequences, with bootstrapping values set as percentage recurrence of from 10,000 iterations shown for nodes with  $\geq 50\%$  recurrence. Candidate leaf rust *R* gene RenRS5\_3 is marked by an asterisk. Examples of cloned wheat leaf rust (*Lr*), stripe rust (*Yr*) and stem rust (*Sr*) genes are included for comparison.

Clade 1 comprised seven full-length NLR contigs encoding proteins of 1414 to 1700 aa (Figure 5, Supplementary Table S8). All members were anchored to position 111.15 Mb on chromosome arm 1RS in the Lo7 reference genome, which also coincided with the genomic block showing the highest association with leaf rust resistance in the GWAS analysis, spanning from 101 to 117 Mb (Table 1, Figure 6a, Supplementary Table S4). This genomic region contained five NLR genes in the Lo7 reference genome. Of the 2558 SNP markers positioned within the region on Lo7 chromosome arm 1RS, 1817 markers mapped to the Weining reference over the 90–190 Mb interval. The region of the Weining reference genome had 19 NLR genes, of which 15 formed three larger clusters 5 NLR genes, each within the region spanning 130–145 Mb (Figure 6b). For comparative analysis with known leaf rust resistance genes mapping to the same chromosome arm 1RS, we mapped restriction fragment length polymorphism (RFLP) markers co-segregating with *Lr26* and *Pr3* to the Lo7 and Weining reference genomes. The *Xmwig68* marker co-segregating with *Lr26* mapped to 16.6 Mb in Lo7 and 19.5 Mb in Weining, with *Lr26* positioned distal to *Xmwig68* towards the telomeric tip, thus clearly excluded from the above candidate interval. The *SCM9* and *Xscm1* markers, co-segregating with *Pr3* and flanking the gene on either side, mapped to 96.7 and 137.6 Mb in Lo7 and to 118.1 and 184.1 Mb in Weining, respectively (Figure 6a,b).



**Figure 6.** Visualization of SNP markers associated by GWAS with leaf rust resistance on chromosome arm 1RS in 92 inbred rye lines. (a) Lo7 reference genome with 2558 associated markers. (b) Weining reference genome with 1817 associated markers. The region harboring the known rye leaf rust resistance gene *Pr3* is marked as a blue bar based on flanking genetic markers *SCM9* and *Xscm1* [14]. The positions of annotated NLR genes are marked as vertical bars. NLR genes showing  $\geq 80\%$  sequence identity to the clade 1 member *RenRS5\_3* candidate *Pr* gene are in pink ('Homolog') and genes with  $\geq 95\%$  sequence identity are in orange ('Variant').

Clade 2 consisted of four NLR contigs encoding near-identical canonical NLR proteins of 1383 to 1417 aa, distinguished by intraspecific polymorphisms in the distal end of the C-terminal LRR domain (Table 2, Supplementary Table S8). Clade 3 contained two NLR contigs encoding canonical NLR proteins of 950 and 1050 aa that exhibited considerable intraspecific polymorphisms. Clade 4 was defined by two identical NLR contigs encoding a noncanonical NLR protein of 713 aa with no N-terminal CC domain.

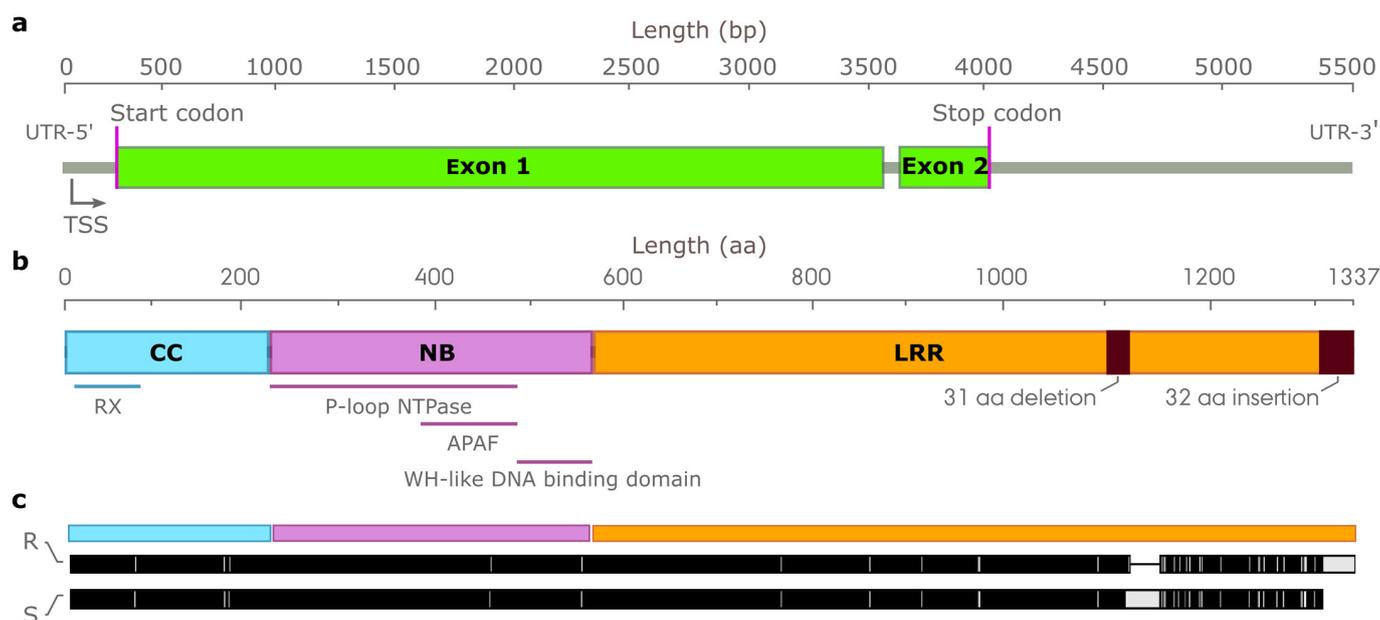
**Table 2.** Nucleotide-binding leucine-rich repeat (NLR) contigs associated with leaf rust resistance in four assigned clades based on phylogenetic relationships identified in resistant inbred rye lines. Contigs were mapped to the rye Lo7 reference genome.

Clade	NLR Contigs	Assemblies		Anchoring Position in Lo7	
		Resistant	SPI-Resistant	Chromosome Arm	Position (Mb)
1	7	RS1, RS3, RS5	RS13	1RS	111.15
2	4	RS1, RS3, RS7		5RL	792.53
3	2	RS3, RS7		NA	NA
4	2	RS7	RS15	5RL	807.97

SPI (single-pustule isolate); 'NA' (not anchored).

