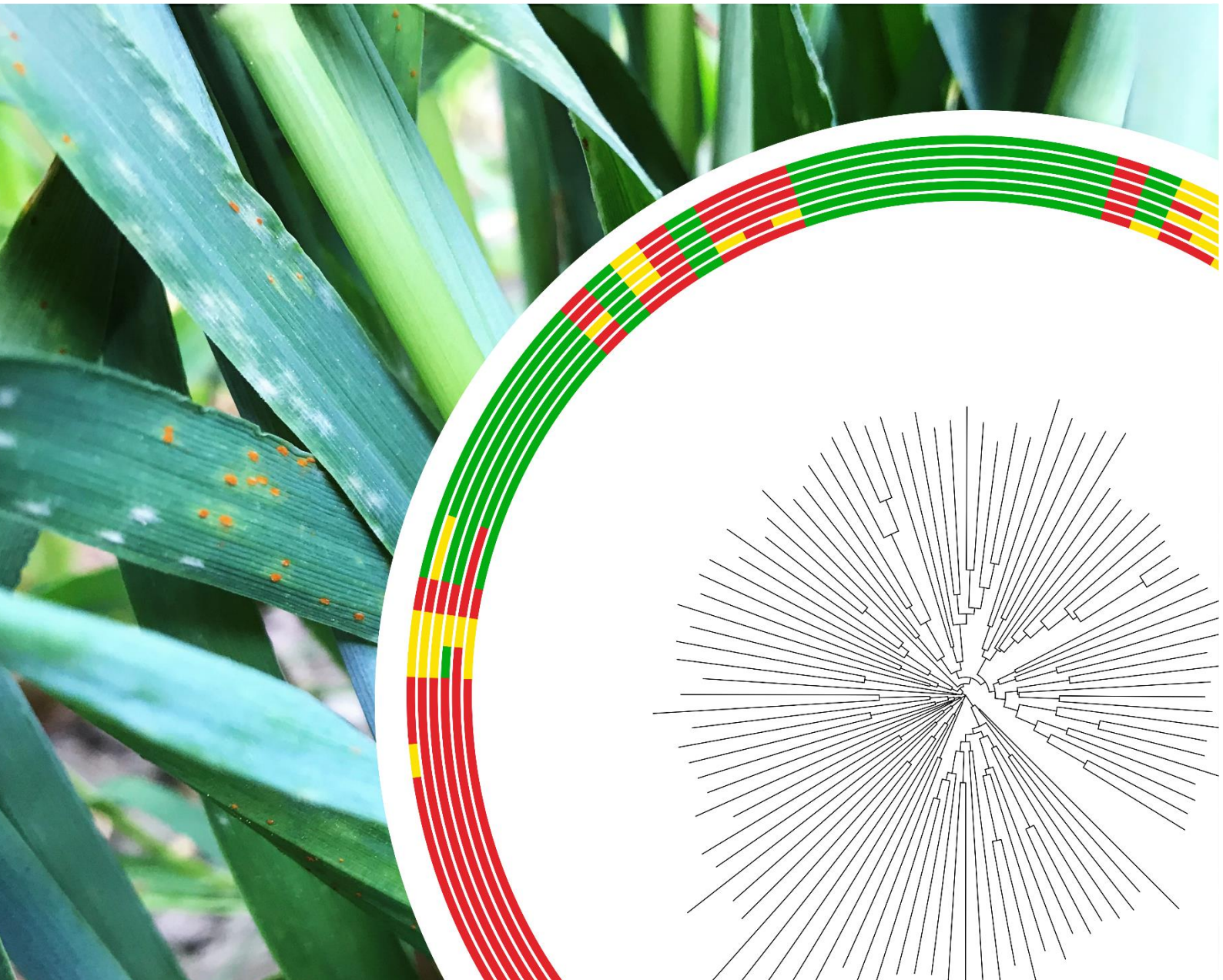
AARHUS UNIVERSITY
FACULTY OF TECHNICAL SCIENCES
DEPARTMENT OF AGROECOLOGY

Selection of rye (*Secale cereale* L.) for powdery mildew and leaf rust resistance through phenotyping, target sequencing, and association genetics

PhD Thesis

Nikolaj Meisner Vendelbo
2019 – 2021



Dedicated to my partner as a token for her unwavering support and band-aid for the battered ears

Aarhus University
Faculty of Technical Sciences
Department of Agroecology
Forsøgsvej 1
4200 Slagelse, Denmark

Nordic Seed A/S
Grindsnabevej 25
8300 Odder, Denmark

Thesis submitted in partial fulfilment of the requirements for Doctor of Philosophy (PhD) degree

Title:	Selection of rye (*Secale cereale* L.) for powdery mildew and leaf rust resistance through phenotyping, target sequencing, and association genetics
Subject area:	Plant Genetics and Breeding
Supervisor:	**Professor Dr. Ahmed Jahoor**, Head of Breeding and Research Nordic Seed A/S, Denmark
	Professor Dr. Mogens Hovmøller, Department of Agroecology, Faculty of Technical Sciences Head of the Global Rust Reference Center (GRRC) Aarhus University, Denmark
Co-supervisor:	**Dr. Jihad Orabi**, Head of Molecular Breeding Nordic Seed A/S, Denmark
	Dr. Annemarie Fejer Justesen, Senior Scientist Department of Agroecology, Faculty of Technical Sciences Aarhus University, Denmark
	Dr. Khalid Mahmood, Researcher, Bioinformatician in Genome Analysis Nordic Seed A/S, Denmark
Assessment Committee:	**Dr. Per Langkjær Gregersen**, Senior Scientist Department of Agroecology, Faculty of Technical Sciences Aarhus University, Denmark
	Professor Dr. Hermann Bürstmayr, Department of Agrobiotechnology and Crop Sciences Head of the Institute of Plant Breeding Head of the Institute of Biotechnology in Plant Production Vienna University of Natural Resources and Life Sciences, Austria
	Professor Dr. Niels Stein Department of Crop Science Center of Integrated Breeding Research (CiBreed), Division of Genomics of Plant Genetic Resources Georg-August-University Göttingen, Germany & Head of the Genomics of Genetic Resources research group Leibniz Institute of Plant Genetics and Crop Plant research, Germany

The project was financed by the Innovation Fund Denmark (Grant ID: 8053-00085B), Pajbjerg Foundation and Nordic Seed A/S.

Table of Contents

Preface

The dissertation has been submitted to the Graduate School of Faculty of Technical Sciences at Aarhus University, Denmark in order to fulfill the requirements of obtaining the degree 'Doctor of Philosophy'. The Industrial PhD was started in September 2018 and ended in September 2021 and was carried out in collaboration between the industrial partner, a Danish plant breeding company Nordic Seed A/S, and Aarhus University. The main industrial supervisor of the project was Professor Dr. Ahmed Jahoor, and main academic supervisor Professor Dr. Mogens Hovmøller. The industrial co-supervisors were Dr. Jihad Orabi and Dr. Khalid Mahmood, and academic co-supervisor Dr. Annemarie Fejer Justesen.

The PhD project included field trials, greenhouse trials, laboratory experiments and extensive bioinformatic analysis of collected data and took its basis in the Gülzow hybrid rye (*Secale cereale* L.) elite breeding germplasm at Nordic Seed. In this project the population structure, fertility control system and genetics underlying powdery mildew and leaf rust resistance were investigated in the germplasm. Outcome of the project were five manuscripts.

List of publications included in the dissertation

I. Vendelbo, N. M., Sarup, P., Orabi, J., Kristensen, P. S., & Jahoor, A. (2020). **Genetic structure of a germplasm for hybrid breeding in rye (*Secale cereale* L.)**. *PloS one*, *15*(10), e0239541.

II. Vendelbo, N. M., Mahmood, K., Sarup, P., Kristensen, P. S., Orabi, J., & Jahoor, A. (2021). **Genomic Scan of Male Fertility Restoration Genes in a 'Gülzow' Type Hybrid Breeding System of Rye (*Secale cereale* L.)**. *International Journal of Molecular Sciences*, *22*(17), 9277.

III. Vendelbo, N. M., Mahmood, K., Sarup, P., Hovmøller, M. S., Justesen, A. F., Kristensen, P. S., ... & Jahoor, A. (2022**). Discovery of a Novel Leaf Rust (*Puccinia recondita*) Resistance Gene in Rye (*Secale cereale* L.) Using Association Genomics.** *Cells*, *11*(1), 64.

IV. Vendelbo, N. M., Mahmood, K., Steuernagel, B., Wulff, B.B.H., Sarup, P., Hovmøller, M., Justesen, A.F., Kristensen, P. S., Orabi, J., & Jahoor, A. (2021). **Discovery of resistance genes in rye by long-read target sequencing and association genetics**. (Draft)

V. Vendelbo, N. M., Mahmood, K., Sarup, P., Kristensen, P. S., Orabi, J., & Jahoor, A. (2021). **Discovery of a novel powdery mildew (*Blumeria graminis*) resistance locus in rye (*Secale cereale* L.)**. *Scientific reports*, *11*(1), 1-15.

Acknowledgement

I would like to start by expressing my deepest appreciation to my company supervisor Ahmed Jahoor and co-supervisor Jihad Orabi. It has been a tremendous journey these past three years, both personally as well as professionally, and when times have been tough or difficult you have never wavered in your support and on several occasions demonstrated your willingness to make compromises for my wellbeing. To my office companion and co-supervisor Khalid Mahmood, thank you for all of the fruitful discussions and for your eagerness to share your experience, it has been a great support in times of doubt and distress.
I would like to extend my since thanks to my university supervisor Mogens Hovmøller and co-supervisor Annemarie Fejer Justesen for sharing their keen experience in the field of plant pathology.

To all of my colleagues at Nordic Seed, I would like to express my gratitude for the friendly work environment and hospitability, even when I had been away for extended periods I never felt foreign. Writing a detailed acknowledgement for all colleagues that have at some point aided me during the past three years, whether it be with the label printer, preparation of SeedMatic trays, adjustment of windows in the greenhouse etc., would be too extensive, but you all have my sincere gratitude. I would, however, like to thank laboratory technician Hanne Svenstrup for her kind assistance in the laboratory and contribution to the genotypic data collection. Special thanks to Pernille Sarup for her radiance of positive energy and excellence in bioinformatic analysis, without you I might still be stuck trying to solve a script in R. I would also like to extend my gratitude to all field technicians and assistant breeders that have been involved in the establishment and maintenance of field trials at Dyngby, Holeby and Nienstädt. I am also grateful to the rye breeder Peter Skov Kristensen and the rye breeding team at Nordic Seed Germany GmbH (Nienstädt) comprised of Johannes Hiller, Anette Deterding and Marlene Walbrodt for their effort in multiplication of lines for use in the field- and greenhouse trials.

I would also like to extend my sincere gratitude to laboratory technician Ellen Jørgensen at the Global Rust Reference Center, Aarhus University (Slagelse, Denmark) for her kind mentorship in the collection, recovery, inoculation and multiplication procedures of cereal rusts. Thanks to trial technician Jacob Sørensen at the Institute of Agroecology, Aarhus University (Slagelse, Denmark) for assistance in the multiplication of collected rust samples. Thanks to lead field

technician Anders Almskou-Dahlgaard at the Institute of Agroecology, Aarhus University (Slagelse, Denmark) for establishment of the field trials at Flakkebjerg.

For analysis of the RenSeq data I would like to extend my deepest gratitude to Burkhard Steuernagel, manager of the Informatic Platform at John Innes Centre (Norwich, UK). With Covid-19 pandemic abruptly ending my stay at the JIC I was adrift for a period and without your mentorship and profound knowledge in bioinformatic processing of RenSeq data I would not have been able to conduct the study. In extension, I cannot begin to express my thanks to Brande B.H. Wulff at King Abdullah University of Science and Technology (Thuwal, Saudi Arabia) for his seemingly unfathomable commitment to the RenSeq project as well as my PhD. It has been a great joy to get a glance 'behind the scene' of the research that your group conducts on the pinnacle of disease resistance improvement in crops. Coming to know you has been a great unanticipated reward from this project and I hope to nurture this relation for years to come.

To my family, while we are situated in either end of the country I have never felt alone in this pursuit, and I have brought your teachings of hard work, dedication and compassion through every step of this PhD. I now look forward to sharing more time with you all and savor the conclusion of this chapter.

To my in-laws I extend the deepest apologies for my absence in gardening, weeding, wood-working and other practical side-jobs on the country estate, and extend my profound gratitude for the stable and sound diets that you have provided me during this final sprint. To Merlin, an blue merle Australian shepherd thank you for your dedicated companionship during writing of the first manuscript, to Cooper, a black bi Australian shepherd, thank you for keeping me awake, and last but not least Sotho, a red merle Australian shepherd, thank you for distracting me and at times putting things into much needed perspective.

Lastly, to my partner, you have shown an exceptional patience, endurance and understanding through this entire endeavor, and your unwavering support has countless times been a life-line that keept me afloat. Looking back on the last three years I am so proud of the companionship that we have forged step by step and I am keen to see what new exciting chapter(s) waits around the corner for us.

Summary

Despite decades of declining rye acreage in Northern and Eastern Europe, rye remains an important crop in the European agriculture. Since the introduction of rye hybrids in 1990s breeding efforts have primarily focused on addressing the inherent issues of male-fertility restoration and ergot susceptibility in the predominant hybrid rye breeding gene pools Petkus and Carsten. As a consequence breeding for resistance towards the foliar diseases powdery mildew and leaf rust were concurrently given a lower priority. Capable of causing considerable grain yield and quality losses we investigated the genetics underlying resistance to powdery mildew and leaf rust in a Gülzow-based hybrid rye (*Secale cereale* L.) elite breeding germplasm distinct from the Petkus and Carsten. In the project commencement a population study of the germplasm was conducted, confirming a strong genetic separation of parental populations suited for hybrid breeding. The investigation, however, revealed a considerable discrepancy in the genetic characteristics of parental populations, emphasizing the need for addressing a low level of genetic diversity in the non-restorer germplasm (NRG) and cytoplasmic-male sterile (CMS) population. For the investigation of resistance a subset of 190 Nordic Seed hybrid rye breeding lines were genotyped on the state-of-the-art 600K high-density single nucleotide polymorphism array and phenotyped in field, and the restorer population (n = 101) in a greenhouse trial. In contrast to the Petkus and Carsten gene pools we observed a high level and diverse spectra of leaf rust resistance in both parental populations. Using genome-wide association study (GWAS) we identified a novel leaf rust resistance (*R*) gene on chromosome arm 7RS and five quantitative trait loci (QTLs) on chromosome arms 1RS, 1RL, 2RL, 5RL, and 7RS. Using *k*-mer association genetics coupled with resistance gene enrichment and sequencing (AgRenSeq) we, furthermore, identified four candidate leaf rust *R* genes, of which one co-localized with *Pr3* and the identified QTL on chromosome arm 1RS. For powdery mildew, as a result of a low level of natural infection in the field trials, a greenhouse trial was conducted using distinct field populations as inoculum leading to observation of a moderate level of resistance in the NRG&CMS population and low level of resistance in the restorer population. Using GWAS a we identified a novel powdery mildew *R* gene on chromosome arm 7RL.

On basis of the Danish official trials records during the past 25 years and current projections of powdery mildew on Triticeae under near-future climate conditions in Northern Europe, these suggest a likely minor role of powdery mildew as a biotic stress factor in rye. While powdery

mildew is not a bygone threat, it is likely to be confined to sporadic years, or geographical regions, with conducive climatic conditions or certain agricultural practices. In contrast, leaf rust is becoming a biotic stress factor of growing importance in rye with current projections concurring on an augmented effect of climatic changes on leaf rust on Triticeae in Northern and Eastern Europe. The increase in restrictions, prohibitions and environmental taxation on the use of pesticide as a result of the rising political incentive on a sustainable transition of agricultural practices in the EU emphasizes the need for improving the low level of leaf rust resistance in released hybrid rye cultivars. On basis of our findings the Gülzow germplasm, therefore, constitute a valuable genetic resource for addressing this potential demand. In order to improve the quality of markers for successful implementation of marker assisted selection and/or marker assisted backcrossing strategies for discovered leaf rust *R* genes and QTLs we propose (i) separation of GWAS on parental populations due to large discrepancy in linkage disequilibrium, (ii) increasing sample size and marker density for GWAS on the restorer population, (iii) establishment of multi-parent advanced generation inter-cross population(s) for discovery of less prevalent leaf rust *R* genes and slow rusting QTLs, and (iv) use of candidate leaf rust *R* gene sequences for potential development of high quality functional markers.

In addition to the study of powdery mildew and leaf rust resistance in the germplasm a investigation of the genetics underlying the Gülzow fertility control was likewise done. By GWAS on a biparental mapping population we identified a novel Gülzow-type major restoration of male-fertility gene on chromosome arm 3RL. Using the recent 'Lo7' rye reference genome we identified a mitochondrial transcription termination factor co-expressed in two Gülzow-type hybrids during flowering as a likely candidate gene. With no previous record of a major restoration of male-fertility gene in fertility control systems of rye on chromosome arm 3RL nor on syntenic regions in barley and wheat our finding constitute a novel discovery with potential value for implementation in other Triticeae hybrid breeding systems

Sammendrag

På trods af at det dyrkede areal af rug i Nord- og Østeuropa har været aftagende i flere årtier, så forbliver rug en betydelig afgrøde i det Europæiske landbrug. Siden introduktionen af rughybrider i 1990erne har forædlingsindsatsen primært adresseret de iboende problemer med restaureringen af hanfertiliteten og modtagelighed overfor meldrøjer i de hyppigst anvendte hybridrugs genpuljer, Petkus og Carsten. Som et følge af dette har forædling af resistens mod bladsygdommene meldug og brunrust været givet en lavere prioritet. Da disse sygdomme kan forårsage et betydelig udbytte- og kvalitetstab, undersøgte vi resistens overfor meldug og brunrust i en Gülzow-baseret hybridrugs forædlingsmateriale ved Nordic Seed afvigende fra Petkus og Carsten genpuljerne. I projektets opstart udførte vi et populationsstudie af forædlingsmaterialet og bekræftede her en betydelig genetisk differentiering af forældrepopulationerne, som på baggrund heraf er velegnede til hybridforædling. I studiet fandt vi ydermere en betydelig differens mellem forældrepopulationernes genetiske karakteristika samt et behov for at adressere den lave genetiske diversitet i non-restorer germplasm (NRG) og cytoplasmic-male sterile (CMS) populationen. Til undersøgelsen af resistens blev et udvalg af 190 forædlingsliner genotypet på en avanceret 600K høj-densitets enkeltnukleotidpolymorphi array og bedømt for resistens i markforsøg og for restorer populationen (n = 101) også i et væksthusforsøg. I kontrast til Petkus og Carsten genpuljerne observerede vi et højt niveau og diverse spektre af brunrustresistens i begge forældrepopulationer. Ved brug af helgenomsassociationsstudier (GWAS) identificerede vi et nyt brunrustresistens (*R*) gen på kromosom arm 7RS og fem kvantitative karakter loci (QTL) på kromosomarmene 1RS, 1RL, 2RL, 5RL, og 7RS. Ydermere, ved brug af subsekvens (*k*-mer) associationsgenetik koblet med resistensgenberigelse- og sekventering (AgRenSeq) identificerede vi fire brunrust *R* kandidatgener, hvoraf én co-lokaliserede med det kendte brunrust *R* gen *Pr3* og den identificerede QTL på kromosom arm 1RS. For meldug var vi nødsaget til, grundet et lavt naturligt smitteniveau i markforsøgene, at bedømme forædlingslinjerne i et væksthusforsøg ved brug af indsamlede markpopulationer som smittestof. Vi observerede her et moderat niveau af meldugresistens i NRG&CMS-populationen og lavt i restorerpopulationen. Ved brug af GWAS identificerede vi et nyt meldug *R* gen på kromosom arm 7RL.

På baggrund af data fra de danske landsforsøg gennem de sidste 25 år samt nuværende prædikteringsmodeller af meldug på Triticeae under nær-fremtidige klimatiske forhold i Nordeuropa så indikeres det, at meldug som biotisk stressfaktorer i rug sandsynlig vil have en

lille betydning i fremtiden. Forekomst af meldug i rug vil dog ikke forsvinde, men sygdommen vil i stedet være begrænset til atypiske år med gunstige klimatiske forhold eller under særlige landbrugspraksis. I kontrast, er brunrust sandsynligvis en biotisk stressfaktorer af tiltagende betydning for dyrkning af rug i Nord- og Østeuropa med alle nuværende prædikteringsmodeller enige om en forstærkende effekt af klimatiske forandringer på forekomsten og betydningen af sygdommen. De tiltagende forbud, restriktioner og miljøafgifter på brug af pesticider, som følge af et stærkt politisk incitament for en bæredygtig omstilling af landbrugspraksis i EU, understreger nødvendigheden for at adressere og forbedre det lave nuværende niveau af brunrustresistens i sortsudvalget af hybridrug. På baggrund af vores observationer udgør Gülzow forædlingsmaterialet en værdifuld genetisk ressource til at adressere dette behov. For at forbedre kvaliteten af identificerede markører til succesfuld implementering af markørassisteret selektion og/eller markørassisteret tilbagekrydsning for de fundne brunrust *R* gener og QTLs forslår vi (i) separering af GWAS på de enkelte forældrepopulationer grundet en stor difference i 'linkage disequilibrium', (ii) forøgelse af linjeantal og markørdensitet for GWAS på restorerpopulationen, (iii) etableringen af multi-parentale avancerede generations inter-cross (MAGIC) population(er) for at identificere mindre hyppige, og muligvis betydningsfulde, brunrust *R* gener og QTLs, og (iv) at gøre brug af identificerede brunrust *R* kandidat gensekvenser til at udvikle mulige højkvalitets, funktionelle markører.