### 3.8. Characterization of Candidate NLR Genes Conferring Resistance to Leaf Rust on Chromosome Arm 1RS

We selected clade 1, with seven NLRs, as the most promising of candidate *Pr* genes for further characterization. The NLRs in this clade mapped to the region on chromosome arm 1RS, displaying the greatest association with leaf rust resistance (Figure 2), which also overlapped with the genomic location of the known leaf rust resistance gene *Pr3* (Figure 6a,b). To investigate gene presence/absence and sequence variation across the captured NLR complements from the 20 rye lines, we mapped all seven NLR contig members from clade 1 to the raw assemblies and scored and visualized the extent of sequence identity between putative homologs (Supplementary Figure S7). We then selected RenRS5\_3 for further study as its complete annotation further facilitated its characterization. The RenRS5\_3 contig was present (with  $\geq 95\%$  sequence identity) in 7 out of the 20 rye lines analyzed (Supplementary Table S9). Without transcript evidence, we predicted the most likely gene structure. We detected a putative transcription start site 45 bp into the contig, as well as two exons spanning the sequence from 199 to 3313 bp and 3404 to 4015 bp (Figure 7a). The corresponding NLR protein was predicted to have 1306 aa with an N-terminal CC domain from aa 17 to 225, a central NB-ARC domain from aa 226 to 586 and a C-terminal LRR domain from aa 587 to 1306 containing 25 LRR motifs (Figure 7b). Manual reference-based assembly of RenRS5\_3 in each of the seven lines carrying this NLR gene revealed two variants. One variant, denoted R, was conserved across the three resistant lines belonging to the IT<sub>1</sub> group with 99.9% sequence identity; the second variant, denoted S, was conserved amongst susceptible or partially resistant lines belonging to the IT<sub>2</sub> category with 97.8% sequence identity (Figure 3b). The two variants displayed a pairwise sequence similarity of 98.8% at the DNA level and were differentiated by 45 SNPs, 39 of which were situated in the genomic region encoding the C-terminal LRR domain. The variants encoded two NLR proteins showing 95.1% sequence identity, differentiated by 32 single amino-acid substitutions, a 31 aa deletion starting at aa 1105 and a 32 aa C-terminal extension of the LRR domain in the R variant (Figure 7c).



**Figure 7.** Characterization of the candidate leaf rust resistance nucleotide-binding leucine-rich repeat (NLR) gene RenRS5\_3. (a) Gene structure, (b) NLR protein motif and domain structure, coiled-coil (CC), nucleotide-binding (NB), leucine-rich repeat (LRR) and (c) Alignment of RenRS5\_3 resistant (R) and susceptible (S) variants with single amino acid substitutions marked as white lines.

We identified NLR contigs with  $\geq 97.8\%$  similarity with the coding sequence of RenRS5\_3 in three lines belonging to the IT categories IT<sub>1</sub> and IT<sub>3</sub>. These contigs exhibited a more complex gene structure and encoded a noncanonical NLR protein of 1314 aa lacking the N-terminal CC domain and showing  $\geq 87\%$  sequence identity with the RenRS5\_3 variants. Using the RenRS5\_3 sequence as a query against the reference genomes, we detected two paralogous NLR genes in the Lo7 genome residing at 111.02 (Lo7\_chr1R\_nlr\_64) and 111.15 Mb (Lo7\_chr1R\_nlr\_65) and sharing 83.5–85.1% sequence similarity with RenRS5\_3 (Figure 6a). In the Weining reference genome, RenRS5\_3 aligned with a larger cluster of five paralogous NLR genes located from 137.1 to 138.9 Mb with  $\geq 80.2\%$  sequence similarity (Figure 6b). One of the NLR genes from Weining (Wei\_chr1R\_nlr\_77) at 138.7 Mb showed 97.1% sequence identity with the R variant and 100% sequence identity with the ‘S’ variant.

## 4. Discussion

### 4.1. Leaf Rust Resistance Genes in Restorer Germplasm

In contrast to the predominant gene pools Petkus and Carsten deployed for hybrid breeding in rye, the restorer germplasm analyzed in this study boasts a high level of resistance to leaf rust (Supplementary Table S1) [2,15]. In relation to a recent study, we observed a strong correlation between resistance at the seedling stage and the adult stage [15], supporting the presence of major effect *R* genes in the restorer panel. Here, we discovered five intermediate-effect genomic regions on chromosome arms 1RS, 1RL, 2RL and 7RS, which are known to harbor *Pr* genes in rye [13–15]. Marker density was high, as the average inter-marker distance was 25.5 kb in the assayed germplasm, which contributed to our ability to distinguish true associations characterized by a distinct marker peak from false-positive outliers [71]. While GWAS has become a routine strategy for the mining of *R* genes and genomic regions associated with resistance in crop species, establishing causality between an SNP marker and a given *R* gene of interest can be impeded by several factors, such as low genetic diversity and the rate of decay of linkage disequilibrium (LD) [32,33]. The restorer population used here was previously shown to exhibit a high genetic diversity, a relatively higher effective population size and lower linkage disequilibrium compared with the seed-mother population [72]. The rate of linkage decay in outcrossing species such as rye is

often rapid and has been observed to occur within 4 kb in a similar hybrid breeding system in maize, providing potential opportunities for single-gene resolution at sufficient marker density [73]. While LD decay has not been investigated in the assayed restorer panel, the LD block surrounding the genomic region on chromosome arm 1RS appeared to span several Mb and contained multiple candidate *Pr* NLR genes (Figure 6). The small sample size and the potential for several rare or less prevalent *Pr* genes both reduce the phenotypic variance explained and, hence, the statistical power of GWAS [32,33,74]. Furthermore, several rye lines also displayed partial resistance against all six SPs tested, suggesting the potential presence of broad-spectrum, slow-rusting *R* gene(s) or QTL. Slow rusting resistance is expressed by a susceptible infection-type of host reaction with reduced infection frequency and severity [75]. Although we failed to identify markers significantly associated with leaf rust resistance by GWAS, this approach nevertheless provided an important insight into the genetic architecture underlying resistance in the rye diversity panel phenotyped here and offered an opportunity for a comparative analysis with RenSeq.

#### 4.2. Test of NLR Capture by the Bait Library

The 600 K Triticeae NLR bait-library was developed using NLR sequences from barley (*Hordeum vulgare* L.), hexaploid bread wheat (*Triticum aestivum* L.), tetraploid pasta wheat (*Triticum durum* L.), red wild einkorn (*T. urartu*), domesticated einkorn (*T. monococcum*) and three goat grass species (*Ae. tauschii*, *Ae. sharonensis* and *Ae. speltooides*) [42]. We first confirmed in silico that the bait library can perform a near complete capture of the NLR repertoire reported in the rye Lo7 reference genome [23]. During the course of our in vitro assessment, we observed a requirement for a minimum alignment identity between bait and target for efficient capture similar to that reported by Jupe et al. [37], thus supporting a key parameter used in the in silico analysis. In practice, only half of all baits aligned per line relative to our in silico analysis, suggesting a partial enrichment of the accessible NLR repertoire. This observation was further supported by the considerably lower number of NLR-annotated contigs per line (with a mean of 646 NLRs) compared to the 1167 NLRs present in the Lo7 reference genome. This discrepancy clearly falls outside the  $\leq 40\%$  expected range of accessional size variation of the NLR repertoire reported in previous RenSeq studies of *Ae. tauschii* and Arabidopsis [35,41]. Our assessment of bait library performance offered a glimpse into NLR capture across the Gülzow rye germplasm. However, additional data and analysis are required to explain the diverging success of NLR capture in this germplasm relative to the in silico analysis conducted on the Lo7 reference genome. Nonetheless, the immediate strength of RenSeq is the ability to mine for novel *R* genes in untapped genetic resources and wild-crop relatives without high-quality reference genomes [35]. This technique relies on the capacity of baits to align with a broad span of orthologous NLR genes with divergence sequences from the original NLR bait-library design. In the pioneering application of RenSeq by Jupe et al. [37], 68.5% of the enriched reads corresponded to genes that did not feature in the bait library design, thus demonstrating the plasticity of bait NLR capture.

#### 4.3. K-Mer Association Genetics with SMRT RenSeq Data

The discovery of *R* genes using association genetics on subsequences depends on the detection of *k*-mer sequences specific to resistant lines, often differentiated by an SNV or InDel in the underlying NLR sequence. In contrast to the pioneering AgRenSeq study using Illumina short-read sequencing [35], here we used the updated PacBio<sup>®</sup> SMRT sequel II HiFi system to achieve a higher precision in SNV and InDel calling [40]. When assessing the effect of sample size on *R* gene detection using AgRenSeq, Arora et al. [35] concluded that a panel of 80 diverse *Ae. tauschii* accessions was sufficient for detection of *SrTA1662* and *Sr46*, two widespread stem rust *R* genes present in 42% of the accessions in the panel. However, increasing the sample size to 140 accessions allowed the detection of two additional rare variants present in as few as 5% of the tested individuals. In this study, we investigated whether the high accuracy of SMRT circular consensus sequences would

enable the detection of *Pr* genes in a limited pool of 20 lines by AgRenSeq analysis. Using a multiple alignment-guided approach to group homologous NLR genes identified in multiple resistant lines, we successfully discovered four nonredundant candidate *Pr* genes. Comparative analysis of mapping positions was consistent with two genomic regions identified by GWAS on chromosome arms 1RS and 5RL.

#### 4.4. Co-Discovery of a Candidate *Pr* Gene on Chromosome Arm 1RS

Serving as a proof-of-concept, one of the candidate *Pr* genes identified by AgRenSeq mapped to a pair of paralogous NLR genes in the Lo7 reference genome located at the center of a genomic region associated with leaf rust resistance on chromosome arm 1RS [14]. The underlying candidate *Pr* gene encoded a canonical NLR protein exhibiting sequence variation in the LRR domain, including a large deletion and a C-terminal extension of equal length. Loss-of-function and gain-of-function mutations are frequently discovered in the LRR domain, likely modulating the ability of the receptor to perceive the pathogen effector [76,77]. In flax (*Linum usitatissimum* L.), loss of a repeated unit within the LRR domain of *M* has been associated with the inactivation of rust resistance [78]. In the Weining reference genome, the paralogous NLR cluster has undergone considerable gene expansion, although one of the paralogs shares an identical sequence to that of the susceptible variant of the candidate *Pr* gene. NLR clusters generate new functional diversity and arise from tandem duplication events often followed by unequal crossover, intracluster rearrangements and gene conversion events [79]. In plants, several NLR clusters form the complex loci of multiple paralogous genes encoding NLR proteins with different resistance specificities against distinct races of a certain pathogen [80,81]. In maize (*Zea mays* L.), the rust resistance locus *Rp1* contains multiple copies of paralogous genes in various haplotype combinations carrying the *Rp1-A* and *Rp1-H* specificities [82]. The evolutionary capacity of complex NLR loci was highlighted in a study by Richter et al. [83], who identified four novel resistance specificities derived from recombination events within the maize *Rp1* complex. The candidate *Pr* gene discovered here may have arisen from such a cluster expansion in an ancestral line, followed by mutation events leading to a gain-of-function phenotype. This hypothesis may similarly explain the presence of the *Pr* gene paralog identified in the panel, accentuating the rapid generation of novel genetic variation in NLR clusters [79].

Chromosome arm 1RS from 'Petkus' rye has been widely deployed for the improvement of rust resistance in wheat through chromosomal translocation, carrying, amongst others, the two leaf rust *R* genes *Pr3* and *Lr26* [70,84]. We used the RFLP markers *SCM9* and *Xscm1*, which genetically co-segregate with *Pr3* [14], to physically delimit the genomic interval of the gene in the Lo7 and Weining reference genomes; the genomic coordinates of *Pr3* spanned the candidate *Pr* gene identified here by AgRenSeq. *Pr3* was initially discovered in a self-incompatible backcross family with the resistant parent of a Russian population of the rye variety 'Jaroslavna', segregating as a single dominant locus [14]. Whether the candidate *Pr* gene is *Pr3*, a paralog, a variant thereof, or merely resides in the same region as *Pr3* remains to be investigated.

The information we present here on rye NLR gene sequences provides a valuable genetic resource for the development of functional markers and for future genetic engineering [36]. With most breeding companies having invested in the infrastructure for MAS, molecular markers can now be easily integrated directly into commercial breeding programs [25].