Ud over undersøgelse af meldug og brunrustresistens i forædlingsmaterialet undersøgte vi ligeledes genetikken bag Gülzow fertilitetskontrolsystemet. Ved brug af GWAS på en biparental F_2 population identificerede vi et nyt dominant Gülzow-type restaurering af hanfertilitetsgen på kromosomarm 3RL. Herefter, ved hjælp af det nylige 'Lo7' rugreferencegenom, identificerede vi en mitochondrial transcription termination factor udtrykt i to Gülzow-hybrider ved blomstring som et sandsynlig kandidatgen. Uden eksisterende litteratur på dominante restaurering af hanfertilitetsgener på kromosomarm 3RL i rug eller synteniske regioner i byg eller hvede, udgør vores fund en ny opdagelse med potential værdi for implementering i andre Triticeae hybrid forædlingssystemer.

Abbreviations

A	
AgRenSeq	Association genetics RenSeq
APR	Adult plant resistance
B	
Bgh	*Blumeria graminis* f. sp. *hordei*
Bgs	*Blumeria graminis* f. sp. *secalis*
Bgt	*Blumeria graminis* f. sp. *tritici*
C	
CC	Coiled-coil domain
CCS	Circular consensus sequence
CMS	Cytoplasmic male sterility
CNL	CC-NLR
CNV	Copy-number variation
D	
DAMPs	Damage associated molecular patterns
DAI	Days after inoculation
DAS	Days after sowing
E	
ETI	Effector triggered immunity
F	
F_{is}	Inbreeding coefficient
F_{st}	Fixation indices
G	
GWAS	Genome-wide association study
G-type CMS	Gülzow type CMS
H	
H_o	Observed heterozygosity
H_s	Within population gene diversity
HR	Hypersensitive response
HiFi	High fidelity
I	
Indel	Insertion-deletion
IT	Infection type
K	
k-mer	Subsequence of *k* nucleotides
L	
LD	Linkage disequilibrium
LDD	Long distance dispersal
LRR	Leucine-rich repeat domain
M	
MABC	Marker-assisted backcrossing
MAF	Minor allele frequency
MAGIC	Multi-parent advanced generation inter-cross
MAS	Marker-assisted selection
MLM	Mixed-linear model
MPP	Multi-parent population
mTERF	Mitochondrial transcription termination factor
N	
NAM	Nested association mapping
NB	Nucleotide-binding domain
NB-ARC	Nucleotide-binding adaptor shared by APAF-1, R proteins, and CED-4 domain
NLR	Nucleotide-binding leucine-rich repeat protein (or NB-LRR)
NLR-ID	NLR gene with integrated domain
NRG	Non-restorer germplasm (or seed mother)
P	
PAMP	Pathogen-associated molecular pattern
PAV	Presence/absence variation
PIC	Polymorphism information content
PPR	Pentatricopeptide repeat protein
Prs	*Puccinia recondita* f. sp. *secalis*
PRR	Pattern recognition receptors
PTI	PAMP-triggered immunity
P-type CMS	Pampa type CMS
Q	
QTL	Quantitative trait loci
R	
R	Restorer (or pollen father)
RenSeq	Resistance gene enrichment sequencing
Rf	Restoration of male-fertility
RFL-PPR	Restoration of male-fertility like PPR protein
RIL	Recombinant inbred line
RLK	Receptor-like kinase
RLP	Receptor-like protein
RNL	RPW8-like CC-NLR
R gene	Resistance gene
S	
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SMRT	Single-molecule real-time
SPI	Single pustule isolate
T	
TIR	Toll-interleukin 1 receptor
TNR	TIR-NLR

General Introduction

1. Rye

1.1 Origin of rye and history of its domestication

About 12,000 years ago western agriculture was born in the Near East region as predecessors of the human race began shifting from a nomadic lifestyle to living in sedentary agricultural communities. Instead of depending on hunting game and gathering wild vegetables and fruits neolithic pioneer farmers cultivated wild progenitors of modern cereal species (Hillman, 1978). Study of archaeobotanical remains have unveiled the subtle unconsciously-applied selection of plants on basis of seed size, ear rachis stiffness and brittleness of glumes in the onset of crop domestication (Salamini et al., 2002). The early domestication of traits central to the agricultural habitat in wild progenitors has been termed 'domestication syndrome' (Sakuma et al., 2011).

In cereals, wild progenitor of founder crops included barley (*Hordeum vulgare* subsp. *spontaneum*), the diploid ancestors of wheat, emmer wheat (*Triticum turgidum* subsp. *dicoccum*), and einkorn wheat (*Triticum monoccum* subsp. *baeoticum*) (Weiss and Zohary, 2011). During this period, wild rye (*Secale cereale* subsp. *vavilovii*) is commonly believed to have invaded the cultivated fields of founder crops as a "hitchhiking" weed, co-harvested, and hence indirectly acquiring the key domestication traits (Weiss et al., 2012; Maraci et al., 2018). The adaptation-strategy of rye to survive in the prehistorian agrarian society has been referred to as 'Vavilovian mimicry' (McElroy, 2014). Originating from these secondary domesticate events, rye (*Secale cereale* L.) hence adapted to the agricultural habitat. First macrobotanical evidence of deliberate cultivation of rye as a crop on its own right were in North-Central Anatolia during the Mesolithic period 3-4000 B.C (Gökgöl, 1944; Hillman, 1978).

While a plethora of routes by which rye migrated into Europe have been proposed, general consensus is that rye migrated alongside cereal founder crops from the Near East to the central Europe via Anatolia and the Balkans (Behre, 1992). The early expansion of rye during the Neolithic period in central Europe was likely subtle, transitioning from rye as an admixed weed to a deliberate crop.

In the East Baltics region earliest macrobotanical remains of rye has been dated to 4th century A.D., with first observation of rye pollen in palynological records dating as far back as 3000 B.C. (Czeczuga and Kossacka, 1973; Gałka et al., 2015). Identification of rye pollen in palynological samples can be done with high confidence due to its unique physiological features

compared to other cereal pollen including prolate grain shape, size and subpopular position of the pore (Köhler and Lange, 1979). Rise in rye pollen amount during the roman period is believed to mark the onset of deliberate rye cultivation as a staple crop in the East Baltics (Szal et al., 2014; Grikpėdis and Matuzevičiūtė, 2016).

In Germany the earliest findings of rye pollen in palynological studies have been date to 3000 B.C. with evidence of deliberate cultivation around 100 B.C. during the roman period (Jahns, 2000; Rösch, 2000; Jahns, 2007). In Northern Germany findings have suggested that rye constituted the dominant staple cereal cultivated from 8th century A.D. and onward (Beug, 2011).

In Denmark earliest findings rye pollen in palnylogical studies have been dated to the 5th century B.C (Grabowski, 2013). Increase in rye pollen recovery and findings of domesticated rye plants in iron smelting furnaces suggest that rye were an central component in the Danish agriculture from 4th century A.D. and onward (Henriksen, 2003; Grabowski, 2013).

While migrating as a 'blind passenger' amongst founder crops to modern wheat and barley, rye became a prominent member of the prehistoric European agricultural landscape. Analysis of medieval and post-medieval cesspits reveal that rye constituted a prominent staple food alongside wheat and barley in Northern Europe from 11th century and onward (Märkle, 2005; Yoder, 2012).

1.2 Taxonomy of the Secale genus

11 million years ago the ancestral Triticeae progenitor diverged, giving rise to the wild progenitors of barley (*Hordeum vulgare* L.) and an intermediate rye-wheat ancestor (Huang et al., 2002). Successively the intermediary ancestor diverged 7 million years ago giving rise to the wild progenitors of wheat (*Triticum aestivum* L.) and rye (*Secale cereale* subsp. *cereale*).

The systematics of the *Secale* genus has undergone several taxonomical revision during the years with the latest by Frederiksen and Petersen (1998) based on morphological characteristics. Their taxonomy was later confirmed by Schreiber et al. (2019), who conducted a larger single nucleotide polymorphism (SNP) genotype based population study on cultivated rye and wild relatives (Figure 1.2).

Rye belongs to a small genus *Secale* with only three taxa, the annual wild *S. silvestre*, the perennial wild *S. strictum* (formerly *S. montanum*) and *S. cereale*. The *S. strictum* contained subspecies *strictum*, *africanum*, *anatolicum* and *kuprijanovii*. The *S. cereale* containes

subspecies *ancestrale*, *segale*, *dighoricum*, cultivated rye subsp. *cereale*, and the wild weedy rye subsp. *vavilovii*.

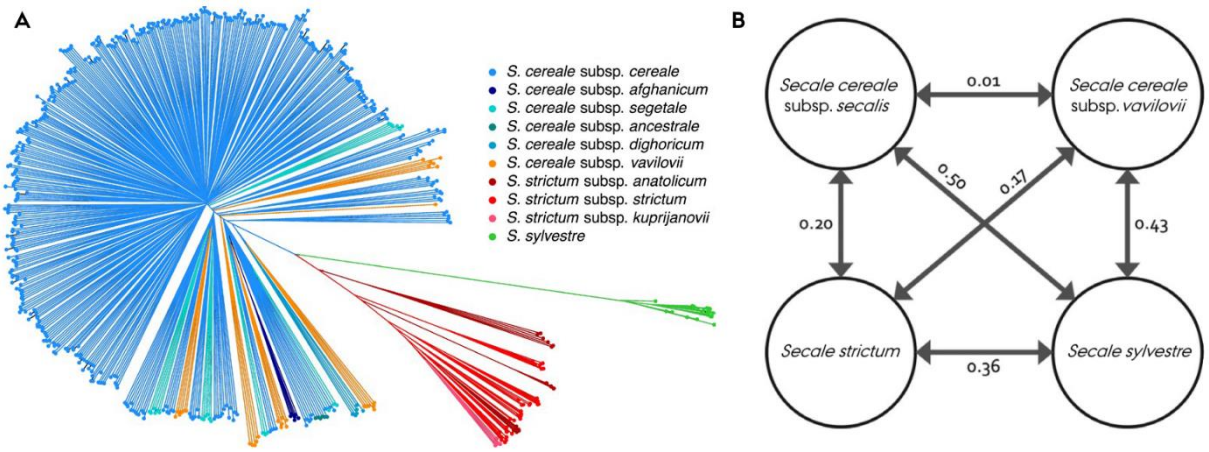


Figure 1.2 Phylogenetic relationship of the *Secale* genus **(A)**, and genetic differentiation (**B**) (Schreiber et al., 2019)

1.3 Rye genetics and biology

Rye is recognized by several conspicuous morphological features distinguishing it from its cereal relatives (Schlegel, 2014). With some varieties of rye capable of reaching 300 cm in height, rye is the tallest of cereals (Figure 1.3). The spikes of rye are long (10-19 cm) and slender with medium-long awns in comparison with wheat (short) and barley (long) (Schlegel, 2014). The vegetation of most rye cultivars and population varieties exhibit a blueish-gray shade caused by a waxy film on the surface. Amongst the cultivated cereal species in Triticeae rye is the only outbreeder, also known as allogamous or cross-pollinator, incapable of fertilization by self-pollen (Lundqvist, 1956). Self-incompatibility constitutes an efficient strategy in flowering plants for promoting outcrossing and maintaining high levels of heterozygosity. In rye, self-incompatibility is controlled by the complementary action of two gametophytic expressed genes, *S* and *Z* (Hackauf and Wehling, 2005). In case both alleles are expressed in the haploid pollen a self-incompatibility response occurs.

Distinguished by its high level of inherent tolerance to abiotic and biotic stresses, rye outperforms wheat and barley under marginal conditions. In Denmark, rye has, therefore, predominantly been cultivated on infertile sandy soils in West Jutland frequently exposed to drought conditions during the growing season. In a recent study by Schittenhelm et al. (2014), rye was observed to exhibit a yield reduction of 54-65% under severe drought conditions,

producing 3.9-4.2 t/ha compared to 3.1 – 3.9 t/ha in triticale, barley and wheat. Adapted to a winter habitat, rye requires a vernalization period to stimulate plants in the vegetative stage to transition to reproductive growth. As a result rye has acquired a high level of cold hardiness surpassing that observed in winter wheat which likely played a major role in the prehistoric expansion of rye cultivation in the northern latitude (Dvorak and Fowler, 1978; Janmohammadi et al., 2018). Lastly, rye is exceedingly tolerant to high levels of free aluminum released in acidic soils which are estimated to cover 1.6 billion hectares of land worldwide (Kim et al., 2002; Santos et al., 2019).

Figure 1.3 Personal picture in the field, with 'tall' traditional type hybrid rye cultivar in the background (cv. Helltop, 180-210 cm), and dwarf-type inbred lines in front

Resulting from the close relationship, rye shares a lot of characteristics with the self-fertilizing cereals barley and wheat (Schreiber et al., 2019). All Triticeae have a haploid set of seven chromosomes and highly collinear, repeat-rich genomes. The collinearity of orthologous chromosomes is, however, complicated by the series of recurrent translocation events that occurred during Triticeae speciation (Devos, 2005). Recent studies have outlined the syntenic genomic regions across rye, wheat, and barley genomes enabling the inference of orthology (Martis et al., 2013; Li et al., 2021). Rye exhibits the largest genome amongst its cereal relatives with 7.9 Gb DNA sequences (Rabanus-Wallace et al., 2021) compared to 4.79 Gb in barley (Mascher et al., 2017) and sub-genomes of hexaploidy bread wheat ranging from 3.95 to 4.94 Gb (Appels et al., 2018).

1.4 Cultivation of rye and its uses

Until mid-20th century, rye constituted a central component of the Northern European agricultural landscape and a staple cereal. Records dating back to 1961 show the dramatic decrease of the total area cultivated with rye in Northern- and Eastern Europe, from 7.07 million ha to 1.79 million ha in 2019 (Figure 1.4A, FAO, 2021). In Denmark, area that had earlier been

cultivated with rye was largely substituted by wheat. Wheat, that had previously been a smaller cereal crop in the Danish agriculture, expanded from 105.000 ha in 1961 to its top at 739.000 ha in 2010, majority of which is sown as winter crop (FAO, 2021). Barley likewise experienced an expansion from 1960 to its top at 1.62 million ha in mid-1980, now having decreased to 621.700 ha in 2021, majority of which is sown as spring crop (DST, 2021). From a historical low-point of 29.700 ha in 2006, cultivation of rye in Denmark has experienced a resurgence in the recent decade, covering 144.900 ha in 2019 (Figure 1.4B). In 2021, rye comprised 12.7% of the winter crop sown in Denmark, and 4.1% of the total cultivated area (DST, 2021). While a minor crop in Denmark, cultivation of triticale (× *Triticosecale* Wittmack) expanded during mid-1980s in Northern- and Eastern Europe, with 1.3 mil. ha in Poland and 358.000 ha in Germany in 2019, having substituted vast areas previously cultivated with rye (FAO, 2021).

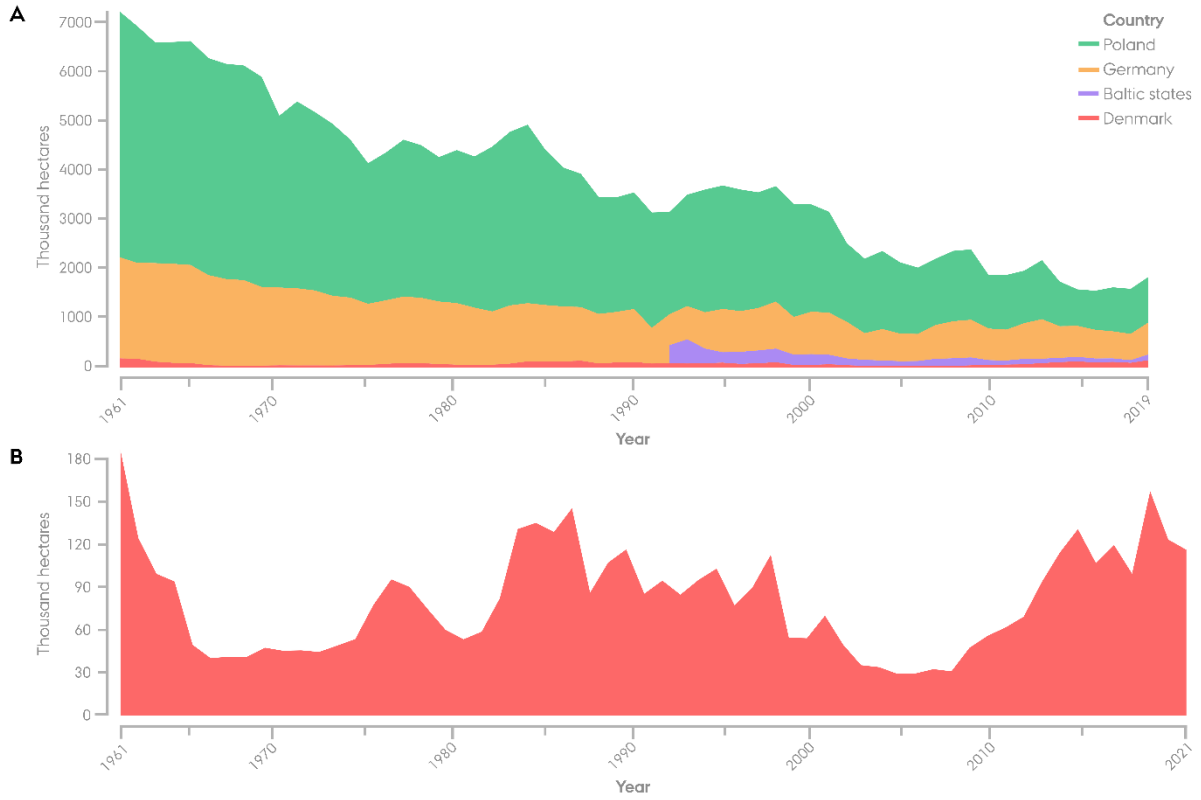


Figure 1.4 Cultivation of rye (*Secale cereale* L.) in **A**) Northern- and Eastern Europe from 1961 – 2019. **B**) Denmark from 1961-2021 (DST, 2021; FAO, 2021)

Since 2000 the organic production has increased dramatically in Europe, with 8.5% of the total utilized agricultural area in EU under organic farming in 2019 (FiBL, 2021). In 2030 the defined objective, published under the EC Farm to Fork and Biodiversity Strategy as part of the European Green Deal, is to have reached a minimum of 25% of EUs agricultural land under organic farming (EC, 2021a). In Denmark the organic production constitutes 11.7% of the cultivated land in 2020, having more than doubled since 2007 (LBST, 2020a). The organic

production of cereals in Denmark covers 100.331 hectare, of which rye constitutes 22.940 hectare corresponding to 22.9% of organic cultivated cereals (LBST, 2020a).

The decline in rye cultivation from mid-1980 coincides with the first release of hybrid rye cultivars in Denmark. In 2013 hybrid cultivars comprised 90% of rye sown in Denmark (Eriksen, 2014). While hybrid rye outperformed the previous population rye varieties, the early hybrid cultivars were prone to lodging and severe infection of ergot disease. Ergot, caused by the fungal pathogen *Claviceps purpurea,* infects the rye ovary leading to development of dark purple sclerotia structures (Miedaner and Geiger, 2015). Containing toxic alkaloids the ergot sclerotia constitutes a serious contaminant of rye grain (Mainka et al., 2007). Currently the ergot contamination threshold set in the EU in unprocessed cereals is 0.05% by weight for human consumption and 0.1% by weight for animal consumption (EC, 2002; EU, 2012). While recent years intense breeding effort has dramatically improved hybrid rye cultivars resistance towards lodging and ergot disease, its history remains a shadow on the standing of hybrid rye in some agricultural communities.

In Denmark rye is used either as animal feed, primarily pigs, or for human consumption. Pig production constitutes a large component of the agricultural production in Denmark, with 17-18 million pigs produced annually (DST, 2021). While pigs were primarily fed a mixture of barley and rye in mid-19th century, modern feed mixtures in Denmark is primarily composed of wheat and barley (Senior, 1947). However, rye has in recent years received a renewed interest as a low input, cost-effective feed source for pigs. Schwarz et al. (2014) demonstrated that rye constitutes a superior feed source to barley, enhancing growth rate and carcass quality in some cases when used for pig fattening while additionally reducing the feed cost (Schwarz et al., 2016). In poultry, rye has similarly been demonstrated to constitute a cost-effective feed source both during the feeding phase as well as for broilers up to a dietary level of 20% (Bederska-Łojewska et al., 2017; Bederska-Łojewska et al., 2019).

To this day rye remains a key constituent in modern Northern European diet and major single source of whole grain in Denmark (Kyro et al., 2012). The rye kernel endosperm contains a high level of dietary fibers (20% dry matter) and with subsequent high extraction rates in Denmark, majority of these fibers remain in the flour used for making the traditional ryebread (Frolich et al., 2013). Intake of whole grain has been found to be associated with a wide range of positive health effects such as a decrease in risk of adiposity, cardiovascular diseases, type 2 diabetes and improved gut microbial profile in recent studies (Andersson et al., 2010; Ounnas et al., 2016; Jonsson et al., 2018; Kyro et al., 2018). Thus, authorities of Denmark has issued a recommended intake of whole grain as part of their ten food-based dietary guideline (MVFM, 2021)

2. Hybrid breeding in rye

As an self-incompatible outbreeder rye has been cultivated in highly heterogeneous and genetically diverse populations since the onset of deliberate rye cultivation in Anatolia 3-4000 BC (Hillman, 1978). Cultivated in spatial secluded areas, rye populations became highly adapted to the local environment in which they were grown, and unconsciously bred, leading to the development of distinct provincial landraces (Persson and von Bothmer, 2002).

During the mid-19th century the first targeted selected of rye was done at the Probsteier Seed Cooperation in Northern Germany. Shortly thereafter rye breeding programs were founded in Germany, Estonia and Russia, marking the commencement of systematic rye breeding (Schlegel, 2016). Most renown during this period were the population varieties developed at Petkus in Central Germany by F. von Lochow based on the Probsteier gene pool, later known as the Petkus gene pool (Schlegel, 2014). In 1908 Rudolf Carsten started a rye breeding program near Kiel in Central Germany based on a rye landrace 'Heinrich Roggen' leading to development of the Carsten gene pool, distinct from Petkus (Miedaner and Laidig, 2019). The introduction of these improved populations of rye varieties such as the Petkus rye in Denmark during the late 19th century led to a gradual replacement of the previously local adapted landraces (Borgen and Grupe, 2012).

2.1 Establishment of hybrid breeding systems

In 1877, a cross between two population varieties 'Swedish Snow rye' and 'Correns rye' was observed to develop more tillers and higher grain yield. Observations like these in the early commencement of rye breeding were the onset of pursuing a controlled system for breeding hybrids in rye (Schlegel, 2016). Few decades later the theorem of hybrid vigor or heterosis observed in these crosses was introduced by George Shull in 1914 (Shull, 1948)

"*The physiological vigor of an organism as manifested in its rapidity of growth, its height and general robustness, is positively correlated with the degree of dissimilarity in the gametes by whose union the organism was founded*" (Shull, 1948)

The biological basis of heterosis has since been studied intensively to understand its role in the evolution and dynamics of natural ecosystems, and how it can be exploited in crop production (Chen, 2013b; Schnable and Springer, 2013).

In self-pollinating plant species, the increased homozygosity have exposed recessive deleterious alleles to natural selection, progressively purging these from the genetic load during

evolution (Hedrick and Garcia-Dorado, 2016). In outbreeding plant species, however, recessive alleles have remained masked due to their highly heterogeneous nature. If repetitively self-fertilized, fitness of outbreeding plant species will, therefore, dramatically decrease due to the genetic load of deleterious alleles expressed at increasing levels of homozygosity, eventually leading to a state of inbreeding depression (Hedrick and Garcia-Dorado, 2016). Crossing with a genetically distinct plant will then restore high levels of heterozygosity in the hybrid progeny giving cause to heterosis by masking the effect of deleterious alleles in the genetic background (Chen, 2013b).

Establishment of a controlled system for breeding of hybrids in rye required several key steps, including identification and introgression of self-compatibility and an efficient fertility control system (Chang et al., 2016). In 1956, self-incompatibility in rye was determined to be controlled by two multiallelic loci, S (1RS) and Z (2RL) (Lundqvist, 1956; Hackauf and Wehling, 2005). Using trisomic of an rye inbred line cv. 'Esto', Melz et al. (1990) identified four dominant self-fertility genes *Sf1* (1R), *Sf2* (5R), *Sf3* (4R), and *Sf4* (6R) (Melz et al., 1987). While several hypothesis have been proposed to explain self-fertility in rye, general consensus is that these are either alleles of self-incompatibility genes *S* and *Z*, or modifying genes (Fuong et al., 1993; Voylokov et al., 1993; Voylokov et al., 1998). The introgression of self-fertility genes enabled the development of inbred lines with a low extent of heterozygosity.

In order to steer the crossing of parental lines during large-scale field multiplication of hybrid rye seed a fertility control system is necessary (Chang et al., 2016; Kim and Zhang, 2018). The system comprises of three components; a cytoplasmic male sterile (CMS) seed mother, a non-restorer germplasm (NRG) capable of maintaining the CMS, and a fertile pollen father lines capable of restoring male-fertility in the progeny (Schlegel, 2014). The male sterilizing factor in CMS lines are a product of mitochondrial genes that causes a defect in the production of viable pollen (Gaborieau et al., 2016). The restorer lines instead carry nuclear restoration of male fertility (Rf) genes that suppress the CMS conferring mitochondrial transcript by interfering with their expression through induced processing or cleavage (Melonek et al., 2021). Lastly, the NRG line is characterized by a fertile cytoplasm and absence of dominant Rf genes, allowing it to maintain the CMS while not restoring the male fertility in progeny.

In 1970, the first male sterile lines were identified in an Argentinian primitive rye 'Pampa', leading to development of the predominant P-type CMS system (Geiger and Schnell, 1970). Shortly thereafter lines capable of restoring male fertility in the P-type CMS system were

identified (Geiger, 1972). In the P-type CMS system, major restoration of male-fertility (Rf) genes have been identified on 1RS, 4RL (*Rfp1, Rfp2, Rfp3*) and a dominant modifier on 6R chromosome, with minor genes on chromosome arms 3RL, 4RL and 5R chromosome (Miedaner et al., 2000; Stracke et al., 2003; Hackauf et al., 2012; Hackauf et al., 2017; Niedziela et al., 2021). Since then several additional CMS systems have been identified in rye. These include the Gülzow (G) type originating from a Austrian population variety 'Schlägler alt' (Melz and Adolf, 1991), R-type originating from a Russian population (Kobylianskii, 1971), and C- (Lapinski, 1972) and S type (Madej, 1975) originating from a Polish population variety 'Smolickie'. The G-type CMS system was developed at the Plant Breeding Institute Gülzow-Gustrow in Northern Germany after comprehensive screenings of international collections during the early 1960s (Melz et al., 2003).

In the G-type, one major Rf gene have been identified on chromosome arm 4RL (*Rfg1*) and two minor on chromosome arm 3RL (*Rfg2*) and 6R chromosome (*Rfg3)* (Melz and Adolf, 1991; Melz et al., 2003). In the C-type, one major Rf gene have been identified on chromosome arm 4RL (*Rfc1*) (Stojalowski et al., 2011).

With introgression of discovered self-fertility alleles in elite background a hybrid breeding system deploying the P-type CMS system was established at Hohenheim university in 1970 (Geiger and Miedaner, 1999). Shortly thereafter breeding companies established hybrid rye breeding programs and the first commercial hybrid rye cultivar 'Akkord' was tested in the Danish official trials in 1987 (Skriver, 1987). In rye all released hybrids are developed following the formula $(A_{CMS} \times B_{NRG}) \times Syn_{Rf}$ (Geiger and Miedaner, 2009). The seed parent is a single-cross hybrid with components originating from the same (A × A) or secluded (A × B) gene pools distinct from the synthetic restorer used as second pollinator. Unlike in maize, rye hybrids are three-way crosses due to high levels of inbreeding (Becker et al., 1982; Saleh et al., 2002).

2.2 Breeding progress and the introduction of hybrids

In Denmark and Germany population varieties were progressively substituted by hybrid rye cultivars after their introduction in late 1980s. In Denmark hybrids comprised 25% of rye cultivated in 1992, and 90% in 2013 (Pedersen, 1996; Eriksen, 2014). In Germany hybrid cultivars already comprised 44% of rye grain harvested in 1994, and 81% in 2014 (Seibel and Weipert, 1994; Laidig et al., 2017). In Poland the first hybrid cv. was introduced in 1995, and based on available literature, hybrid rye did not experience the same progressive expansion here

as in Denmark and Germany (Arseniuk and Oleksiak, 2003). Rye is cultivated as a low-input crop in Poland and the Baltic states, and farmers apprehension to use of hybrids is likely foremost the increased cost and dependency on buying hybrid sow seed (Arseniuk and Oleksiak, 2003; Hardzei and Urban, 2003).

During the period 1961 to 2019 the average yield of rye in Denmark increased from 2.8 to 6.5 t/ha, and in Germany 2.0 to 6.1 t/ha (Figure 2.2, FAO, 2021). Concurrently, yield increased from 1.4 to 3.2 t/ha in Poland and 1.7 to 3.7 t/ha in the Baltic states. The considerable yield difference between Northern European (Germany, Denmark) and Eastern European (Poland, Baltic states) rye producing countries is a product of multiple factors, including highly different agricultural systems and production intensity (Schils et al., 2018). Another key contributing factor is likely the predominance of hybrid rye cultivation in Northern Europe and population varieties in Eastern Europe. In a comprehensive investigation of breeding progress during 1989 to 2014 in Germany, Laidig et al. (2017) reported a 23.3% increase in grain yield of tested rye hybrids from 8.1 to 10.0 t/ha, compared to 18.8% increase in population varieties from 7.2 to 8.5 t/ha. At the trial terminus in 2014 the relative yield difference between rye hybrids and population varieties were 19%. In the Danish official trial in 2020, the top yielding hybrid produced 11.4 t/ha compared to the 8.6 t/ha in the only population variety tested, resulting in a 25% yield difference (Sortsinfo, 2021).

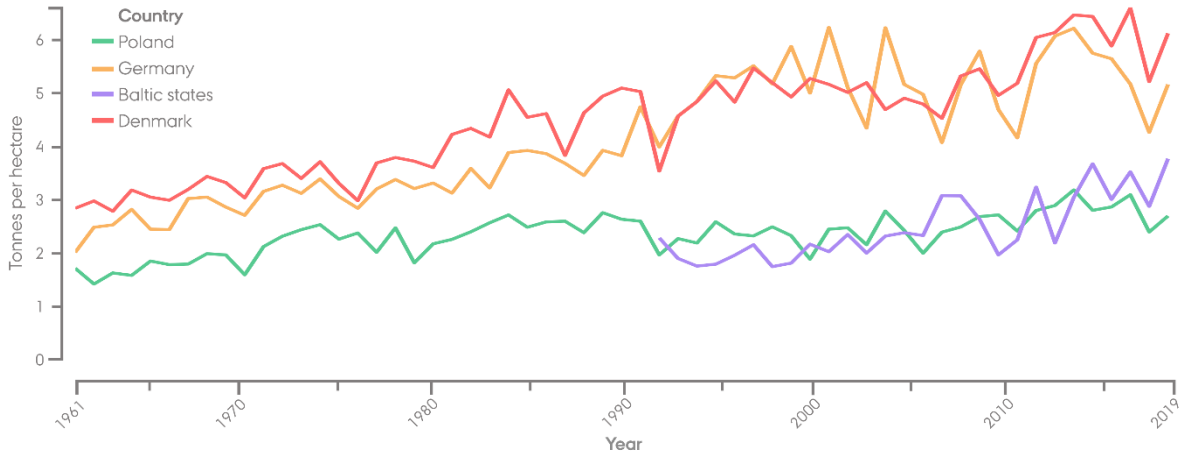


Figure 2.2 Yield progression of rye (*Secale cereale* L.) in Northern- and Eastern Europe from 1961 – 2019 (FAO, 2021)

2.3 Exploiting heterosis

The transition of rye breeding from population varieties to hybrid cultivars offered opportunities for greater uniformity, higher selection intensities and the exploitation of heterosis (Schlegel,

2016). Inbred lines in rye respond strongly to changes in heterozygosity with hybrids demonstrating a heterosis effect on all development and yielding characteristics (Miedaner and Laidig, 2019). While few studies have quantified the precise effect of heterosis in rye hybrids a heterosis effect of 110 – 140% on grain yield relative to midparent performance have been reported (Geiger and Miedaner, 1999).

In contrast, the effect of heterosis in maize (*Zea mays* L.) hybrids, a system highly similar to rye, has received ample scientific attention (Labroo et al., 2021). In a comprehensive study by Flint-Garcia et al. (2009), the effect of heterosis was investigated on seventeen traits in 300 single cross hybrids in maize. On most of the vegetative and reproductive traits a low-intermediate midparent heterosis (MPH) effect of 4 to 30% was observed. Biomass productivity showed the highest MPH effect of 185%, followed by yielding characteristics from 30-144% with the maximum being total grain weight (Flint-Garcia et al., 2009). Similar observations have been made in several successive studies (Yu et al., 2020; Meena et al., 2021). The effects of heterosis on tolerance to abiotic stresses in maize hybrids have also been widely documented (Chairi et al., 2016). Under severe drought conditions MPH of up to 210% have been reported on grain yield, and similarly 170% MPH under nitrogen deficit conditions (Riache et al., 2021). The enhanced tolerance to drought and nitrogen deficit conditions are likely a result of the considerable heterosis effect on root system development (Hoecker et al., 2006; Chairi et al., 2016).

To harness the full potential of heterosis in hybrid crosses, parental lines must originate from genetically distinct gene pools (Larièpe et al., 2017). In rye hybrid breeding the predominant heterotic gene pools are the Petkus and Carsten (Schlegel, 2014). In maize, the relationship between MPH and specific combining ability have been demonstrated to be positively correlated with interparental genetic distance (Reif et al., 2003; Balestre et al., 2008). While genetic distance can serve as a simple predictor of hybrid performance, efficient selection of crossing parents amongst hundreds or thousands of possible combinations remain a major challenge in hybrid breeding for exploiting the maximum potential heterosis. At present complex statistical models have been developed for prediction of hybrid performance incorporating G × E, M × E and dominance effects in determination of general and specific combining ability (Acosta-Pech et al., 2017; Cuevas et al., 2017; Ferrão et al., 2020). Genomic prediction models have also been developed for hybrids in rye (Auinger et al., 2016; Bernal-Vasquez et al., 2017).

3. Leaf rust and powdery mildew in rye

3.1 Leaf rust

Encompassing nearly 8000 species the rust fungi constitute a vital component of the global terrestrial ecosystem (Helfer, 2014). As obligate biotrophic parasites, rust fungi function as drivers of community dynamics and diversification in terrestrial ecosystems through co-evolution with their host plants (Gilbert, 2002). The rust fungi belongs to the highly diverse phylum Pucciniales (formerly Uredinales) of Basidiomycota, represented in 166 genera and 14 families with majority residing in the genera *Puccina* and *Uromyces* (Kirk et al., 2008). Rusts obtained their name from their conspicuous brown-orange uredinia pustules, resembling ferrous iron oxidation (Figure 3.1). In cereals economical important rusts include stripe rust (*P. striiformis*), stem rust (*P. graminis*) and leaf rust. The classification of leaf rust fungi on grasses has a long history of controversy and revision due to reduced morphological characteristics used to infer taxonomical relationships (Liu et al., 2013). In rye, leaf rust is caused by *Puccinia recondita* f. sp. *secalis* Roberge ex Desmaz (*Prs*), while in barley (*Puccinia hordei*), and wheat (*Puccina tritici*) (Figure 3.1). Formerly, nomenclature of the leaf rust pathogen on wheat was *Puccinia recondita* f. sp. *tritici*, however, based on molecular phylogenetic relationship, interfertility, and telial host range the lineages of wheat and rye leaf rust were separated (Anikster et al., 1997; Liu et al., 2013).

Figure 3.1 Leaf rust (*Puccinia recondita* f. sp. *secalis*) on winter rye (*Secale cereale* L.) in field. **A**) Orange-brown colored pustules with urediniospores, **B**) Dark colored telium with teliospores

3.1.2 Life cycle of leaf rust on Triticeae

While the leaf rust pathosystem on rye has received little scientific attention, the system on wheat and barley are well characterized (Bolton et al., 2008; Park et al., 2015).

Leaf rust on cereals are heteroecious, requiring two taxonomically unrelated hosts to complete their life cycle, and macrocyclic entailing five distinct spore types corresponding to different life stages (Figure 3.1.2).

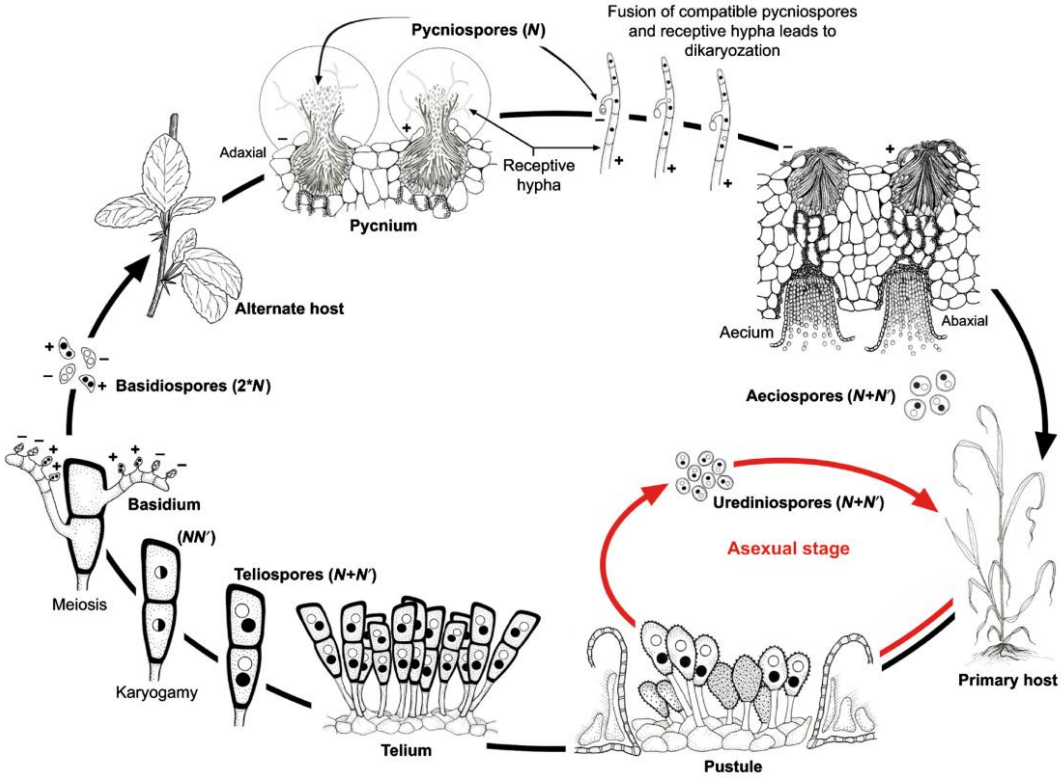


Figure 3.1.2 Exemplary life cycle of leaf rusts on cereal crops belonging to the Triticeae tribe. Adapted from the *Puccinia graminis* f. sp. *tritici* life cycle (Kolmer, 2013). Original illustration from Jacolyn A. Morrison at the USDA-ARC Cereal Disease Laboratory, St Paul, MN, USA.

Upon landing on a susceptible host leaf the urediniospores germinates within 4-8 hours at 20ºC and 100% relative humidity (Zhang and Dickinson, 2001). The germination triggers the development of a germ tube, forming several exploratory branches as it elongates (200 µm ≤) on its search for a stoma (Hu and Rijkenberg, 1998a). After recognition of a stomata an septate oval or oblong appressorium structure is formed, adhering closely to the host cuticle and stoma guard cells. Through accumulated hydrostatic pressure (turgor) in the appressorium a penetration peg is forced through the stomata into a substomatal chamber. These pre-penetration processes are concluded within 12 hours post inoculation (hpi). Within the substomatal chamber the terminal end of the penetration peg swells to form a substomatal vesicle (Hu and Rijkenberg, 1998a). From the substomatal vesicle a primary infection hypha is formed, developing a septate haustorial mother cell adjacent to a mesophyll cell in the intercellular space or an epidermal cell. As the haustorial mother cell develops, its cell wall thickens, in particular the inner wall adjacent to the host cell with a significant increase around

the penetration site (Hu and Rijkenberg, 1998b). Here a penetration tube is forced through the host cell wall leading to the invagination of the host plasmalemma to give way to the development of a feeding structure, the haustorium (explained in further detail in section 3.2.1). From the primary infection hypha several additional secondary infection hypha are formed, and new haustoria mother cells develop around 24 hpi resulting in a progressively branching fungal mycelia (Hu and Rijkenberg, 1998b). Around seven to ten days post inoculation anamorph uredinia structures are developed in the sub-epidermal mycelia. At maturity the uredinia pustule rupture the epidermis, releasing the conspicuous orange-brown urediniospores. Urediniospores are dikaryotic, spheroid (27µm x 20 µm) and exhibit spines on their surface (Anikster et al., 2005).