In conclusion, we successfully demonstrated the use of SMRT AgRenSeq using high-fidelity long-read sequencing technology on a reduced sample size for the discovery of *Pr* genes in rye. In contrast to the predominant gene pools 'Petkus' and 'Carsten' used for hybrid breeding in rye, we observed a high level of leaf rust resistance in the tested germplasm. We identified five genomic regions situated on chromosome arms 1RS, 1RL, 2RL, 5RL and 7RS by GWAS with a high-density SNP array. While an in silico assessment of the 600 K Triticeae bait-library demonstrated a near-complete theoretical

capture of the Lo7 NLR repertoire, our in vitro assessment showed partial target capture in the restorer germplasm. To explain this diverging NLR capture, additional data and analysis are required. We successfully identified four candidate *Pr* genes using SMRT AgRenSeq, including a canonical NLR gene whose genomic location overlapped with the top SNP associated with leaf rust resistance on chromosome arm 1RS. The nearby situated leaf rust resistance gene *Pr3* may be the underlying candidate *Pr3* gene or a paralog or variant thereof. In order to validate the candidate *Pr* genes identified in this study further investigations are required. Genetic validation could involve using a different type of genetically structured population, e.g., biparental, in combination with genetic markers to independently show a correlation between the segregation of markers and the resistant phenotype [85]. Functional validation could involve using knock out, EMS [86] or CRISPR-Cas9 [87] and/or transgenesis to express the functional unit in a susceptible background and confer the resistant phenotype to that background [88].

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/cells11081273/s1>. Supplementary Table S1: Leaf rust severity of top-yielding hybrid rye (*Secale cereale* L.) and hybrid rye/population of rye mixtures tested in the Danish official trials in 2019 across nine locations. Supplementary Table S2: National sampling of *Puccinia recondita* f. sp. *secalis*, causative agent of leaf rust disease in rye (*Secale cereale* L.) in Denmark 2018–2019. Supplementary Table S3: Infection type response (0–9) of 92 inbred rye (*Secale cereale* L.) lines phenotyped using six geographically distinct Danish *Puccinia recondita* f. sp. *secalis* single-pustule isolates (SPI) collected in 2018 (SPI-1, SPI-2 and SPI-3) or 2019 (SPI-4, SPI-5 and SPI-6). Supplementary Table S4: Association of 600 K SNP markers with leaf rust resistance on chromosome arms 1RL, 2RL, 5RL and 7RL across 92 inbred rye (*Secale cereale* L.) lines phenotyped using *Puccinia recondita* f. sp. *secalis* single-pustule isolates. Supplementary Table S5: Metrics of in silico assessment of the 600 K Triticeae nucleotide-binding leucine-rich repeat (NLR) bait-library on the rye (*Secale cereale* L.) reference genome Lo7 and of in vitro assessment of 20 inbred rye lines analyzed using single-molecule real-time resistance gene enrichment sequencing. Supplementary Table S6: Metrics of single-molecule real-time (SMRT) resistance gene enrichment sequencing (RenSeq) analysis, assembly and NLR annotation of 20 inbred rye (*Secale cereale* L.) lines. Supplementary Table S7: Candidate leaf rust resistance nucleotide-binding leucine-rich repeat (NLR)-annotated contigs in resistant rye (*Secale cereale* L.) lines analyzed using resistance gene enrichment sequencing (RenSeq) identified using association genetics RenSeq (AgRenSeq) and AgRenSeq-GLM. Supplementary Table S8: Characteristics of nucleotide-binding leucine-rich repeat-annotated contigs associated with leaf rust resistance in 20 inbred rye (*Secale cereale* L.) lines analyzed using resistance gene enrichment sequencing and association genetics. Supplementary Table S9: BLASTN results between the candidate leaf rust resistance gene RenRS5\_3 long-read resistance gene enrichment sequencing (RenSeq) raw circular consensus sequence data from 20 inbred rye (*Secale cereale* L.) lines. Data File S1. Genotype information for 261,406 informative 600 K SNP array markers in 92 inbred rye (*Secale cereale* L.) lines (.Rdata). Supplementary Figure S1: Geographical location of the origin of the six *Puccinia recondita* f. sp. *secalis* single-pustule isolates collected in Denmark in 2018 and 2019. Supplementary Figure S2: Manhattan plots for genome-wide association study (GWAS) for leaf rust resistance in 92 hybrid rye (*Secale cereale* L.) restorer breeding lines using 261,406 informative SNP markers mapped to the Lo7 reference genome. The entire germplasm was phenotyped for resistance against six distinct *Puccinia recondita* f. sp. *secalis* single pustule isolates (SPI) in a greenhouse with two replicate trials. The Bonferroni-adjusted significance threshold based on informative markers was 6.72. Supplementary Figure S3: Characteristics of NLR target sequencing in 20 inbred rye (*Secale cereale* L.) lines using a 60 K Triticeae NLR bait library. The reference genome Lo7 was included as an in silico reference with baits filtered for >80% alignment identity. (a) Distribution of bait alignment in lines with raw consensus sequence (CCS) data. (b) Distribution of bait alignment identity against raw CCS data. (c) Unique baits per NLR annotated contig. (d) Unique NLR annotated contigs per bait. (e) Total number of unique baits per breeding line and distribution of bait categories; ‘rare’ baits aligned to NLR sequence present in fewer than five lines, ‘common’ baits to 5–12 lines and ‘core’ baits to more than 12 lines. Supplementary Figure S4: Identification of resistance associated nucleotide-binding leucine-rich repeat (NLR) annotated contig across 20 inbred rye (*Secale cereale* L.) lines by association genetics resistance gene enrichment sequencing (AgRenSeq) analysis. (a) Standard AgRenSeq with candidate

leaf rust resistance gene highlighted in red. (b), AgRenSeq with NLR contigs mapped to the Lo7 reference genome. (c), AgRenSeq-GLM. Supplementary Figure S5: Association genetics resistance gene enrichment sequencing (AgRenSeq) analysis for the identification of NLR contigs harboring resistant-specific *k*-mers in 20 inbred rye (*Secale cereale* L.) lines. Supplementary Figure S6: Association genetics resistance gene enrichment sequencing (AgRenSeq) GLM analysis for the identification of NLR contigs harboring resistant-specific *k*-mers in 20 inbred rye (*Secale cereale* L.) lines. Supplementary Figure S7: Distribution of alignment identity (%) of seven leaf rust resistance associated nucleotide-binding leucine-rich repeat (NLR) contigs belonging to clade 1 in 20 inbred rye (*Secale cereale* L.) lines long-read resistance gene enrichment sequencing (RenSeq) raw circular consensus sequence (CCS) data.