Resulting from the complex life cycle of leaf rust on Triticeae crop species the sexual cycle on the alternate host remains largely unstudied (Gold et al., 1979; Bolton et al., 2008). In wheat leaf rust, the alternative host is *Thalictrum speciosissimum* L., native to the Mediterranean flora (Anikster et al., 1997). In rye leaf rust, the alternate host of *Prs* are members of the *Anchusa* genera. Anikster et al. (1997) demonstrated that basidiospores of *Prs* were capable of infecting *Anchusa undulata* L. and *Anchusa arvensis* L., leading to aecial and pycnial development on the species. *A. arvensis* L. (In Danish 'krumhals') is widespread in the Northern European flora (Frederiksen et al., 2012).

The sexual stage of the life cycle is initiated by the formation of teleomorph structures, telium, towards the end of the growing season or during periods of unconducive conditions during the summer (Bolton et al., 2008). As the telium matures it acquires a conspicuous dark-brown color (Figure 3.1B). Rising perpendicular from the telium structure, two-celled, oblong (50µm x 20 µm) dikaryotic teliospores are formed (Anikster et al., 2005). Early in the teliospore development the two haploid nuclei undergo karyogamy to form a diploid nucleus. Hereafter, when the conditions are conducive, a hyphal protrusion called the promycelium is developed from one or both of the teliospore cells. Through a successive meiotic and mitotic cell division, binucleate haploid basidiospores of opposing mating types emerge from the germinated teliospore cell, now referred to as the basidium. The asymmetrical basidiospore (10µm x 8µm) is attached to the promycelium by a spine-like structure called the sterigmata (Anikster et al., 2005). Few hours after being formed the mature basidiospores are actively ejected from the sterigmata and carried by wind to a nearby alternate host. The thin-walled basidiospore is highly sensitive to desiccation, depending on surface moisture and high relative humidity to germinate

(Gold and Mendgen, 1991). If conditions are conducive the basidiospore germinates within few hours leading to the development of a basidiospore germtube. Upon recognition of a suitable site an appressorium is developed, host cuticle and cell wall penetrated, and a intraepidermal vesicle developed inside the a epidermal cell (Gold and Mendgen, 1991). Penetration can also occur at an stomata leading to the development of a substomatal vesicle instead. From the vesicle a primary hypha is formed, rapidly developing into a branched, multicellular network of mycelia without formation of haustoria. Here a globose pycnial primordia develops, erupting through the epidermal cells upon maturation with paraphyses and flexuous hyphae extending through its ostiole (Gold et al., 1979). From the ostiole ellipsoidal, thin-walled pycniospores (4µm x 2µm) are exuded in droplets of a honeydew like substance (Anikster et al., 2005). Insect facilitated transmission of pycniospore to a different pycnium of opposing mating type hereafter leads to the pairing of a receptive hyphae by plasmogamy. Followed by fusion of the two haploid nuclei by karyogamy the dikaryotic state is restored and a abaxial globose aecium structure is formed (Gold et al., 1979). In the primordial aecium spheroid, thick-walled, dikaryotic aeciospore (26 µm x 23µm) with a spiked surface are formed. The mature aeciospore is spread by wind to a nearby primary host plant, completing the life cycle of cereal leaf rusts.

3.2 Powdery mildew

Powdery mildews are some of the world's most ubiquitous plant pathogenic fungi, infecting the stems, leaves, flowers and fruits of nearly 10.000 species of angiosperms, of which 634 belongs to the *Poaceae*, the grass family (Braun et al., 2001). Their familiar and conspicuous profuse development of superficial mycelia and production of conidia have given them their name powdery mildews (Figure 3.2.1, Glawe, 2008).

Figure 3.2 Powdery mildew (*Blumeria graminis* f. sp. *secalis*) mycelia and condia on leaves of winter rye (*Secale cereale* L.)

Since the birth of powdery mildew taxonomy by Linnaeus in the 18th century, several classification systems based on anamorph and/or teleomorph morphology have been proposed to understand their phylogenetic relationships. In cultivated Triticeae specie, powdery mildew is caused by *Blumeria graminis* (DC.) Speer (formerly *Erysiphe graminis*) occupying a distinct clades within the ascomycete order Erysiphales (Menardo et al., 2017). Powdery mildews on cultivated Triticeae species are highly specialized, with *B. graminis* having evolved into minimum five distinct *formae speciales* (f. sp.), each specifically adapted to a single host specie (Menardo et al., 2017). In rye, powdery mildew is caused by *B. graminis* f. sp. *secalis* (*Bgs*), closely related to *B. graminis* f. sp. *tritici* on bread wheat and *B. graminis* f. sp. *hordei* on barley (Inuma et al., 2007).

3.2.1 Life cycle of powdery mildew on Triticeae

While the powdery mildew pathosystem on rye has received little scientific attention, the systems on barley and wheat are well characterized (Glawe, 2008).

The powdery mildews on Triticeae crop species are obligate biotrophic pathogens, depending on a host for survival and reproduction, forming two morphological distinct spore stages throughout its life cycle (Figure 3.2.1).

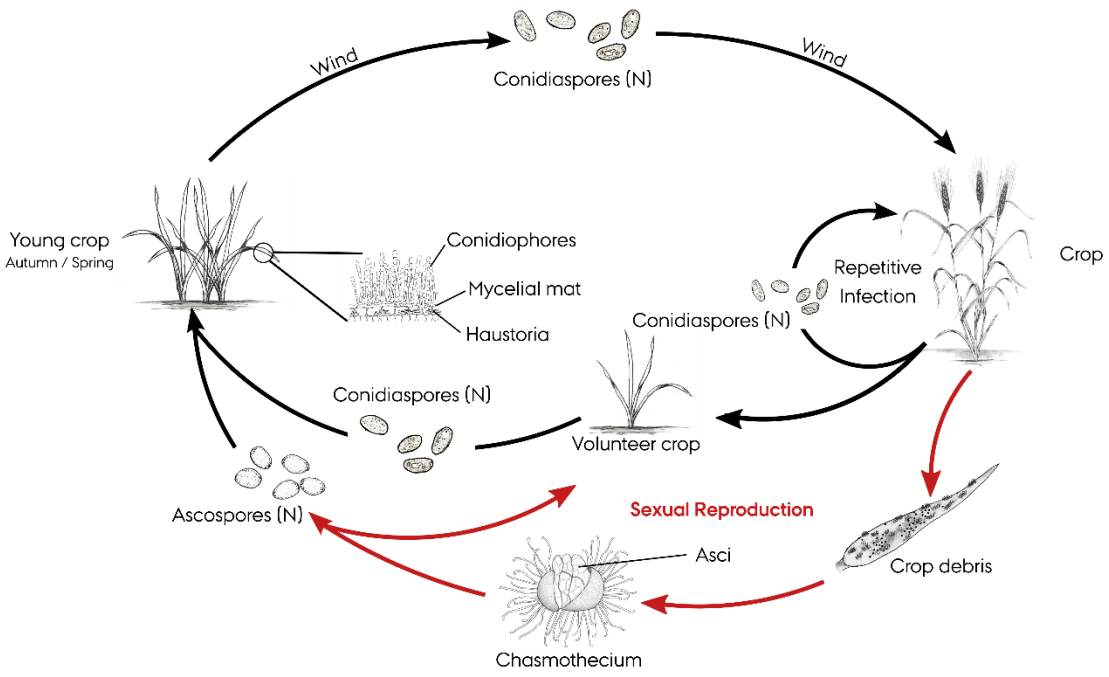


Figure 3.2.1 Life cycle of *Blumeria graminis* on cereal crops belonging to the Triticeae tribe (AU, 2021). Original illustration from Kylie Fowler

Shortly after landing on a susceptible host the infection process of powdery mildew is initiated (Nielsen et al., 2000). The germination of asco- or conidiospore leads to the initial formation of a short (5-10µm) primary germ tube that breaches through the cuticle layer, serving as anchorage to the host, an access point to host water and recognition of host surface factors (Zhang et al., 2004). Hereafter a longer (≈ 20 µm) secondary septate appressorial germ tube is produced (Jankovics et al., 2015). Upon recognition of a suitable location on the host epidermis elongation of the appressorial germ tube is haltered and an lobed apical appressorium formed (Nielsen et al., 2000; Hansjakob et al., 2012). Through accumulation of hydrostatic force (turgor) in the mature appressorium a short penetration peg is used to breach the host cuticle and cell wall. These pre-penetration processes are completed 12-15 hpi (Zhang et al., 2005). Once inside the epidermal cell lumen the penetration tip swells, initiating a invagination process of the host plasma membrane to allow for the fungal feeding structure, the digitate haustorium, to develop (Green et al., 2002). Encoding the haustorium is a extrahaustorial membrane arising from the modified host cell plasmalemma with a vast surface area of around 5000 μm^2 (Zhang et al., 2005). The extrahaustorial membrane constitutes the interaction interface used by *Bg* to manipulate host innate immune system through secretion of virulence factors and facilitate nutrient uptake (Bozkurt and Kamoun, 2020). From the haustoria secondary hyphae are formed within 24 hours after infection leading to the development of differentiated secondary hyphal appressoria and new generation of haustoria (Glawe, 2008). Progressing at an exponential pace the host epidermal surface is gradually covered in a dense mycelial mat.

Several days after host infection anamorph structures, conidiophores, are formed, rising perpendicular from the mycelial mat (Figure 3.2.1). Conidiophores are simple and unbranched vegetative structures, adding each new asexual conidium to the base of a growing chain. The conidia spores are ovoid, hyaline (10-15 µm x 30-40 µm), uninucleate and covered by spinelike protrusions (Carver et al., 1995; Zhang et al., 2005). Production of condia appears to follow a diurnal pattern with spores predominantly released by wind from mid-morning to early afternoon (Eversmeyer et al., 1973). Optimal conditions for sporulation is a temperature around 20ºC and high relative humidity, with evidence of a sudden drop in relative humidity proposed as the main environmental factor triggering active discharge of conidia (Ward and Manners, 1974; Adams Jr et al., 1986). Sporulation commences when the colony is about 4 days old with a single colony producing up to 200.000 wind-dispersed asexual spores (Zhang et al., 2005).

The sexual cycle of *B. graminis* is initiated towards the end of the growing season or triggered by stress conditions during summer leading to the development of anteridium and ascogonium structures (Glawe, 2008). Like most ascomycetes *B. graminis* is heterothallic, requiring the concurrence of a anteridium and ascogonium of opposing mating type for initiation of the sexual reproduction (Brewer et al., 2011). Upon interaction of compatible mating types plasmogamy is initiated enabling a nucleus to move from the anteridium to the ascogonium in a process known as dikaryotization (Glawe, 2008). Hereafter a peridium is developed encapsulating the multinucleate ascogonium, now referred to as chasmothecia (formerly cleistothecia). Asci then developed from dikaryotic cells, followed by karyogamy and meiosis in the ascus. The young chasmothecia (100-200 µm) are light-colored, darkening to a conspicuous brown-black color when mature. Serving as a oversummering or overwintering survival structure the chasmothecial is developed using stored nutrients absorbed from the host during the early stages of the infection (Götz and Boyle, 1998). Differentiation of the ascospore is triggered by water, maturing after 3-5 days of immersion or in a moist environment after which they are actively discharged from the asci (Jankovics et al., 2015). The ascospore are globose to ovoid, hyaline (10 µm x 20 µm), uninucleate with a clean surface. Germination of chasmothecia occurs in the late summer or early autumn leading to the infection of young newly established autumn-sown cereals as well as volunteer plants, enabling *B. graminis* to overwinter as dormant mycelium.

3.3 Epidemiology of leaf rust and powdery mildew

The capacity of plant pathogens to cause disease and give rise to severe epidemics as a product of the interrelationship between three factors was elegantly framed in the 'Disease triangle' by McNew (1960), defining these as:

'*The inherent susceptibility of the host, the inoculum potential of the parasite, and the impact of the environment on parasitism and pathogenesis'* (McNew, 1960)

Within the pathogen factor several life history traits influences a pathogenic species capacity to cause disease on a large spatial and temporal scale (Le May et al., 2020). The foliar diseases on cultivated Triticeae species, leaf rust and powdery mildew, share several decisive life history traits related to their ability to cause severe epidemics. This includes an asexual polycyclic phase in their life cycle, their ability to develop highly aggressive pathotypes, capacity to disperse spores over long distances and lastly, survival strategies to insure their persistence during off-seasons. While these life history traits have received little scientific interest in

relation to the epidemiology of the leaf rust and powdery mildew pathosystem on rye, they are well characterized on both wheat and barley.

3.3.1 Epidemiology of leaf rust on Triticeae

Long-distance dispersal (LDD) constitute an important survival strategy for obligate biotrophs, enabling them to migrate between habitats were the host plant are seasonally absent and invade new territories (Brown and Hovmøller, 2002). Several cases of LDD in rusts on cultivated Triticeae species have been documented. In leaf rust on wheat, urediniospores have been observed to migrate 500 kilometers by wind across the North Sea from the British Isles to Denmark (Hermansen et al., 1978). In the wheat stem rust pathogen a novel aggressive race 'Ug99' spread from Uganda in 1998 and across the East- and Southern African countries, and into the Middle east, following wind trajectory models (Singh et al., 2011b). In the wheat stripe rust pathogen, introduced non-European origin races 'Warrior', 'Kranich' and 'Triticale aggressive' are believed to have migrated by wind from the near-Himalayan region of Asia (Hovmøller et al., 2016). In leaf rust on wheat, urediniospores have been observed to migrate 500 kilometers by wind across the North Sea from the British Isles to Denmark (Hermansen et al., 1978). This LDD constitutes a corridor of genetic exchanged between elsewise spatial secluded populations, having a profound effect on spatiotemporal disease dynamics, genetic diversity and structure of the rust populations.

Genetic variation is an essential component in the evolutionary capacity of plant pathogens enabling them to adapt changes in their environment and host (Marais and Charlesworth, 2003; Figueroa et al., 2020). Through recombination during the sexual reproductive cycle novel gene combinations are created with the potential evolution of aggressive pathotypes (Heitman, 2015). In leaf rust on wheat several studies have, however, found little evidence of sexual recombination suggesting a predominantly clonal propagation of the pathogen (Goyeau et al., 2012; Kolmer, 2015). Similar observation has been made in the yellow rust pathogen on wheat (Hovmøller et al., 2002b).

While the alternate host, *Berberis* sp., has been actively eradicated in intense campaigns across Europe to hinder the sexual reproduction of wheat yellow rust, the alternate host of wheat leaf rust, *T. speciosissimum*, is present in the Mediterranean region. However, it is likely that the sexual reproduction in wheat leaf rust plays an insignificant role in the epidemiology and development of novel genetic variation due to a low occurrence of the alternate host (Bolton et al., 2008). In clonal populations novel genetic variation is instead generated by other means

such a mutations, somatic hybridization and internuclear exchange (Figueroa et al., 2020). The evolutionary plasticity of the wheat leaf rust pathogen in the absence of sexual reproduction was demonstrated by Kolmer et al. (2007), identifying 72 pathotypes amongst 797 single pustule isolates in the United States in 2005 (Kolmer et al., 2007).

The alternate host of rye leaf rust, *Anchusa arvensis,* is, however, widespread in the Northern European flora (Anikster et al., 1997; Frederiksen et al., 2012). While suggesting the likelihood of frequent sexual recombination, its role in the epidemiology of rye leaf rust remains to be investigated. In a study on the virulence dynamics of rye leaf rust by Miedaner et al. (2011), 317 pathotypes were identified amongst of 827 single pustule isolates collected during 2000-2002 in Germany. It is to expected that the heterogeneous nature of rye and low level of leaf rust resistance inherent to the predominant hybrid rye breeding systems are contributing factors positively influencing the diversity in virulence phenotypes (Miedaner et al., 2002; Goyeau et al., 2006). The high virulence complexity and frequency of distinct virulence phenotypes observed in the German leaf rust population, however, could suggest a mixed reproduction mode, alternating between asexual and sexual. This is supported by the large formation of telia structures on field crop (Figure 3.1B).

In wheat yellow rust, introduced non-European origin races have been determined to originate from a sexual reproducing and genetically diverse population in the near-Himalayan region of Asia (Hovmøller et al., 2016).

Leaf rusts are polycyclic disease capable of undergoing a large number of clonal generation per growing season. The repetitive infections by asexual urediniospores are an essential component in the epidemiology of leaf rusts, driving an exponential progression of disease within field that is further accelerated by wind dispersal. In rye leaf rust, clonal reproduction have been demonstrated to drive a rapid expansion of virulent pathotypes reaching up to 60% of the population within weeks (Miedaner et al., 2011). Facilitated by LDD these virulent pathotypes can be disseminated on a large spatial scale, contributing to risk of severe epidemics across continents (Hovmøller et al., 2002a).

3.3.2 Epidemiology of powdery mildew on Triticeae

While the powdery mildew fungi on cultivated Triticeae species shares several life history traits with the leaf rust pathogen as elaborated in detail above, it differentiates on several epidemiology related traits.

Several studies in barley and powdery mildew have demonstrated the capacity of both conidia- and ascospore to migrate hundreds or thousands of kilometers by wind (Limpert et al., 1999; Zhu et al., 2016). Similar to leaf rust, conidiospores have been observed to migrate across the North Sea from the British Isles to Denmark (Hermansen et al., 1978). In 1988 a virulent pathotype of barley powdery mildew was observed in England believed to have migrated by LDD from Czechoslovakia across the European continent and North Sea (Wolfe et al., 1992).Virulence complexity of wheat and barley powdery mildew have been observed in several studies to correlated with prevailing wind direction, emphasizing the impact of LDD on powdery mildew epidemiology (Cowger et al., 2016; Zhu et al., 2016).

Unlike leaf rusts, sexual recombination is frequent in the powdery mildews on cultivated Triticeae species, independent of an co-occurring alternate host (Cowger et al., 2016; Dreiseitl, 2018). Through sexual recombination foreign pathotypes, introduced by LDD, are introgressed into existing local populations leading to development of novel virulence gene combinations and thereof new pathotypes (Heitman, 2015). The effect of sexual recombination on pathotype diversity was demonstrated in a study on the Central European barley powdery mildew population, identifying 279 distinct pathotypes out of 309 isolates tested from 2015-2017 (Dreiseitl, 2018). Similar observation was made in the wheat powdery mildew where Cowger et al. (2016) identified 210 distinct haplotypes amongst 238 isolates sampled in United States from 2007 to 2010. The considerable extent of gene flow enabled by this wind corridor is likely a key contributing factor to the large spatial variability observed in the genetic structure of wheat and barley powdery mildew (Limpert et al., 1999).

Concurrent to the increase in pathotype diversity, LDD and frequency sexual recombination are likely drivers of the extreme levels of genetic diversity and pathotype complexity observed in the US, European and Chinese wheat and barley powdery mildew populations (Zhu et al., 2016; Cowger et al., 2018; Dreiseitl, 2018).

3.4 Virulence of leaf rust and powdery mildew on Triticeae

In natural and agricultural ecosystems a ceaseless battle is fought between plants and pathogens in their struggle for survival and reproduction (Thompson and Burdon, 1992). This struggle drives a concurrent arms-race with selective forces on host defense and pathogen virulence framed elegantly in the gene-for-gene coevolution theory by Flor (1971).

'*For each gene that conditions reaction in the host there is a corresponding gene that conditions pathogenicity in the parasite'* (Flor, 1971)

In order to infect the host, pathogens need to surpass two layers of defense, the preformed and the induced (Anderson et al., 2010). The preformed refers to a constitutive defense that discourages, or provides a physical or chemical barrier preventing pathogen infection. The induced defense constitute two types of immune receptors that upon recognition of the pathogen initiates a resistance response. First layer of inducible defense are the pattern recognition receptors (PPRs) located on the plant cell surface. The PRRs recognizes pathogen associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs). PAMPs are conserved components of pathogen cell structure such as chitin in the fungal cell walls or cell walls or flagellin in bacterial flagellum, whereas DAMPs are released during alteration of host cell integrity (Pel and Pieterse, 2013; Raaymakers and Van den Ackerveken, 2016; Peng et al., 2018). Second layer of inducible defense are intracellular resistance (R) proteins that recognizes effector molecules (Laflamme et al., 2020).

To breach the physical barrier pathogens secrete extracellular hydrolase enzymes degrading the waxy cuticle and epidermal cell wall facilitating penetration by specialized structures such as the appressoria of cereal rusts and mildews (Feng et al., 2011). Hereafter, metabolic mechanisms allow the pathogen to circumvent host deployed antimicrobial phytochemicals (Pedras and Abdoli, 2017). In order to suppress or circumnavigate PAMP or DAMP triggered immunity pathogens secrete extracellular effector molecules (Varden et al., 2017). Suppression of host defense response is central to the survival of obligate biotrophic pathogens, such as powdery mildew and leaf rust, allowing them to inhabit living cells and feed on host derived metabolites. The secreted effector molecules interact with host R proteins in accordance to the gene-for-gene coevolution theory (Bourras et al., 2015). For every *R* gene in the host, a ligand in the form of a avirulence gene exist in the pathogen population giving cause to a resistance response, referred to as an incompatible interaction (Zhong et al., 2017; Peng et al., 2018; Yoo et al., 2020). Virulent pathotypes capable of successfully infecting the host instead carry a virulence gene allowing them to suppress PAMP triggered immunity and evade *R* gene recognition, this is referred to as an compatible interaction (Lanver et al., 2017). The host-pathogen interactions give cause to a wide spectrum of compatibility or incompatibility reactions on the phenotypic level related to the molecular basis of the response (Kosman et al., 2019). The interaction is furthermore influenced by plant fitness, related to the agricultural

practices and environmental factors. Determining the virulence phenotype is often done using quantitative infection type (IT) scales, with IT categories defined by lesion type, lesion size, chlorosis, necrosis, spore production etc. (Kosman et al., 2019). Virulence surveys of pathogens is done on basis of collected isolated IT spectrum, or profile, towards a differential set of host cultivars with complementing *R* gene repertoire, allowing a differentiation of isolates into distinct pathotypes (Hovmøller et al., 2017).

3.4.1 Infection types of rye leaf rust

While differential sets for virulence phenotyping of leaf rust in rye exists, access to the inbred breeding lines owned by private or governmental breeding institutes is often restricted (Miedaner et al., 2011).

In rye a ten-step 0-9 IT scale was developed on basis of incompatible and compatible interactions observed in a screening of 101 inbred rye lines towards six distinct *Puccinia recondita* f. sp. *secalis* single pustule isolates (Manuscript IV, Figure 3.4.1).

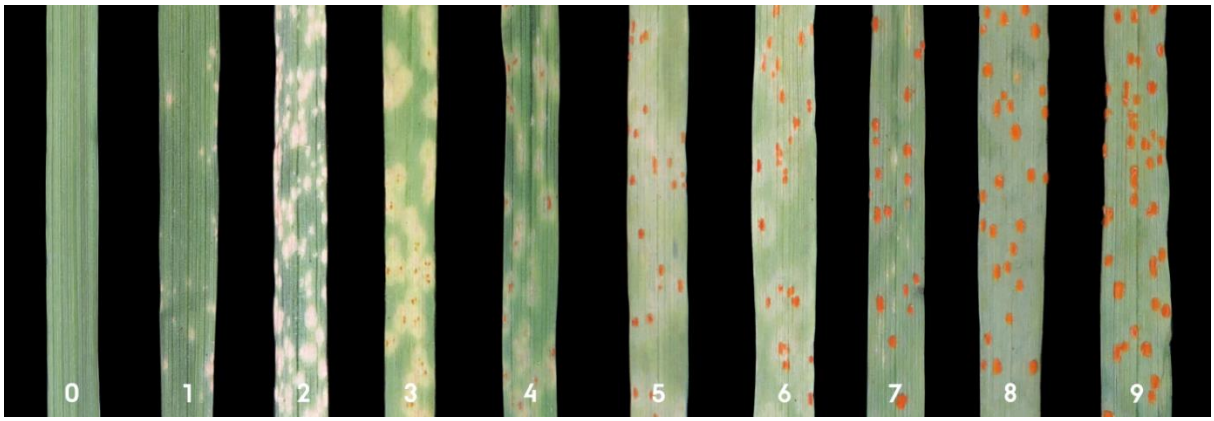


Figure 3.4.1: Infection type response (0-9) for leaf rust in rye (*Secale cereale* L.) caused by the fungal pathogen *Puccinia recondita* f. sp. *secalis.* Inspired by Hovmøller., *et al.* (2017) and McNeal., *et al.* (1971). IT 0-2 are considered to represent 'Resistant', IT 3-4 'Partial resistant', IT 5-6 'Partial susceptible', and IT 7-9 'Susceptible' host plants. In terms of virulence/avirulence, IT 0-6 are considered 'avirulent' and 7-9 'virulent'.

3.4.2 Infection types of powdery mildew on Triticeae

To our knowledge no differential sets has been developed for virulence phenotyping of powdery mildew in rye. In barley and wheat virulence phenotyping of powdery mildew has predominantly been done according to the ten step 0-4 IT scale by Torp et al. (1978) (Sørensen et al., 2019).

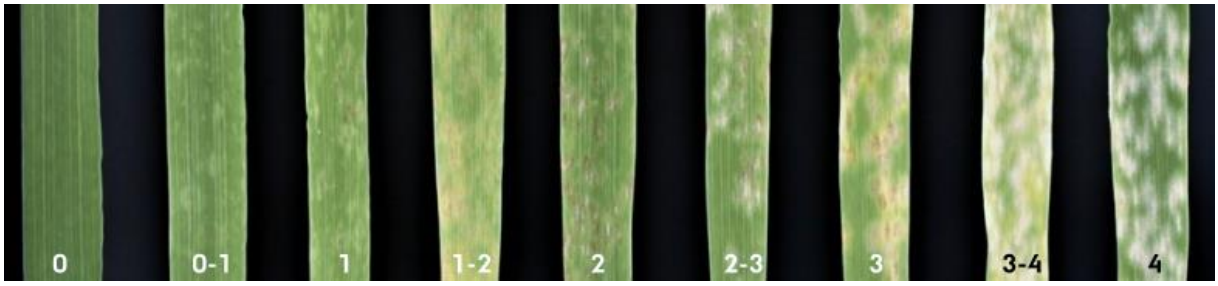


Figure 3.4.2: Infection type response (0-4) for powdery mildew in barley (*Hordeum vulgare* L.) caused by the fungal pathogen *Blumeria graminis* f. sp. *hordei*, After (Torp et al., 1978), with original pictures from Chris Khadgi Sørensen at Aarhus University Department of Agroecology, Flakkebjerg, Denmark. IT 0 to 1 are considered 'resistant', IT 1-2 to 2 'Partial resistant', IT 2-3 to 3 'partial susceptible' and IT 3 to 4 'Susceptible'. IT less than 3 are considered 'avirulent' and IT 3 to 4 'virulent'.

4. Host resistance in rye to leaf rust and powdery mildew disease

In natural as well as agricultural ecosystems plants coexist with a multitude of microorganisms (Compant et al., 2019). The plants interaction with the surrounding microbial community residing in the rhizosphere (root interface), phyllosphere (aerial interface) and endosphere (internal interface) habitat are crucial determinants of plant functioning and health (Finkel et al., 2017). Members of these communities comprise beneficial mutualists, neutral commensals and pathogenic microorganisms (Trivedi et al., 2020). In order to distinguish 'foe' from 'friend' plant exhibit an intricate immune system with several layers of preformed and induced defense (Anderson et al., 2010).

4.1 Plant innate immune system

In plants the first layer of defense is the waxy cuticle and epidermal cell wall, constituting a protective physical barrier (Ziv et al., 2018). This preinvasion resistance prevents most non-adapted pathogens without the appropriate infection structures or enzymatic repertoire from infecting the plant (Lee et al., 2017). Beside the physical properties of the barrier, accumulation of secondary metabolites in the plant epidermal cells and cuticle constitutes a metabolic or chemical layer of defense. The secondary metabolites constitute antimicrobial phytochemicals either toxic to the pathogen or exhibiting a inhibitory effect on pathogen infection (Calmes et al., 2015). The diverse array of plant antimicrobial compounds are divided in two groups, phytoanticipins that are produced and store constitutively in plant tissue and phytoalexins that are synthesized *de novo* in response to infection (Piasecka et al., 2015). To avoid auto-toxic effects some phytoanticipins are stored as inactivated forms, activated upon recognition of pathogen invasion (Chen et al., 2020). The early sensing of pathogen invasion is associated with plant recognition of PAMPs or DAMPs by a cell surface-associated PRR proteins mainly located at the plasma membrane (Saintenac et al., 2018). Plant PRRs are often represented by receptor-like protein kinases (RLKs) or receptor-like proteins (RLPs) (Jamieson et al., 2018). RLKs consist of a N-terminus extracellular domain, a central single-pass transmembrane domain, and a C-terminus intracellular kinase domain, with RLPs exhibiting the same overall structure, however, absent of a kinase domain (Wu et al., 2018). Some RFLs additionally have a short extracellular or intracellular juxtamembrane domain (Albert et al., 2019; Gong et al., 2019). The extracellular ligand recognition domain consist of a lysine motif, lectin domain, epidermal growth factor-like domain, or the more common a leucine-rich repeat (LRR) domain (Jamieson et al., 2018). Absent of a intracellular kinase domain responsible for triggering the

downstream signal transduction, RLPs are believed to be associated with RLKs (Wu et al., 2018). Recognition of a PAMP or DAMP lead to a PAMP-triggered immune (PTI) response. PTI is characterized by a cascade of rapidly deployed cellular responses including the activation of mitogen-activate protein kinase cascade (Bi et al., 2018), calcium-dependent protein kinases (Marcec et al., 2019), generation of reactive oxygen species and nitrogen species (Zhang et al., 2018), ion flux (Thor et al., 2020), callose deposition (Wang et al., 2021), phytohormone production (Lal et al., 2018), transcriptional reprogramming (Xu et al., 2017) production of phytoalexins, and activation of phytoanticipins (Piasecka et al., 2015).

In addition to the direct or indirect perception of pathogen-derived molecules on the cell surface by PRR proteins, seven other mechanisms for R protein functioning have recently been presented by Kourelis and Van Der Hoorn (2018) (Figure 4.1A).

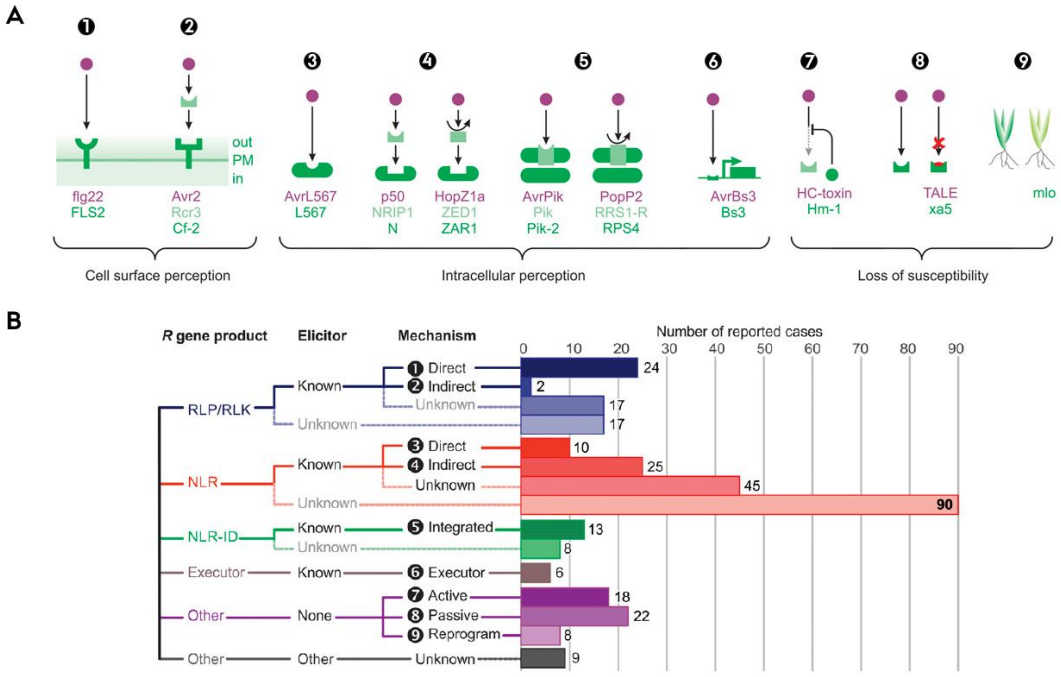


Figure 4.1 Molecular mechanisms of resistance (*R*) gene functioning in plants; including **A**) Direct (1) or indirect (2) recognition of pathogen-derived molecules by cell surface localized RLK or RLP, direct (3) or indirect (4) intracellular recognition of pathogen-derived molecules by nucleotide-binding leucine rich repeat receptors (NLR), or through integrated domains (5), recognition of transcription activator-like effectors through activation of executor genes (6), and active (7) or passive (8) or host programming-mediated (9) loss of susceptibility. **B**) Grouping and frequency of identified *R* genes based on their proposed mechanism. After (Kourelis and Van Der Hoorn, 2018)

The sensing of a cytoplastic effector proteins (3-5) by intracellular nucleotide-binding leucine rich repeat (NLR) proteins induces a effector triggered immunity response (ETI) response (Figure 4.1A). While several of the downstream molecular events triggered during ETI are similar to PTI, it is distinguished by the triggering of a hypersensitive response (HR) (Peng et

al., 2018; Lu and Tsuda, 2021). The HR is a type of programmed cell death induced by production bursts of reactive oxygen species and nitric oxide effective against biotrophic pathogens, such as cereal leaf rust and powdery mildews, depending on living host cells (Balint-Kurti, 2019).

The canonical NLR gene is composed of three domains, a C-terminal leucine-rich repeat (LRR) domain responsible for effector recognition and NLR autoinhibition (Qi et al., 2012; Monino-Lopez et al., 2021) and a conserved central nucleotide-binding and ARC (present in Apaf-1, R proteins, and CED-4) (NB-ARC) domain involved in regulation of NLR activation (Williams et al., 2011). On basis of the N-terminal domain involved in mediating the downstream signaling, NLR genes are divided into three broad classes, (i) CNLs carrying a coiled-coil (CC) domain, (ii) TNLs carrying a Toll-interleukin 1 receptor (TIR), and (iii) RNLs carrying a RPW8-like CC domain (Shao et al., 2016). In monocots the TNL-type NLR genes are absent (Tamborski and Krasileva, 2020). The sensing of pathogens effectors by intracellular NLR genes occur by one of three mechanisms, (i) direct interaction by effector binding to the LRR domain, (ii) indirect interaction by either effector binding to or mediated modification of a host component perceived by the NLR, or (iii) indirect interaction by effector binding to or mediated modification of a domain that it integrated in a host NLR (NLR-ID) (Figure 4.1A, Ortiz et al., 2017; Zhang et al., 2017; Kourelis and Van Der Hoorn, 2018). Host components and integrated domains that have lost their original function, serving only as ‘effector baits’ are often referred to as decoys, while those that have retained their functions in plant immunity are referred to as guards (Baggs et al., 2017). The fused integrated domain can be located in either end of the NLR gene.

The NLRs constitute a large gene family in flowering plants, with 1,169 NLRs identified in the reference genome of a German inbred grain-type rye line ‘Lo7’ and 1,447 in a Chinese inbred forage-type rye line ‘Weining’, corresponding to 3.2-3.4 % of the gene complement in rye (Li et al., 2021; Rabanus-Wallace et al., 2021). In plants, NLRs are predominantly located in the subteleomeric region organized either in pairs, often in a head-to-head arrangement, or large clusters spanning several megabases (Rabanus-Wallace et al., 2021). NLR pairs situated in head-to-head arrangement have been observed to function as coupled ‘helper’ and ‘sensor’ in three different combinations, (i) paired NLR and NLR-ID sensor and helper, (ii) RNL helper and CNL or TNL sensor, and (iii) NLR Required for Cell Death (NRC) helper and CNL sensor (Feehan et al., 2020).

In hexaploidy bread wheat 60% of the NLR repertoire is organized in gene clusters (Andersen et al., 2020). These NLR gene clusters fall into two categories, (i) functionally linked genes with little sequence similarity, often diverging by means of gene conversion events and intra-cluster rearrangements and (ii) homologous genes (paralogs) arising from duplication events often followed by unequal crossing over (van Wersch and Li, 2019). With the occurrence of gene clustering in eukaryotes being uncommon the conservation of clustering amongst NLR genes suggest an evolutionary benefit of the gene arrangement (van Wersch and Li, 2019).

Structural variations plays a large role in the evolution and diversity of NLR genes from small scale insertion-deletion (indel) polymorphism, to copy number variation (CNV) and presence/absence variation (Baggs et al., 2017). PAV is an extreme case of structural variation where fragments in the size of genes are missing and has been associated with an over-representation of NLR genes (Gabur et al., 2020). NLRs that have lost their fitness effect due to absence of the corresponding virulence factor in the pathogen population would favor PAV, likewise explaining observations of NLR PAV between accessions originating from distinct environments (Stam et al., 2019).

Individual NLRs have been associated with high levels of allelic diversity with the enrichment of non-synonymous mutations enriched in the LRR domain, signature a positive selection pressure (He et al., 2020). The diversification of NLRs through point mutations, intragenic and intergenic recombination and gene conversion are likely drivers of chimeric LRRs with novel recognition specificities (Lindner et al., 2020; Tamborski and Krasileva, 2020). Domain fusion likewise constitute a means of diversification, adding to the heterogeneity of NLRs (Baggs et al., 2017). The rapid evolution of certain NLRs is likely linked to their accumulation around recombination hotspots in the subteleomeric region (Nieri et al., 2017; Chen et al., 2018).

While pivotal to plant health, resistance to pathogens has a metabolic cost and the diversification of NLRs a pleiotropic effect on plant fitness (Brown and Rant, 2013). Expression of NLRs is, therefore, tightly regulated to prevent autoimmunity or retarded plant growth while retaining resistance (Newman et al., 2019). In the plant NLR proteins are believed to exist in an equilibrium between a resting and active state with activity regulated by the central NB-ARC domain through a ‘trigger-and-switch’ mechanism (Borrelli et al., 2018). Upon binding of an effector molecule the NLR protein transitions from the resting state by release of a bound ADP to an active by binding of a ATP, inducing a conformational change in the LRR domain, stabilizing the activated NLR (Wang et al., 2019a).

In addition to the PTI and ETI triggered by perception of pathogen-derived molecules by receptor proteins, *R* genes can also confer resistance by loss of susceptibility (Figure 4.1, Kourelis and Van Der Hoorn, 2018). Loss of susceptibility is either caused by host proteins that are directly involved in disarming the pathogen, passive loss of host susceptibility factor interacting with the pathogen effector, or host reprogramming by mutations in components of the cellular pathway (Wang et al., 2019c; de Toledo Thomazella et al., 2021). The resistance, often broad spectrum, conferred by loss of susceptibility is characterized the absence of visual cell death, referred to as type-I non-host resistance, while type-II is characterized by HR and visual cell death (Lee et al., 2016). A well-characterized example of a loss-of-function mutant is the barley *Mildew resistance locus o* (*Mlo*), conferring a broad-spectrum resistance towards the powdery mildew pathogen (Acevedo-Garcia et al., 2014).

Most of the genes conferring adult plant resistance (APR) expressed in the later stages of plant life are related to host reprogramming induced loss-of-susceptibility (Kourelis and Van Der Hoorn, 2018). APRs often confer a broad spectrum resistance effective towards a large range of pathogen races (Chen, 2013a). A well-characterized example of a loss-of-susceptibility mutant conferring APR is the wheat leaf rust resistance gene *Lr34*, encoding a ABC transporter (Krattinger et al., 2011).

While the resistance mechanism conferred by major *R* genes have been described extensively, the molecular mechanisms controlling variation in quantitative, or partial, resistance is much less well-known (Kourelis and Van Der Hoorn, 2018). The study of quantitative resistance is hindered by the biological complexity underlying resistance, often exhibiting a polygenic architecture with many genes of small to moderate effects (Corwin and Kliebenstein, 2017). Quantitative resistance is expressed as a susceptible infection type of host reaction with reduced infection frequency and severity (Figure 3.4.1, Figure 3.4.2, Sucher et al., 2017).

4.2 Resistant to leaf rust disease in rye

While leaf rust resistance in rye has attracted little scientific attention in recent years, several earlier studies have investigated qualitative and quantitative resistance towards the disease in rye (Reichel, 1981; Musa et al., 1984; Miedaner et al., 2002; Solodukhina, 2002). In 2004 a systematic rye nomenclature for leaf rust *R* genes '*Pr'* was introduced by Roux et al. (2004) as reference to the causative agent of leaf rust in rye *Puccinia recondita* f. sp. *secalis*. Currently five major leaf rust *R* genes conferring race-specific resistance have been identified in rye on chromosome arms 1RS (*Pr3*), 1RL (*Pr4, Pr5*), 6RL (*Pr1*), and 7RL (*Pr2*) (Wehling et al., 2003;

Roux et al., 2004). Previously these *R* genes were denoted in accordance to wheat nomenclature as *Lr-a* (*Pr1*), *Lr-c* (*Pr2*), *Lr-c* (*Pr3*), *Lr-g* (*Pr4*), and *Lr-h* (*Pr5*) (Roux et al., 2000). Additional three major leaf rust *R* genes have been identified in wheat-rye substitution and translocation lines with rye as resistance donor. In accordance to wheat gene nomenclature these leaf rust *R* genes in rye are designated as *Lr26* on chromosome arm 1RS (1BL-1RS), and *Lr25* and *Lr45* on chromosome arm 2RL (McIntosh et al., 1995; Friebe et al., 1996). The former gene *Lr26* is located on the renown 1BL.1RS translocation from Petkus rye that has been widely deployed in commercial wheat cultivars since mid-1950s (Schlegel and Korzun, 1997). The chromosome arm likewise carries yellow rust (*Yr9*), stem rust (*Sr31*), and powdery mildew (*Pr8*) *R* genes (Mago et al., 2005).