**Author Contributions:** Conceptualization, N.M.V., K.M., B.B.H.W., J.O. and A.J.; methodology, N.M.V., K.M., P.S., J.O. and A.J.; software, N.M.V., K.M., B.S., B.B.H.W.; validation, N.M.V., K.M., B.S., P.S., B.B.H.W., J.O. and A.J.; formal analysis, N.M.V.; investigation, N.M.V., K.M., B.S., B.B.H.W., J.O. and A.J.; resources, B.B.H.W. and P.S.K.; data curation, N.M.V. and B.S.; writing—original draft preparation, N.M.V.; writing—review and editing, N.M.V., K.M., B.S., B.B.H.W., P.S., M.S.H., A.F.J., P.S.K., J.O. and A.J.; visualization, N.M.V.; supervision, M.S.H., A.F.J., J.O. and A.J.; project administration, J.O. and A.J.; funding acquisition, M.S.H., A.F.J., J.O. and A.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was funded by Innovation Fund Denmark (grant no. 8053-00085B), the Pajbjerg Foundation, the Danish Green Development and Demonstration Program (project Jr. NR. 34009-19-1603), and the Biotechnology and Biological Sciences Research Council (BBSRC) Designing Future Wheat Cross-Institute Strategic Programme (BBS/E/J/000PR9780), and King Abdullah University of Science and Technology.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The authors declare that the main data supporting the findings of this study are available within the article and its supporting information. All data and R scripts used to conduct the analysis and construct graphical outputs have been provided at <https://doi.org/10.5281/zenodo.5725078> (accessed on 17 February 2022).

**Acknowledgments:** We thank Hanne Svenstrup at Nordic Seed A/S for her contribution to the genotypic data collection; Ellen Jørgensen at the Global Rust Reference Center, Aarhus University (Slagelse, Denmark) for her assistance in the recovery, inoculation and multiplication procedures of cereal rusts; Jacob Sørensen at the Institute of Agroecology, Aarhus University (Slagelse, Denmark) for assistance in the multiplication of rust samples; Johannes Hiller, Anette Deterding and Marlene Walbrodt at Nordic Seed Germany GmbH (Nienstädt, Germany) for seed multiplication; and GenomeDK at Aarhus Genome Data Center for rapid assistance and trouble-shooting.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Miedaner, T.; Sperling, U. Effect of leaf rust and yield components of winter rye hybrids and assessment of quantitative resistance. *J. Phytopathol.* **1995**, *143*, 725–730. [[CrossRef](#)]
2. Miedaner, T.; Gey, A.-K.M.; Sperling, U.; Geiger, H.H. Quantitative-genetic analysis of leaf-rust resistance in seedling and adult-plant stages of inbred lines and their testcrosses in winter rye. *Plant Breed.* **2002**, *121*, 475–479. [[CrossRef](#)]
3. Sortsinfo. The Danish Official Trial Records. Available online: <https://sortinfo.dk/> (accessed on 10 February 2021).
4. Miedaner, T.; Klocke, B.; Flath, K.; Geiger, H.H.; Weber, W.E. Diversity, spatial variation, and temporal dynamics of virulences in the German leaf rust (*Puccinia recondita* f. sp. *secalis*) population in winter rye. *Eur. J. Plant Pathol.* **2011**, *132*, 23–35. [[CrossRef](#)]
5. Anikster, Y.; Bushnell, W.; Roelfs, A.; Eilam, T.; Manisterski, J. *Puccinia recondita* causing leaf rust on cultivated wheats, wild wheats, and rye. *Can. J. Bot.* **1997**, *75*, 2082–2096. [[CrossRef](#)]
6. Hanzlik, K.; Gerowitt, B. Occurrence and distribution of important weed species in German winter oilseed rape fields. *J. Plant Dis. Prot.* **2012**, *119*, 107–120. [[CrossRef](#)]
7. De Mol, F.; Von Redwitz, C.; Gerowitt, B. Weed species composition of maize fields in Germany is influenced by site and crop sequence. *Weed Res.* **2015**, *55*, 574–585. [[CrossRef](#)]
8. Andreasen, C.; Stryhn, H. Increasing weed flora in Danish beet, pea and winter barley fields. *Crop Prot.* **2012**, *36*, 11–17. [[CrossRef](#)]