In wheat, the resistance conferred by monogenic inherited major *Lr R* genes have been associated with a short effective life-span (Wang et al., 2010; Kolmer et al., 2013). In contrast, a combination of minor and major *R* genes conferring non-race specific quantitative resistance, also referred to as slow rusting resistance in cereal rusts, has contributed considerably to the control of leaf rust in wheat by providing a layer of more durable resistance (Huerta-Espino et al., 2011; Singh et al., 2011a). Slow rusting *R* genes and QTLs often confer a APR-type resistance expressed in later stages of plants as a susceptible infection type of host reaction with reduced infection frequency and severity (Sucher et al., 2017). While individual slow rusting *R* genes often exert an insufficient level of resistance at high disease pressure, a high level of resistance can be obtained by the pyramiding of multiple *R* genes (Singh et al., 2011a; Pilet-Nayel et al., 2017). In wheat more than 12 major slow rusting *R* genes and 80 QTLs conferring APR have been identified and (Li et al., 2014).

In rye, while slow-rusting resistance towards leaf rust has been reported, no *R* genes or QTLs have to our knowledge been identified and deployed in commercial hybrid cultivars (Reichel, 1981; Miedaner and Sperling, 1995; Miedaner et al., 2002). In a smaller collection of inbred rye lines belonging to the Petkus and Carsten gene pools, Miedaner et al. (2002) found that four lines displayed a APR resistance to leaf rust.

The inherent level of leaf rust resistance is low in the predominant heterotic gene pools Petkus and Carsten used for hybrid breeding in rye (Figure 4.2, Miedaner et al., 2002). Already in 2004, (Roux et al., 2004) reported that virulence towards *Pr3*, *Pr4*, and *Pr5* was common in rye growing regions in Germany. While less common pathotypes virulent to *Pr1* has also been observed in Germany (Wehling et al., 2003; Roux and Wehling, 2010).

During the past 25 years several outbreaks of leaf rust has been recorded in the Danish official trials across the top-yielding hybrid rye cultivars and population varieties tested (Figure 4.2A, Sortsinfo, 2021). Since 1997 the average disease severity of leaf rust has been 7.6 ± 6.5% leaf

area covered by the disease across the tested panel in Denmark. Leaf rust has, however, had a positive incline in the Danish official trials during the period with disease severity increasing by 44% from a mean of 6.2% in 1997 to 8.9% leaf area covered in 2020, converted to a yearly increase in severity of 1.8%. In 2019 the evaluated top-yielding hybrid cultivars demonstrated a maximum severity score between 10 to 50% with severity highly dependent on trial location, spanning from an average of 2.9 ± 4.9% at a trial site in Central Jutland to 21.8 ± 11.8% in West Zealand (Figure 4.2B). In years with prolonged periods of conducive conditions, leaf rust has been observed to cause a 11-27% reduction in grain yield in addition to considerable quality losses (Miedaner and Sperling, 1995). In Denmark, yield losses of 7 to 12% have been reported in leaf rust inoculated field trials (Jørgensen et al., 2013). Higher yield losses of up to 60% have, however, been recorded in winter wheat (Morgounov et al., 2015). While the effect of leaf rust on modern hybrid rye cultivars has not been investigated recently, it is reasonable to infer on basis of the low level of inherent resistance, frequency of outbreaks and average disease severity across top-yielding hybrids that leaf rust is causing considerable grain yield losses in rye (Figure 4.2).

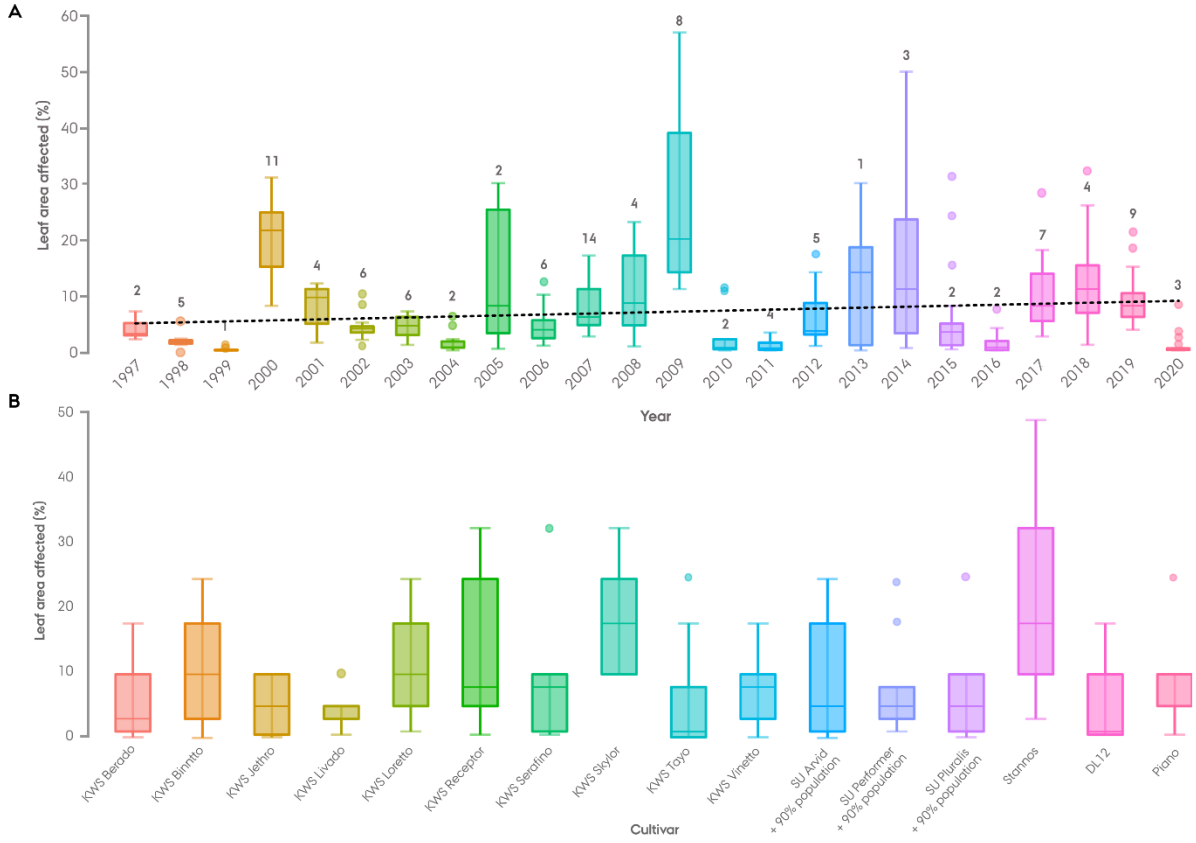


Figure 4.2 Leaf rust in rye (*Secale cereale* L.) hybrid cultivars and population varieties tested in the Danish official trials (Sortsinfo, 2021). **A**) Distribution of mean leaf rust severity across test panel from 1997 to 2020 with number of trial locations per year in bold above the boxplot. Trend line is represented by a block dotted line, **B**) Leaf rust disease severity across 16 hybrid rye cultivars at nine locations in Denmark in 2019.

4.3 Resistance to powdery mildew disease in rye

Rye remains to this day an important genetic resource for improvement of powdery mildew resistance in wheat (Hao et al., 2018; Ma et al., 2020; Ren et al., 2020). Through wheat-rye translocation- or substitution lines rye chromatin harboring powdery mildew *R* genes and resistance QTLs have been introgressed into wheat (Crespo-Herrera et al., 2017). Chromosomal translocation lines resistant towards powdery mildew have likewise been developed using wild accessions of the perennial *S. sylvestre* as resistance doner (He et al., 2021). Currently 23 *Pm R* genes have been identified in rye, of which majority in wheat-rye translocation- and substitution lines (Manuscript V, Table 1). Most of the identified *Pm R* genes confer a complete all-stage resistance characteristic for an NLR-type *R* gene. Currently, *Pm8* and *Pm1* homologs in wheat (*Pm3, Pm17*) have been identified as CNL-type NLR genes, conferring race-specific resistance to powdery mildew (Hurni et al., 2013; Singh et al., 2018). In wheat the resistance conferred by such monogenic inherited major *Pm R* genes have been associated with a short effective life-span (Wu et al., 2019; Xue et al., 2021).

While more than a hundred QTLs conferring resistance to *Bgt* in wheat have been identified, little is known on the genetics underlying quantitative resistance to powdery mildew in rye (Li et al., 2014; Kang et al., 2020). However, investigations of powdery mildew resistance in the Petkus and Carsten based hybrid rye breeding germplasm at Hohenheim University led to findings of prevalent quantitative resistance (Kast and Geiger, 1982; Miedaner et al., 1993).

During the last 25 years powdery mildew has constituted a less important biotic stress factor in the Danish rye production (Figure 4.3A). In this period powdery mildew disease severity has had a negative incline of 46% in the Danish official trials from a mean of 2.9% to 1.6% leaf area covered by the disease in 2021. Outbreaks on national scale in Denmark have been rare with only a single year with elevated disease severity recorded in 2010 (Sortsinfo, 2021). Certain regions in Denmark are, however, prone to exhibit more severe recurrence of powdery mildew such as the Southern region of Jutland as can be seen in the Danish official trial records from 2017 (Figure 4.3B). At the trial site the tested top-yielding hybrid rye cultivars displayed a 7% mean leaf area covered by powdery mildew, with 5 cultivars scoring more than 10% (Sortsinfo, 2021). While no recent studies have investigated the effect of powdery mildew in top-yielding hybrid rye cultivars Matzen et al. (2019) reported a 16% yield reduction in triticale at disease severity of ≤ 10% leaf area covered by powdery mildew under field conditions in

2017. The sowing of susceptible hybrid rye cultivars in predisposed regions of Denmark is, therefore, likely associated with risks of considerable grain yield and quality losses.

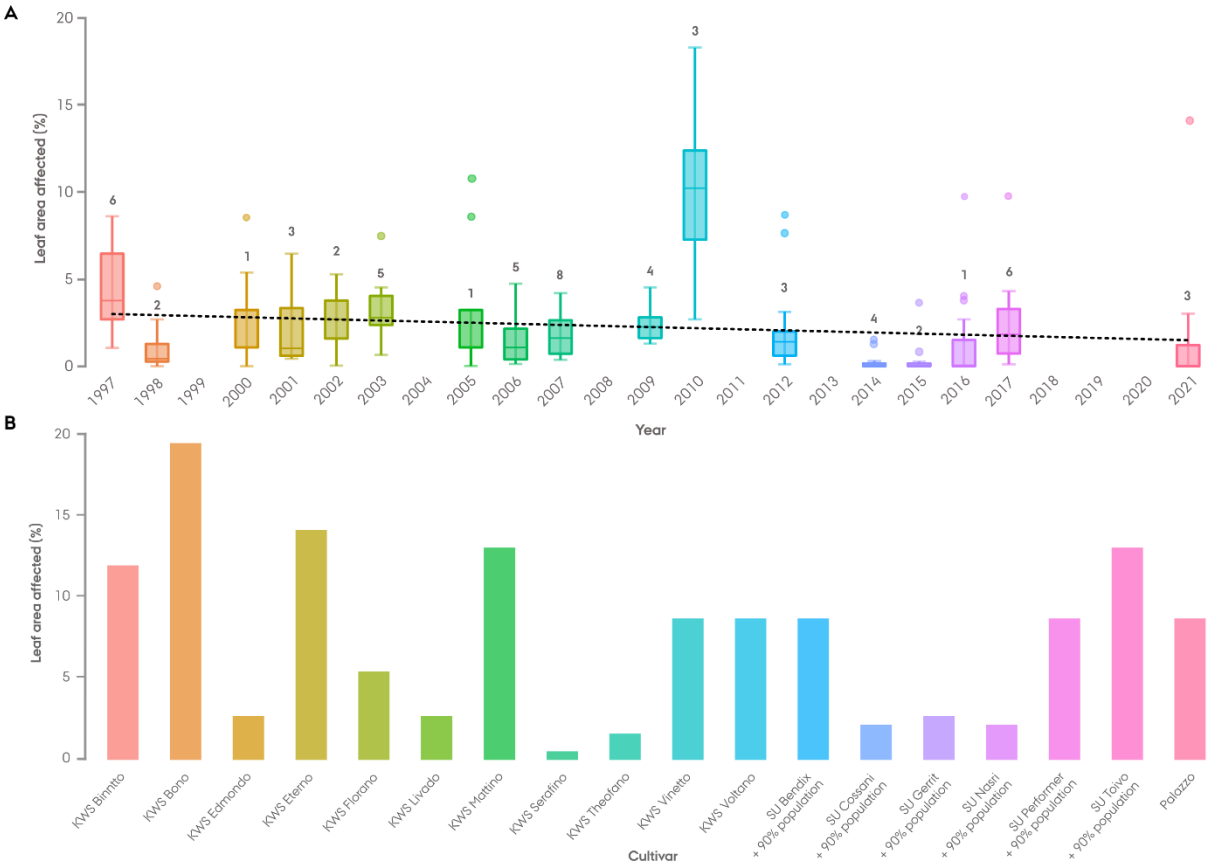


Figure 4.3 Powdery mildew in rye (*Secale cereale* L.) hybrid cultivars and population varieties tested in the Danish official trials (Sortsinfo, 2021). **A**) Distribution of mean powdery mildew severity across test panel from 1997 to 2021 with number of trial locations per year in bold above the boxplot, Trend line is represented by a block dotted line. **B**) Powdery mildew severity across 18 hybrid rye cultivars at the Tinglev trial site in Southern Denmark in 2017.

5. Molecular plant breeding for resistance

The integration of advances in genomic research, crop biotechnology, and molecular marker application with conventional plant breeding practices has transitioned breeding for disease resistance in plants (Moose and Mumm, 2008). In recent years several significant milestones have been reached in rye genomic resources, contributing greatly to the application of molecular breeding techniques in the small-grain cereal. These include the advent of chromosomal-scale reference genomes of an German grain-type inbred line 'Lo7' (Rabanus-Wallace et al., 2021) and Chinese forage-type rye inbred line 'Weining' (Li et al., 2021). Furthermore, a 600K high-density single nucleotide polymorphism (SNP) array (Bauer et al., 2017) has been developed as a successor or compliment to the previous 5K array (Haseneyer et al., 2011).

The application of marker-assisted selection (MAS) is highly suited for disease resistance often conferred by mono-, di- or oligogenic inherited *R* genes and QTLs (Chhetri et al., 2017; Ma et al., 2018). MAS has been proved effective for (i) resistance that depends on a specific environment or host developmental stage (Nsabiyera et al., 2018), (ii) guided backcrossing of *R* genes or QTLs into elite background (Yadav et al., 2015), and (iii) pyramiding of multiple *R* genes or QTLs for broad-spectrum resistance and enhanced durability (Liu et al., 2020). The success of MAS is highly dependent on the quality of the trait-linked markers (Cobb et al., 2019). Recently, Platten et al. (2019) proposed a set of five core metrics to describe and quantify marker reliability to correctly classify germplasm as QTL[+]/[-]. The core metrics included, (i) call rate, (ii) clarity, (iii) false positive rate, (iv) false negative rate, and (v) utility. In this context the recent advent of a high-density SNP array in rye constitutes an important tool, providing a vastly higher marker resolution for the identification of reliable markers situated in close proximity to the *R* gene or QTL of interest (Bauer et al., 2017). Despite the use of high quality markers MAS nonetheless remains constrained by the fixation of large genomic regions with potential linkage drag of undesirable QTLs with adverse effects on other agronomically important traits, *e.g.*, yield penalty (Miedaner et al., 2017; Chitwood-Brown et al., 2021). However, the rate of linkage decay in outcrossing species such as rye is often rapid due to their inherent heterogeneity with a linkage decay within 4 kb observed in a similar hybrid breeding system in maize (Wu et al., 2016). The rapid rate of decay, while providing the means for a more targeted introgression of a smaller genomic region with reduced risk of linkage drag requires a considerably higher marker resolution than needed for self-pollinating crops such as barley and wheat (Brar et al., 2019; Huang et al., 2021).

In Europe most private breeding companies have invested in the infrastructure for high-throughput MAS, constituting an integral part in commercial breeding programs (Miedaner and Korzun, 2012). To increase the selection intensity and gain integrated plant breeding platforms have been developed combining MAS strategies with speed breeding approaches, allowing a rapid development of resistance cultivars with multiple *R* genes and QTLs (Hickey et al., 2017). Use of genome-wide association studies (GWAS) has become a routine strategy in plant breeding for investigating the genetic basis of resistance and identification of markers linked to *R* genes and QTLs (Liu and Yan, 2019; Alqudah et al., 2020). GWAS uses linkage disequilibrium between a non-causal genetic variant and a gene of interest to determine a statistical association (Liu and Yan, 2019). The power of GWAS to identify true causal relationships between a genomic variant, often a SNP, and a phenotypic trait depends on the phenotypic variance explained by the SNP. The phenotypic variance is determined by the how strongly the two allelic variants of the SNP differ in their phenotypic effect and their frequency (Korte and Farlow, 2013).

The successful application of GWAS, however, depends on several biological and statistical factors. These include (i) phenotypic variation, (ii) sample size and genetic heterogeneity, (iii) allele frequency, (iv) population structure, and (v) linkage disequilibrium (Korte and Farlow, 2013; Liu and Yan, 2019; Alqudah et al., 2020). Depending on the genetic characteristics of the plant species, *e.g.*, genetic heterogeneity, phenotypic variation and linkage disequilibrium, and complexity of trait, size sample in the range of 100 to 500 individuals are needed for GWAS (Korte and Farlow, 2013). The discovery of rare *R* genes and low to intermediate effect QTLs in collections of breeding lines and diversity panels are limited due to a low level of phenotypic variance explained, impeding the establishment of a statistical marker-trait association in GWAS. If the analytical objective is the identification of such rare *R* genes or low to intermediate effect QTLs, coupling GWAS with multi-parent mapping populations have proven a successful strategy (Ren et al., 2017; Novakazi et al., 2020; Scott et al., 2020). Population structure and unequal relatedness amongst individual introduces a confounding effect that might give cause to spurious marker associations, introducing a risk of false positives (Vilhjalmsson and Nordborg, 2013). To account for population structure most analytical software used for GWAS, such as Genome Association and Prediction Integrated Tool (GAPIT) (Wang and Zhang, 2021), integrates a mixed linear model (Sul et al., 2018). The presence of long-distance LD, however, introduces a uncorrectable confounding factor, complicating the

disentanglement of actual causal variants from linked neutral markers leading to potential spurious associations.

5.1 Resistance gene enrichment sequencing

In 2013 the *R* gene enrichment and sequencing (RenSeq) workflow was developed by Jupe et al. (2013), for discovery and reannotation of the NLR repertoire in plants. Exploiting the conserved nature of NLRs, capture libraries comprised of biotinylated 110-130 nt oligo baits are designed on basis of known NLR sequences (Figure 5.11, Jupe et al., 2012). Allowing up to 20% sequence discrepancy between bait and NLR sequence, the bait library can successfully be deployed in plants exhibiting considerable NLR sequence divergence from the original bait library design (Manuscript IV Figure 3B, Jupe et al., 2013).

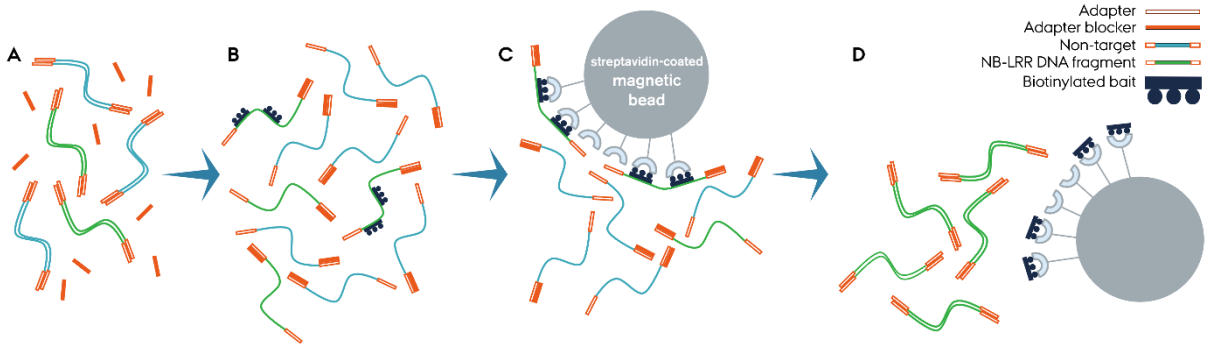


Figure 5.11 Target capture of nucleotide-binding leucine rich (NB-LRR) sequences in genomic plant DNA for Resistance gene enrichment sequencing (RenSeq) analysis. **A**) pooling of reaction components, **B**) hybridization of blockers to adapters prior to introduction and hybridization of 110-130 nucleotide biotinylated NB-LRR specific oligo baits to DNA library. **C**) target capture of NB-LRR DNA fragments by streptavidin-coated magnetic beads, sequestered with a magnet, **D**) non-target DNA is washed away and NB-LRR library is enriched (DAB, 2021).

At present, RenSeq has been deployed in mouse-ear cress (*Arabidopsis thaliana*) (Van de Weyer et al., 2019), strawberry (*Fragaria ananassa*) (Barbey et al., 2019), potato (*Solanum tuberosum* L.) (Witek et al., 2016), tomato (*Solanum lycopersicum* L.) (Seong et al., 2019), hexaploid bread wheat (*Triticum aestivum* L.) (Steuernagel et al., 2016), wheat-*Dasypyrum villosum* L. introgression line (Xing et al., 2018), and the wheat wild relatives *Aegilops tauschii* (Arora et al., 2019), and *Ae. peregrina* (Narang et al., 2020).

For the discovery of stem rust *R* genes in a wild *Ae. tauschii* diversity panel Arora et al. (2019) combined *k*-mer based association genetics with RenSeq leading to the identification of *Sr33*, *Sr45*, *Sr46* and *SrTA1662* (Figure 5.12). With no dependency on a high quality reference genome, AgRenSeq constitute a promising tool for accessing the genetic reservoir of novel *R* genes in wild ancestors for crop improvement (Li et al., 2020; Zhang and Batley, 2020). In

contrast to laborious positional cloning or mutational genomics approaches AgRenSeq enables a rapid *R* gene cloning (Steuernagel et al., 2016; Lu et al., 2020). The discovered *R* gene(s) can be used for the development of high quality functional markers for marker-assisted breeding, or a resource for future engineering of genetically modified crops (GMO) (Schenke and Cai, 2020; Tripathi et al., 2020).

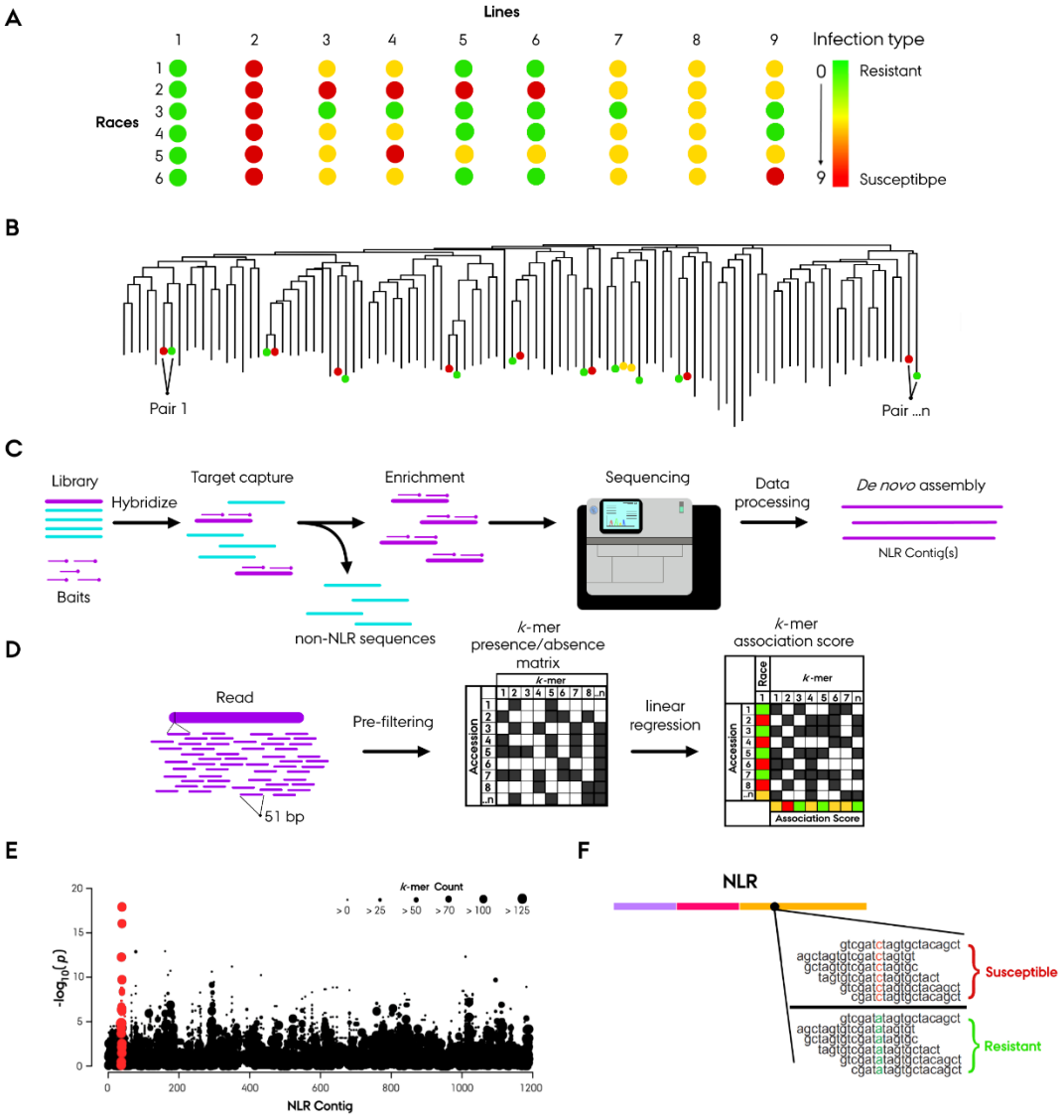


Figure 5.12 Combined association genetics and resistance (*R*) gene enrichment and sequencing. **A**) phenotyping of a diverse panel of breeding lines using distinct pathogen isolates or races, **B**) Pair-wise selection of closely related lines with diverging infection type, **C**) Selected lines are subjected to RenSeq and NLR repertoire assembled *de novo*, **D)** subreads (*k*-mers) of 51 bp are extracted from raw reads, pre-filtered based on the correlation of their presence/absence to the level of resistance or susceptibility in the phenotyped panel. Each pre-filtered *k*-mer is given a *P* value based on its ability to predict the phenotype using linear regression to control for population structure. **E**) *k*-mers are plotted in an association matrix according to their sequence identity to NLRs from a given resistant line (x-axis) and the measure of their association with phenotype (y-axis). A candidate *R* gene contig is illustrated by a red-dot column. **F**) multiple alignment of candidate *R* gene across susceptible and resistant lines for identification of resistance-specific single nucleotide variant or insertion/deletion. After Arora et al. (2019).

Objectives

The following objectives were pursued in the project

I. Analyze the genetic diversity and population structure of Nordic Seed hybrid rye elite breeding material

II. Investigate the genetics underlying male-fertility restoration in the Gülzow type CMS breeding system of rye

III. Investigate *Puccinia recondita* f. sp. *secalis* population structure by appropriate DNA marker analysis and virulence phenotyping

IV. Establishment of efficient disease resistance phenotyping method for powdery mildew and leaf rust resistance in Nordic Seed hybrid rye components

V. Phenotyping of agronomic important traits and resistance to biotic stresses in Nordic Seed hybrid rye components with emphasis on powdery mildew and brown rust disease

VI. Identify trait-linked SNP markers by means of genome-wide association study (GWAS) for direct implementation in marker assisted selection in Nordic Seed hybrid rye breeding program

VII. Establishment of bioinformatic pipeline for identifying resistance genes in rye components by single-molecule real-time (SMRT) Association Genetics Resistance Gene Enrichment Sequencing (AgRenSeq) analysis

Hypothesis

The following hypothesis were explored in the project

I. Genetic separation of Nordic Seed hybrid rye elite breeding component populations

II. Moderate level of genetic diversity in the Nordic Seed hybrid rye elite breeding component populations

III. Presence of major and minor restoration of male fertility genes in the Gülzow type fertility control system

IV. Spatiotemporal dependent variation in the *Puccinia recondita* f. sp. *secalis* population in Denmark

V. Variation in the interaction between rye lines and *Puccinia recondita* f. sp. *secalis* isolates

VI. Variation in the interaction between rye lines and *Blumeria graminis* f. sp. *secalis* populations

VII. SNP markers linked to important agronomic traits or biotic resistance can be identified by GWAS in rye components

VIII. Leaf rust resistance genes in rye components can be identified by SMRT AgRenSeq

Manuscript I

Genetic structure of a germplasm for hybrid breeding in rye (*Secale cereale* L.)

By

Nikolaj M. Vendelbo[1,3*], Pernille M. Sarup[1], Jihad Orabi[1], Peter S. Kristensen[2], and Ahmed Jahoor[1,4]

[1] Nordic Seed A/S, Odder, Denmark

[2] Nordic Seed Germany GmbH, Nienstädt, Germany

[3] Department of Agroecology, Aarhus University, Slagelse, Denmark

[4] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

Publication details

Published in PloS one, vol. 15, issue 10, e0239541 , 2020

PLOS ONE

RESEARCH ARTICLE

Genetic structure of a germplasm for hybrid breeding in rye (*Secale cereale* L.)

Nikolaj M. Vendelbo[1,2]*, Pernille Sarup[1], Jihad Orabi[1], Peter S. Kristensen[3], Ahmed Jahoor[1,4]

1 Nordic Seed A/S, Odder, Denmark, **2** Department of Agroecology, Aarhus University, Slagelse, Denmark, **3** Nordic Seed Germany GmbH, Nienstädt, Germany, **4** Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

* nive@nordicseed.com

OPEN ACCESS

Citation: Vendelbo NM, Sarup P, Orabi J, Kristensen PS, Jahoor A (2020) Genetic structure of a germplasm for hybrid breeding in rye (*Secale cereale* L.). PLoS ONE 15(10): e0239541. https://doi.org/10.1371/journal.pone.0239541

Editor: Dragan Perovic, Julius Kuhn-Institut, GERMANY

Received: May 31, 2020

Accepted: September 9, 2020

Published: October 9, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0239541

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The research was funded by Innovation Fund Denmark (grant no. 8053-00085B),

Abstract

Rye (*Secale cereale* L.) responds strongly to changes in heterozygosity with hybrids portraying strong heterosis effect on all developmental and yielding characteristics. In order to achieve the highest potential heterosis effect parental lines must originate from genetically distinct gene pools. Here we report the first comprehensive SNP-based population study of an elite germplasm using fertilization control system for hybrid breeding in rye that is genetically different to the predominating P-type. In total 376 inbred lines from Nordic Seed Germany GmbH were genotyped for 4419 polymorphic SNPs. The aim of this study was to confirm and quantify the genetic separation of parental populations, unveil their genetic characteristics and investigate underlying population structures. Through a palette of complimenting analysis, we confirmed a strong genetic differentiation (F_{ST} = 0.332) of parental populations validating the germplasms suitability for hybrid breeding. These were, furthermore, found to diverge considerably in several features with the maternal population portraying a strong population structure characterized by a narrow genetic profile, small effective population size and high genome-wise linkage disequilibrium. We propose that the employed male-sterility system putatively constitutes a population determining parameter by influencing the rate of introducing novel genetic variation to the parental populations. Functional analysis of linkage blocks led to identification of a conserved segment on the distal 4RL chromosomal region annotated to the *Rfp3* male-fertility restoration 'Pampa' type gene. Findings of our study emphasized the immediate value of comprehensive population studies on elite breeding germplasms as a pre-requisite for application of genomic-based breeding techniques, introgression of novel material and to support breeder decision-making.

Introduction

Rye (*Secale cereale* spp. *cereale*., genome RR, 2n = 2x = 14) is a cereal crop species belonging to the botanical tribe *Triticea* within the grass family *Poaceae* and a close relative to common wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.) [1]. Rye is believed to have

Pajbjergfoundation, and Nordic Seed A/S. The funder provided support in the form of salaries for authors NMV, PMS, JO, PSK, and, AJ, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing interests: All authors are employees in the plant breeding company Nordic Seed A/S. The employment does not alter the authors' adherence to all of the PLOS ONE policies on sharing data and materials.

evolved as a hitchhiking weed of its progenitor, an annual wild rye (*S. cereale* spp. *vavilovii*) in cultivated fields of founder crops around 12.000 years ago in the Fertile Crescent of the Near East, acquiring key agronomic traits by co-domestication events [2–4]. First evidence of deliberate rye cultivation dates to the European bronze age 5000 years ago and rye has since remained a prominent crop in the temperate northern hemisphere covering 4.12 million hectares in 2018 [5, 6]. Recognized for its tolerance to abiotic and biotic stress, rye exhibits a superior high yield potential under marginal conditions. In mid-19th century, targeted breeding of rye was initiated in Germany, Poland and Russia leading to the generation of secluded germplasms of which the best known are the 'Petkus' and 'Carsten' gene pools [7, 8]. As an allogamous plant species, self-incompatibility has only recently been overcome in cultivated rye after introgression of a dominant self-fertility gene [9]. Breeding lines in rye are therefore predisposed to severe inbreeding depression, when repetitively self-fertilized due to the high load of recessive deleterious mutations accumulated in highly heterozygous species [10]. Inbred rye lines consequently respond very strongly to changes in heterozygosity with hybrids demonstrating a heterosis effect on all developmental and yielding characteristics with *e.g.* grain yield displaying a 110–140% average relative to midparent performance [11–13].With the discovery of cytoplasmic male-sterility (CMS) alleles, breeding effort shifted from open-pollinating varieties (OPVs) to hybrids [14]. Breeding of hybrids in rye rely on a three-way cross of the formula (A_{CMS} x B_{NRG}) x R, involving three components, a maternal line carrying the CMS allele, a non-restorer germplasm (NRG) for maintaining the CMS and a restorer (R) pollen father carrying the restorer of fertility (Rf) allele [15]. In order to achieve the highest potential of heterosis effect, parental lines should originate from genetically separate, heterotic gene pools [16]. In recent years, technological and scientific advances have expedited the implementation of genome-based breeding strategies in rye for a more efficient exploration of the genetic potential [17, 18]. Implementation of marker assisted selection and genomic selection relying on molecular DNA markers has enhanced the selection efficiency leading to increased genetic gains surpassing traditional breeding methods [12, 13, 19]. In 2011 Haseneyer *et al.* [20] published a 5K single nucleotide polymorphism (SNP) array for rye, succeeded by the 600K array by Bauer *et al.* [1]. Recently Rabanus-Wallace *et al.* [21] published the first chromosome-scale assembly of the 7.9 Gbp rye genome. In conjunction, these genomic resources represent significant milestones, providing indispensable tools for elucidating genes underlying important agronomic traits, implementation of genomic-based breeding techniques and dissecting population genetics in rye [22]. Through a comprehensive understanding of the genetic architecture, wild relatives, OPVs, and landraces can be efficiently exploited through selected introgression, broadening the genetic basis of elite breeding germplasm and providing access to a rich reservoir of genetic diversity for ensuring continued genetic gains in hybrid breeding [16, 23, 24]. SNP arrays constitute a high-throughput genotyping platform readily implemented by breeding programs for assaying the genetic variation and identifying trait-linked markers by genome-wide association studies for marker assisted selection in elite breeding germplasm. Numerous population studies have been conducted on the *Secale* genus, open-pollinating rye varieties and inbred populations using simple sequence repeats (SSR), diversity array technology (DArT), and, random amplification of polymorphic DNA (RAPD) marker systems [25–27]. In the extensive study by Bauer *et al.* [1], they validated their 600K SNP array in a smaller population study on a diverse panel of wild *Secale* sp. accessions, including inbred lines from heterotic gene pools.

In this study, we report a comprehensive SNP-based study of genetic diversity and population structure in a hybrid rye elite breeding germplasm as prerequisite for application of genomic-based breeding techniques, introgression of novel material and to support breeder decision-making. Our aim was to I) confirm and quantify the genetic separation between

parental populations, II) gain a comprehensive understanding of the genetic characteristics, features and architecture of the parental populations, and, III) Investigate underlying sub-population structures.

Materials & methods

Plant material

In total 376 Nordic Seed A/S inbred hybrid rye (*Secale cereale* L.) elite breeding component lines were selected for the study, comprising 250 restorer, 119 non-restorer germplasm (NRG) and 7 cytoplasmic male-sterile (CMS) lines. The restorer lines originated from the 'Petkus' and 'Carsten' gene pools with a predominance of the latter suggested by available pedigree data, whereas information on the precise origin of the NRG lines is not-existent. The CMS male sterility is based on the 'Gülzow' (G) type cytoplasm originating from the German population rye variety 'Schlägler alt' [28].

DNA extraction

Four seeds of each line were sown in 104 hole plates (51.5 x 31 x 4.5cm), containing fine-grain sphagnum substrate. Plants were cultivated at Nordic Seed A/S greenhouse facilities under natural light and manual irrigation, with night temperatures of 14–16˚C and day temperatures of 18–24˚C. Seven days after sowing, the lowest section of two coleoptiles and primary leaves were cut, equivalent to 75 mg plant material, and placed in a 96-well Micro-Dilution Tube System (STARLAB International GmbH) containing two 4mm glass beads per 1.2 mL tubes. Plant tissue samples were stored at -20˚C for two days prior to freeze drying for an additional two days in a 9L Coolsafe™ (LaboGene) apparatus. Preceding the DNA extraction, leaves were crushed at 4000 RPM using a TissueLyser II (Qiagen®) bead mill, and DNA extracted using an SDS-based extraction method. DNA concentration and 260/280nm ratio of samples were measured using Epoch™ microplate spectrophotometer (Biotek®) and quality, *i.e.* evidence of fragmentation by size-visualization on a 1.2% agarose gel.

Molecular marker resources and SNP genotyping

Samples containing 100 ng and 1.7–1.9 260/280nm ratio long-stranded DNA were sent for single nucleotide polymorphism (SNP) genotyping at TraitGenetics GmbH using a pre-designed Illumina Infinium $15K_{wheat}$ and $5K_{Rye}$ SNP iSelect ultra HD chip. Custom rye specific SNPs were comprised of 2698 markers from the 5K array by Haseneyer *et al.* [20] denoted as '5K-set' and 2059 markers from the 600K array by Bauer *et al.* [1] denoted as '600K-set'. Wheat SNPs comprised of 12908 markers originating from the 90K array by Wang *et al.* [29] denoted 'wheat-set'. The markers were prior to analysis filtered for marker allele frequency ≥ 0.005, missing individual score ≤ 0.2, and, missing marker score of ≤ 0.1.

Data analysis

Population genetic analysis of SNP marker data was done in R studio (v. 1.1.463) interface in R statistical software (v. 3.6.3) by application of various predesigned packages [30, 31]. Marker minor allele frequency (MAF), polymorphic information content (PIC), and, effective population size (N_e) was calculated using SnpReady (v. 0.9.6) R package [32]. Calculation of the genetic characteristics of accessions, *i.e.* observed heterozygosity (H_o), within population gene diversity (H_s), and Wrights F statistics, inbreeding coefficient (F_{IS}) and fixation indices (F_{ST}) per population and identified sub-populations was done using Hierfstat (v. 0.04–22) R package [33, 34]. Parental pool differentiating SNPs were identified by calculating the marker F_{ST} using

Pegas (v. 0.13) R package [35]. Principal component analysis (PCA) was conducted using inherent functions in R with 3D PCA constructed using rgl (v. 0.100.50) R package [36]. Shared ancestry between accessions were estimated by Admixture model using STRUCTURE (v. 2.3.4) software, assigning each locus probabilistically to an ancestral founder (or subpopulation) K per breeding line on basis of allele frequencies [37]. The model was run with burning period set to 200.000 and Markov chain Monte Carlo to 400.000 iterations, with the results visualized in R studio. Estimate of the most likely number of K was done using the DeltaK method by Evanno *et al.* [38] in Structure Harvester (v. 0.6.94) at K set to 1 to 5 with 5 replicates per level [39]. The analysis was conducted on the complete germplasm to investigate the number of genetically distinct ancestral founders and for each of the parental populations to identify the most likely number of subpopulations. Phylogenetic analysis comprised of a neighbor-joining clustering of breeding lines with Euclidean genetic distance measure using ape (v. 5.3) R package [40]. The tree was constructed after 10.000 bootstrapping iterations with weak nodes (80% recurrence) collapsed into multifurcations and generated using iTOL (v. 5) online tool (http://itol.embl.de/) [41]. Fisher's exact test was conducted to test whether the structural composition, *i.e.* no. of 'clades' and 'singletons' in the phylogenetic tree diverged between populations using R inherent functions. Analysis of pairwise linkage disequilibrium (LD) in the parental populations was done using SnpStats (v. 1.36.0) R package with depth set to 6 and LD estimated as the coefficient of determination (r^2) [42]. Prior to analysis, a filtration step was included to remove non-informative population-specific monomorphic markers. Heatmap of the pairwise LD was constructed using LDheatmap (v. 0.99–7) R package [43]. Genetic map of markers and visualization of linkage decay was done using Sommer (v. 2.9) R package with filtration for marker correlation (r^2) significance at $p<0.0001$ [44]. The intersection of the LOESS curve fit to the baseline of population-specific critical value of r^2, *i.e.* the estimated linkage disequilibrium (LD) background was set as the delimiter of the linkage decay interval [45]. LD and rate of decay was calculated per parental population the 5K-set and 600K-set markers separately due to divergence in the arrays genetic distance measure.