9. Figueroa, M.; Dodds, P.N.; Henningsen, E.C. Evolution of virulence in rust fungi-multiple solutions to one problem. *Curr. Opin. Plant Biol.* **2020**, *56*, 20–27. [[CrossRef](#)]
10. Kolmer, J.A.; Ordonez, M.E.; German, S.; Morgounov, A.; Pretorius, Z.; Visser, B.; Goyeau, H.; Anikster, Y.; Acevedo, M. Multilocus genotypes of the wheat leaf rust fungus *Puccinia triticina* in worldwide regions indicate past and current long-distance migration. *Phytopathology* **2019**, *109*, 1453–1463. [[CrossRef](#)]
11. Hovmoller, M.S.; Yahyaoui, A.H.; Milus, E.A.; Justesen, A.F. Rapid global spread of two aggressive strains of a wheat rust fungus. *Mol. Ecol.* **2008**, *17*, 3818–3826. [[CrossRef](#)]
12. Nelson, R.; Wiesner-Hanks, T.; Wissler, R.; Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **2018**, *19*, 21–33. [[CrossRef](#)] [[PubMed](#)]
13. Wehling, P.; Linz, A.; Hackauf, B.; Roux, S.R.; Ruge, B.; Klocke, B. Leaf-rust resistance in rye (*Secale cereale* L.). 1. Genetic analysis and mapping of resistance genes Pr1 and Pr2. *Theor. Appl. Genet.* **2003**, *107*, 432–438. [[CrossRef](#)] [[PubMed](#)]
14. Roux, S.R.; Hackauf, B.; Linz, A.; Ruge, B.; Klocke, B.; Wehling, P. Leaf-rust resistance in rye (*Secale cereale* L.). 2. Genetic analysis and mapping of resistance genes Pr3, Pr4, and Pr5. *Theor. Appl. Genet.* **2004**, *110*, 192–201. [[CrossRef](#)] [[PubMed](#)]
15. Vendelbo, N.M.; Mahmood, K.; Sarup, P.; Orabi, J.; Hovmøller, M.; Justesen, A.F.; Kristensen, P.S.; Jahoor, A. Discovery of a novel leaf rust (*Puccinia recondita*) resistance in rye (*Secale cereale* L.) Using Association Genomics. *Cells* **2021**, *11*, 64. [[CrossRef](#)]
16. McIntosh, R.A.; Frieb, B.; Jiang, J.; Gill, B.S. Cytogenetical studies in wheat XVI. Chromosomal location of a new gene for resistance to leaf rust in a Japanese wheat-rye translocation line. *Euphytica* **1995**, *82*, 141–147. [[CrossRef](#)]
17. Friebe, B.; Jiang, J.; Raupp, W.J.; McIntosh, R.A.; Gill, B.S. Characterization of wheat-alien translocations conferring resistance to diseases and pests: Current status. *Euphytica* **1996**, *91*, 59–87. [[CrossRef](#)]
18. Kourelis, J.; Van Der Hoorn, R.A. Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell* **2018**, *30*, 285–299. [[CrossRef](#)]
19. Qi, D.; DeYoung, B.J.; Innes, R.W. Structure-function analysis of the coiled-coil and leucine-rich repeat domains of the RPS5 disease resistance protein. *Plant Physiol.* **2012**, *158*, 1819–1832. [[CrossRef](#)]
20. Monino-Lopez, D.; Nijenhuis, M.; Kodde, L.; Kamoun, S.; Salehian, H.; Schentsnyi, K.; Stam, R.; Lokossou, A.; Abd-El-Haliem, A.; Visser, R.G. Allelic variants of the NLR protein Rpi-chc1 differentially recognize members of the Phytophthora infestans PexRD12/31 effector superfamily through the leucine-rich repeat domain. *Plant J.* **2021**, *107*, 182–197. [[CrossRef](#)]
21. Williams, S.J.; Sornaraj, P.; de Courcy-Ireland, E.; Menz, R.I.; Kobe, B.; Ellis, J.G.; Dodds, P.N.; Anderson, P.A. An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Mol. Plant-Microbe Interact.* **2011**, *24*, 897–906. [[CrossRef](#)]
22. Maekawa, T.; Cheng, W.; Spiridon, L.N.; Töller, A.; Lukasik, E.; Saijo, Y.; Liu, P.; Shen, Q.-H.; Micluta, M.A.; Somssich, I.E. Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* **2011**, *9*, 187–199. [[CrossRef](#)] [[PubMed](#)]
23. Rabanus-Wallace, M.T.; Hackauf, B.; Mascher, M.; Lux, T.; Wicker, T.; Gundlach, H.; Báez, M.; Houben, A.; Mayer, K.F.X.; Guo, L.; et al. Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. *Nat. Genet.* **2021**, *53*, 564–573. [[CrossRef](#)] [[PubMed](#)]
24. Li, G.; Wang, L.; Yang, J.; He, H.; Jin, H.; Li, X.; Ren, T.; Ren, Z.; Li, F.; Han, X.; et al. A high-quality genome assembly highlights rye genomic characteristics and agronomically important genes. *Nat. Genet.* **2021**, *53*, 574–584. [[CrossRef](#)] [[PubMed](#)]
25. Miedaner, T.; Korzun, V. Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* **2012**, *102*, 560–566. [[CrossRef](#)]
26. Beukert, U.; Thorwarth, P.; Zhao, Y.; Longin, C.F.H.; Serfling, A.; Ordon, F.; Reif, J.C. Comparing the potential of marker-assisted selection and genomic prediction for improving rust resistance in hybrid wheat. *Front. Plant Sci.* **2020**, *11*, 1650. [[CrossRef](#)]
27. Liu, R.; Lu, J.; Zhou, M.; Zheng, S.; Liu, Z.; Zhang, C.; Du, M.; Wang, M.; Li, Y.; Wu, Y. Developing stripe rust resistant wheat (*Triticum aestivum* L.) lines with gene pyramiding strategy and marker-assisted selection. *Genet. Res. Crop Evol.* **2020**, *67*, 381–391. [[CrossRef](#)]
28. Hickey, L.T.; Germán, S.E.; Pereyra, S.A.; Diaz, J.E.; Ziem, L.A.; Fowler, R.A.; Platz, G.J.; Franckowiak, J.D.; Dieters, M.J. Speed breeding for multiple disease resistance in barley. *Euphytica* **2017**, *213*, 1–14. [[CrossRef](#)]
29. Ghosh, S.; Watson, A.; Gonzalez-Navarro, O.E.; Ramirez-Gonzalez, R.H.; Yanes, L.; Mendoza-Suárez, M.; Simmonds, J.; Wells, R.; Rayner, T.; Green, P. Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nat. Protoc.* **2018**, *13*, 2944–2963. [[CrossRef](#)]
30. Bajgain, P.; Rouse, M.N.; Tsilo, T.J.; Macharia, G.K.; Bhavani, S.; Jin, Y.; Anderson, J.A. Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. *PLoS ONE* **2016**, *11*, e0155760. [[CrossRef](#)]
31. Rollar, S.; Serfling, A.; Geyer, M.; Hartl, L.; Mohler, V.; Ordon, F. QTL mapping of adult plant and seedling resistance to leaf rust (*Puccinia triticina* Eriks.) in a multiparent advanced generation intercross (MAGIC) wheat population. *Theor. Appl. Genet.* **2021**, *134*, 37–51. [[CrossRef](#)]
32. Alqudah, A.M.; Sallam, A.; Baenziger, P.S.; Börner, A. GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: Lessons from Barley—A review. *J. Adv. Res.* **2020**, *22*, 119–135. [[CrossRef](#)] [[PubMed](#)]
33. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **2013**, *9*, 1–9. [[CrossRef](#)] [[PubMed](#)]

34. Florence, J.; Vernaldi, S.; Maekawa, T. Evolution and conservation of plant NLR functions. *Front. Immunol.* **2013**, *4*, 297:1–297:16. [[CrossRef](#)]
35. Arora, S.; Steuernagel, B.; Gaurav, K.; Chandramohan, S.; Long, Y.; Matny, O.; Johnson, R.; Enk, J.; Periyannan, S.; Singh, N.; et al. Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* **2019**, *37*, 139–143. [[CrossRef](#)] [[PubMed](#)]
36. Dong, O.X.; Ronald, P.C. Genetic engineering for disease resistance in plants: Recent progress and future perspectives. *Plant Physiol.* **2019**, *180*, 26–38. [[CrossRef](#)] [[PubMed](#)]
37. Jupe, F.; Witek, K.; Verweij, W.; Sliwka, J.; Pritchard, L.; Etherington, G.J.; Maclean, D.; Cock, P.J.; Leggett, R.M.; Bryan, G.J.; et al. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *Plant J.* **2013**, *76*, 530–544. [[CrossRef](#)] [[PubMed](#)]
38. Witek, K.; Jupe, F.; Witek, A.I.; Baker, D.; Clark, M.D.; Jones, J.D. Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol.* **2016**, *34*, 656–660. [[CrossRef](#)]
39. Xing, L.; Hu, P.; Liu, J.; Witek, K.; Zhou, S.; Xu, J.; Zhou, W.; Gao, L.; Huang, Z.; Zhang, R. Pm21 from *Haynaldia villosa* encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Mol. Plant* **2018**, *11*, 874–878. [[CrossRef](#)]
40. Wenger, A.M.; Peluso, P.; Rowell, W.J.; Chang, P.-C.; Hall, R.J.; Concepcion, G.T.; Ebler, J.; Fungtammasan, A.; Kolesnikov, A.; Olson, N.D. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat. Biotechnol.* **2019**, *37*, 1155–1162. [[CrossRef](#)]
41. Van de Weyer, A.L.; Monteiro, F.; Furzer, O.J.; Nishimura, M.T.; Cevik, V.; Witek, K.; Jones, J.D.G.; Dangl, J.L.; Weigel, D.; Bemm, F. A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. *Cell* **2019**, *178*, 1260–1272.e1214. [[CrossRef](#)]
42. Steuernagel, B.; Periyannan, S.K.; Hernandez-Pinzon, I.; Witek, K.; Rouse, M.N.; Yu, G.; Hatta, A.; Ayliffe, M.; Bariana, H.; Jones, J.D.; et al. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* **2016**, *34*, 652–655. [[CrossRef](#)] [[PubMed](#)]
43. Seong, K.; Seo, E.; Witek, K.; Li, M.; Staskawicz, B. Evolution of NLR resistance genes with noncanonical N-terminal domains in wild tomato species. *New Phytol.* **2020**, *227*, 1530–1543. [[CrossRef](#)] [[PubMed](#)]
44. Barbey, C.R.; Lee, S.; Verma, S.; Bird, K.A.; Yocca, A.E.; Edger, P.P.; Knapp, S.J.; Whitaker, V.M.; Folta, K.M. Disease resistance genetics and genomics in octoploid strawberry. *G3 Genes Genomes Genet.* **2019**, *9*, 3315–3332. [[CrossRef](#)] [[PubMed](#)]
45. USDA. United States Department of Agriculture: Wheat and Barley DNA Extraction Protocol (96-Well Plate Format). Available online: [https://www.ars.usda.gov/ARSUserFiles/60701500/SmallGrainsGenotypingLaboratory/Protocols/wheat%20and%20barleyDNA%20extraction\\_original.pdf](https://www.ars.usda.gov/ARSUserFiles/60701500/SmallGrainsGenotypingLaboratory/Protocols/wheat%20and%20barleyDNA%20extraction_original.pdf) (accessed on 7 July 2021).
46. Pallotta, M.A.; Warner, P.; Fox, R.L.; Kuchel, H.; Jefferies, S.J.; Langridge, P. Marker assisted wheat breeding in the southern region of Australia. In *Proceedings of the Tenth International Wheat Genetics Symposium*; Istituto Sperimentale per la Cerealicoltura: Puglia, Italy, 2003; pp. 789–791.
47. McNeal, F.H.; Konzak, C.F.; Smith, E.P.; Tate, W.S.; Russell, T.S. *A Uniform System for Recording and Processing Cereal Research Data*; No. REP-10904; CIMMYT: El Batan, Mexico, 1971.
48. Hovmøller, M.S.; Rodriguez-Algaba, J.; Thach, T.; Sorensen, C.K. Race typing of *Puccinia striiformis* on wheat. In *Wheat Rust Diseases: Methods and Protocols*; Humana Press: New York, NY, USA, 2017; pp. 29–40.
49. Thach, T.; Ali, S.; Justesen, A.; Rodriguez-Algaba, J.; Hovmøller, M. Recovery and virulence phenotyping of the historic ‘Stubbs collection’ of the yellow rust fungus *Puccinia striiformis* from wheat. *Ann. Appl. Biol.* **2015**, *167*, 314–326. [[CrossRef](#)]
50. Bauer, E.; Schmutzer, T.; Barilar, I.; Mascher, M.; Gundlach, H.; Martis, M.M.; Twardziok, S.O.; Hackauf, B.; Gordillo, A.; Wilde, P.; et al. Towards a whole-genome sequence for rye (*Secale cereale* L.). *Plant J.* **2017**, *89*, 853–869. [[CrossRef](#)]
51. Vendelbo, N.M.; Mahmood, K.; Sarup, P.; Orabi, J.; Kristensen, P.S.; Jahoor, A. Discovery of a novel powdery mildew (*Blumeria graminis*) resistance locus in rye (*Secale cereale* L.). *Sci. Rep.* **2021**, *11*, 1–15.
52. RStudio Team. Rstudio: Integrated Development for R. RStudio, Inc., Boston. Available online: <http://www.rstudio.com> (accessed on 2 August 2021).
53. R Core Team. *R A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria. Available online: <https://www.R-project.org/> (accessed on 8 February 2021).
54. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.A.; Buckler, E.S.; Zhang, Z. GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **2012**, *28*, 2397–2399. [[CrossRef](#)]
55. NCBI. National Center for Biotechnology Information. Available online: <https://www.ncbi.nlm.nih.gov> (accessed on 13 June 2021).
56. Paradis, E.; Schliep, K. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **2019**, *35*, 526–528. [[CrossRef](#)]
57. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, 256–259. [[CrossRef](#)]
58. Nurk, S.; Walenz, B.P.; Rhie, A.; Vollger, M.R.; Logsdon, G.A.; Grothe, R.; Miga, K.H.; Eichler, E.E.; Phillippy, A.M.; Koren, S. HiCanu: Accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *Genome Res.* **2020**, *30*, 1291–1305. [[CrossRef](#)]
59. Steuernagel, B.; Jupe, F.; Witek, K.; Jones, J.D.; Wulff, B.B. NLR-parser: Rapid annotation of plant NLR complements. *Bioinformatics* **2015**, *31*, 1665–1667. [[CrossRef](#)] [[PubMed](#)]

60. Steuernagel, B.; Witek, K.; Krattinger, S.G.; Ramirez-Gonzalez, R.H.; Schoonbeek, H.J.; Yu, G.; Baggs, E.; Witek, A.I.; Yadav, I.; Krasileva, K.V.; et al. The NLR-annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiol.* **2020**, *183*, 468–482. [[CrossRef](#)] [[PubMed](#)]
61. Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **2011**, *3*, 180–185. [[CrossRef](#)]
62. Marçais, G.; Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **2011**, *27*, 764–770. [[CrossRef](#)] [[PubMed](#)]
63. GeneiousPrime. Geneious Prime (v. 2020.2.3). 2020. Available online: <https://www.geneious.com> (accessed on 11 December 2021).
64. Stanke, M.; Steinkamp, R.; Waack, S.; Morgenstern, B. AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res.* **2004**, *32*, W309–W312. [[CrossRef](#)] [[PubMed](#)]
65. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; Hunter, S. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
66. Martin, E.C.; Sukarta, O.C.; Spiridon, L.; Grigore, L.G.; Constantinescu, V.; Tacutu, R.; Goverse, A.; Petrescu, A.-J. LRRpredictor—A new LRR motif detection method for irregular motifs of plant NLR proteins using an ensemble of classifiers. *Genes* **2020**, *11*, 286. [[CrossRef](#)]
67. Toparlan, E.; Karabag, K.; Bilge, U. A workflow with R: Phylogenetic analyses and visualizations using mitochondrial cytochrome b gene sequences. *PLoS ONE* **2020**, *15*, e0243927. [[CrossRef](#)]
68. Yu, G. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinform.* **2020**, *69*, e96. [[CrossRef](#)]
69. UniProt, C. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, 480–489. [[CrossRef](#)]
70. Mago, R.; Miah, H.; Lawrence, G.J.; Wellings, C.R.; Spielmeier, W.; Bariana, H.S.; McIntosh, R.A.; Pryor, A.J.; Ellis, J.G. High-resolution mapping and mutation analysis separate the rust resistance genes Sr31, Lr26 and Yr9 on the short arm of rye chromosome 1. *Theor. Appl. Genet.* **2005**, *112*, 41–50. [[CrossRef](#)] [[PubMed](#)]
71. Gaikpa, D.S.; Kessel, B.; Presterl, T.; Ouzunova, M.; Galiano-Carneiro, A.L.; Mayer, M.; Melchinger, A.E.; Schön, C.-C.; Miedaner, T. Exploiting genetic diversity in two European maize landraces for improving Gibberella ear rot resistance using genomic tools. *Theor. Appl. Genet.* **2021**, *134*, 793–805. [[CrossRef](#)] [[PubMed](#)]
72. Vendelbo, N.M.; Sarup, P.; Orabi, J.; Kristensen, P.S.; Jahoor, A. Genetic structure of a germplasm for hybrid breeding in rye (*Secale cereale* L.). *PLoS ONE* **2020**, *15*, e0239541. [[CrossRef](#)] [[PubMed](#)]
73. Wu, Y.; San Vicente, F.; Huang, K.; Dhaliwayo, T.; Costich, D.E.; Semagn, K.; Sudha, N.; Olsen, M.; Prasanna, B.M.; Zhang, X.; et al. Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theor. Appl. Genet.* **2016**, *129*, 753–765. [[CrossRef](#)] [[PubMed](#)]
74. Gibson, G. Rare and common variants: Twenty arguments. *Nat. Rev. Genet.* **2012**, *13*, 135–145. [[CrossRef](#)] [[PubMed](#)]
75. Sucher, J.; Boni, R.; Yang, P.; Rogowsky, P.; Buchner, H.; Kastner, C.; Kumlehn, J.; Krattinger, S.G.; Keller, B. The durable wheat disease resistance gene Lr34 confers common rust and northern corn leaf blight resistance in maize. *Plant Biotechnol. J.* **2017**, *15*, 489–496. [[CrossRef](#)]
76. Xie, J.; Guo, G.; Wang, Y.; Hu, T.; Wang, L.; Li, J.; Qiu, D.; Li, Y.; Wu, Q.; Lu, P. A rare single nucleotide variant in Pm5e confers powdery mildew resistance in common wheat. *New Phytol.* **2020**, *228*, 1011–1026. [[CrossRef](#)]
77. Halterman, D.A.; Wise, R.P. A single-amino acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. *Plant J.* **2004**, *38*, 215–226. [[CrossRef](#)]
78. Anderson, P.A.; Lawrence, G.J.; Morrish, B.C.; Ayliffe, M.A.; Finnegan, E.J.; Ellis, J.G. Inactivation of the flax rust resistance gene M associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell* **1997**, *9*, 641–651.
79. Barragan, A.C.; Weigel, D. Plant NLR diversity: The known unknowns of Pan-NLRomes. *Plant Cell* **2021**, *33*, 814–831. [[CrossRef](#)]
80. Wei, F.; Gobelmann-Werner, K.; Morroll, S.M.; Kurth, J.; Mao, L.; Wing, R.; Leister, D.; Schulze-Lefert, P.; Wise, R.P. The Mla (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* **1999**, *153*, 1929–1948. [[CrossRef](#)] [[PubMed](#)]
81. Jones, D.; Dickinson, M.; Balint-Kurti, P.; Dixon, M.; Jones, J. Two complex resistance loci revealed in tomato by classical and RFLP mapping of the Cf-2, Cf-4, Cf-5, and Cf-9 genes for resistance to *Cladosporium fulvum*. *Mol. Plant Microbe Interact.* **1993**, *6*, 348. [[CrossRef](#)]
82. Smith, S.M.; Pryor, A.J.; Hulbert, S.H. Allelic and haplotypic diversity at the rp1 rust resistance locus of maize. *Genetics* **2004**, *167*, 1939–1947. [[CrossRef](#)] [[PubMed](#)]
83. Richter, T.E.; Pryor, T.J.; Bennetzen, J.L.; Hulbert, S.H. New rust resistance specificities associated with recombination in the Rp1 complex in maize. *Genetics* **1995**, *141*, 373–381. [[CrossRef](#)] [[PubMed](#)]
84. Lukaszewski, A.J. Frequency of 1RS. 1AL and 1RS. 1BL translocations in United States wheats. *Crop Sci.* **1990**, *30*, 1151–1153. [[CrossRef](#)]
85. Qiu, L.; Wang, H.; Li, Y.; Wang, W.; Liu, Y.; Mu, J.; Geng, M.; Guo, W.; Hu, Z.; Ma, J. Fine mapping of the wheat leaf rust resistance gene LrLC10 (Lr13) and validation of its co-segregation markers. *Front. Plant Sci.* **2020**, *11*, 470. [[CrossRef](#)]
86. Wang, X.; Zhang, H.; Nyamesorto, B.; Luo, Y.; Mu, X.; Wang, F.; Kang, Z.; Lagudah, E.; Huang, L. A new mode of NPR1 action via an NB-ARC–NPR1 fusion protein negatively regulates the defence response in wheat to stem rust pathogen. *New Phytol.* **2020**, *228*, 959–972. [[CrossRef](#)]

- 
87. Zhu, H.; Li, C.; Gao, C. Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 661–677. [[CrossRef](#)]
  88. Chen, S.; Rouse, M.N.; Zhang, W.; Zhang, X.; Guo, Y.; Briggs, J.; Dubcovsky, J. Wheat gene Sr60 encodes a protein with two putative kinase domains that confers resistance to stem rust. *New Phytol.* **2020**, *225*, 948–959. [[CrossRef](#)]