Results

Molecular marker resources and SNP genotyping

Quality filtration of markers for low minor allele frequency (MAF), missing marker, and missing individual across the entire panel led to the identification of 4419 high-quality SNP markers distributed evenly amongst the seven rye chromosomes. The filtered marker panel comprised of 2364 5K-set markers, 1629 600K-set markers, and 426 wheat-set markers. In terms of marker-set performance and informativeness across the entire germplasm a mean MAF of 0.254, 0.250, 0.201 and mean PIC of 0.268, 0.262, and 0.225 was determined the 5K-set, 600K-set and wheat-set markers respectively. Mean MAF was 0.188 across the markers in the NRG&CMS population, omitting 1870 monomorphic markers, with ~50% of the polymorphic markers exhibiting a MAF ≤ 0.1 (Fig 1A). The restorer population portrayed a homogeneous distribution of MAF frequencies with a mean of 0.251, omitting only 38 monomorphic markers. Quality filtration was likewise conducted on the separate parental populations to determine the effect of population size difference. The test yielded an additional 82 high-quality markers for the NRG&CMS population, while 47 for the restorer population. Mean polymorphic information content (PIC) was 0.202 in the NRG&CMS post filtration of monomorphic markers with $\approx$ 35% exhibiting a PIC ≤ 0.1 (Fig 1B). In the restorer population, mean PIC across polymorphic markers was 0.261, with $\approx$ 50% exhibiting a PIC value between 0.3 and 0.4. Overall mean PIC was estimated to 0.261.

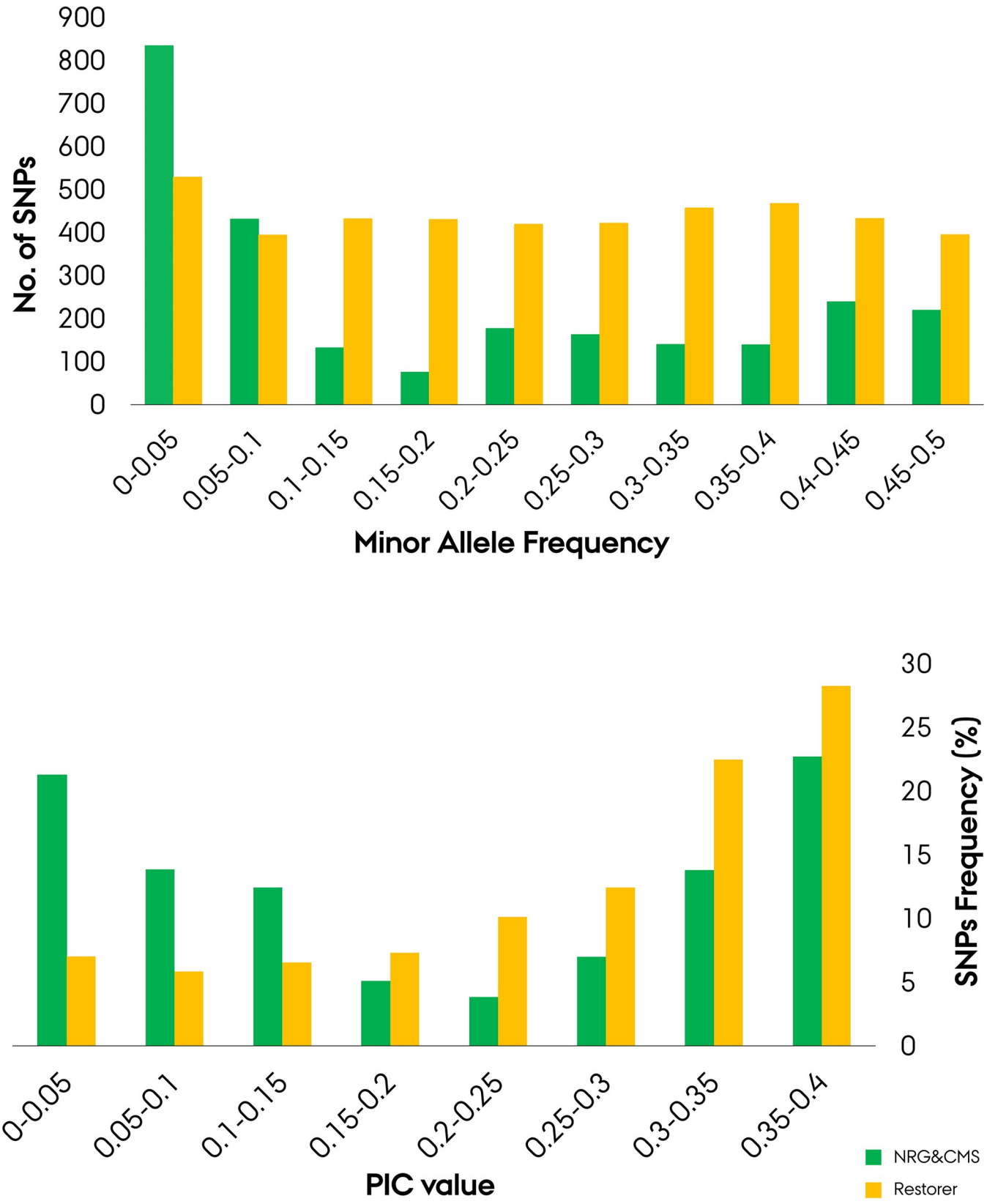


Fig 1. Population-wise distribution of minor allele frequency (MAF) (**A**) and polymorphic information content (PIC) (**B**) for 4419 SNP markers in the Nordic Seed elite hybrid rye breeding germplasm comprising a seed mother (NRG&CMS) and pollen father (Restorer) population.

https://doi.org/10.1371/journal.pone.0239541.g001

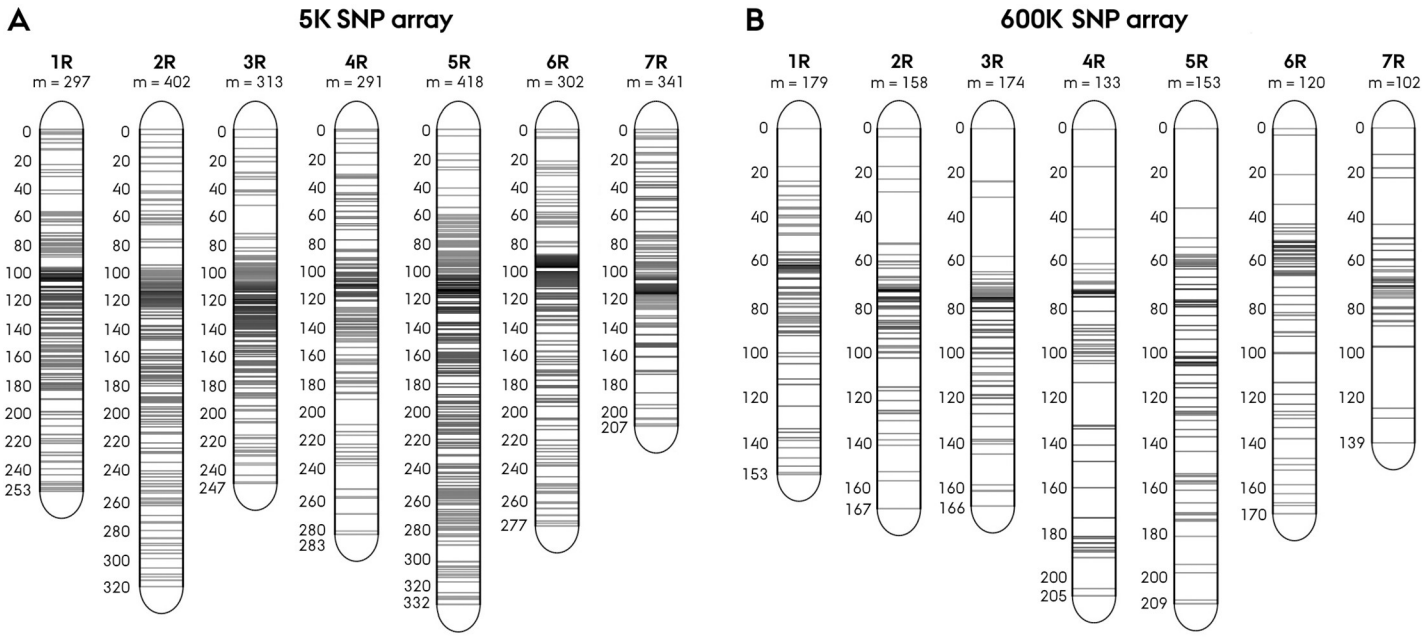


Fig 2. Genetic map of SNP markers in rye. **A**) 2364 markers from the 5K array and **B**) 1019 from the 600K array.

https://doi.org/10.1371/journal.pone.0239541.g002

Genetic mapping of the 3383 mapped markers revealed an excess of markers mapping to the peri-centromeric region in both the 5K-set and 600K-set markers, increasing the inter-loci distance at the proximal and distal region (Fig 2A and 2B). None of the wheat-set markers had a mapping position in the rye genome.

Genetic characteristics

Calculation of fundamental genetic characteristics led to the initial discovery of 11 lines portraying evidence of an inadvertent cross fertilization event between the two parental populations. All these 11 lines displayed a high proportion of the opposing ancestral component in an analysis of inferred ancestry at K set to 2 with a mean of 56%. Moreover, 4 lines displayed a mean residual observed heterozygosity (H_o) of 41%, suggesting a recent inadvertent cross fertilization (S1 Table). Together, these results led to the decision to discard these lines from further analysis leaving a trimmed population comprised of 242 restorer, 116 NRG and 7 CMS component lines.

In the remaining trimmed population, the NRG&CMS population displayed a mean H_o of 3.6%, within population gene diversity (H_s) of 0.250, inbreeding coefficient (F_{IS}) of 0.852, and effective population size (N_e) of 72, equivalent to a N_e/N of 0.587. The restorer population displayed a H_o of 3.8%, H_s of 0.333, F_{IS} of 0.886, and, N_e of 137 equivalent to a N_e/N of 0.565. Fixation indices (F_{ST}) between the two populations was estimated to 0.332.

Principal component analysis

Initial analysis conducted on the complete set of 376 Nordic Seed hybrid rye elite breeding component lines comprised of a principal component (PC) analysis plot visualizing the genetic architecture across various PC combinations (Fig 3). While the analysis provides an estimate of the proportion of variation explained by the individual PCs, the PCA primarily served in our study to get a firsthand visual impression of the genetic diversity and architecture within

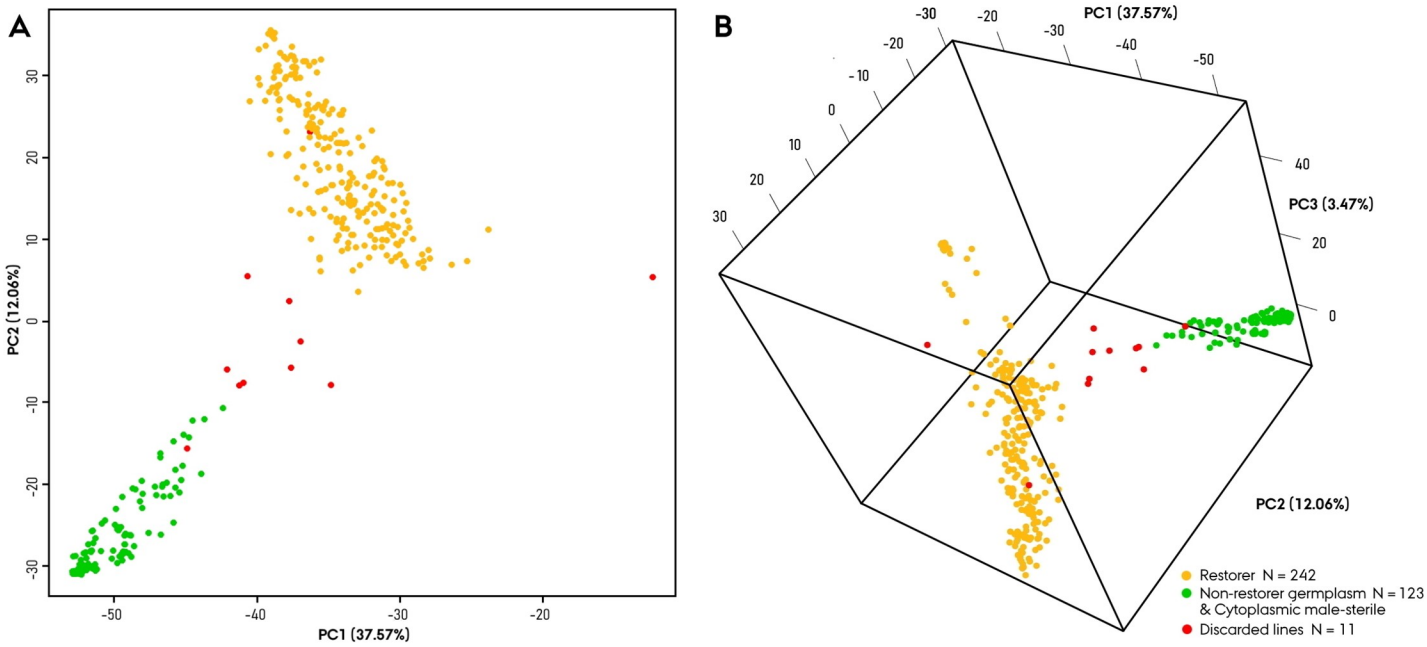


Fig 3. Principal component analysis of the Nordic Seed hybrid rye elite breeding germplasm (N_{pop} = 376), based on 4419 SNP markers, (**A**) 2-D plot of principal component (PC) 1 and 2, (**B**) 3-D plot of PC1, 2 and 3. Red dots indicate discarded lines.

https://doi.org/10.1371/journal.pone.0239541.g003

the material. In the constructed 2-D PCA plot depicting PC1 (37.57%) and PC2 (12.06%), lines were observed to group into two distinct clusters in accordance to their population, restorer and NRG&CMS (Fig 3A). With addition of PC3 (3.47%) a 3-dimensional PCA plot led to further unfolding of the restorer populations genetic architecture, while the NRG&CMS population remained closely grouped. Whilst informative as a static graph, visualization of the PC in a 3D interactive software facilitates a more comprehensive mining of data structures. Lastly, the PCA, furthermore, served to validate the discontinuation of 11 lines with majority of these observed in the intermediary space diluting the distinct genetic separation between two parental population.

Phylogenetic analysis

Hierarchical clustering analysis of the trimmed Nordic Seed hybrid rye elite breeding population supported the distinct genetic separation of the two parental populations visualized in a circularized neighbor-joining dendrogram (Fig 4). In order to validate the phylogenetic tree, weak nodes showing less than 80% recurrence were collapsed into multifurcations after 10.000 bootstrapping iterations. This stringent validation-step dramatically reduced the number of inconsistent family clades within the dendrogram, and hence the overall structural complexity in the germplasm. In the constructed neighbor-joining dendrogram the populations were found to exhibit a significantly different structural composition ($p = 0.028$) in a Fisher's exact test. The NRG&CMS population portrayed 25 clades comprising on average of 5 lines and 11 singletons relative to the restorer population that portrayed 41 clades of similar mean size and 51 singletons.

Inferred ancestry & sub-population discovery

In order to dissect the underlying sub-population structure within the trimmed elite Nordic Seed hybrid rye breeding germplasm, an Admixture analysis of inferred ancestry was

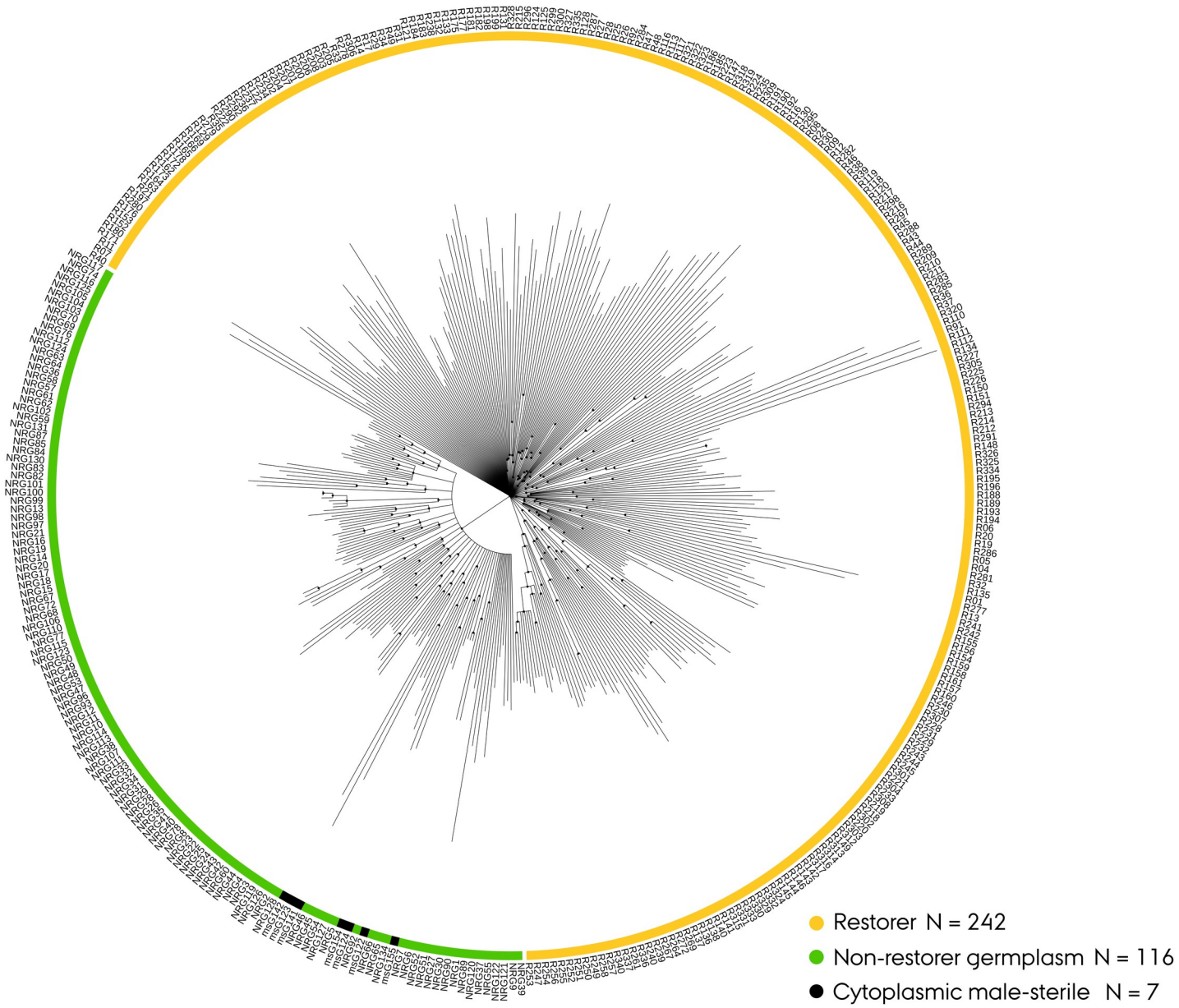


Fig 4. Circularized neighbor-joining dendrogram of the trimmed Nordic Seed hybrid rye elite breeding germplasm (N_{pop} = 365) based on 4419 SNP markers, after bootstrapping with recurring nodes shown as black dots.

https://doi.org/10.1371/journal.pone.0239541.g004

conducted. This model-based clustering method assigned individual lines based on their allele frequencies to K clusters, *i.e.* ancestral founder populations or sub-populations. Output of the admixture model is a frequency distribution of K ancestral sub-population components for each component breeding line easily presented in a bar plot and colorized according to K clusters (Fig 5). Component breeding lines were aligned in accordance to their phylogenetic relationship in order to support the identification of underlying sub-population structures (Fig 5A and 5F). First step was to validate the existence of two genetically distinct ancestral founders which the parental populations originate by conducting the Admixture model on the entire trimmed population at K set to 2 (Fig 5B and 5G). This was done using Evanno's DeltaK method confirming a historical separation of the two parental gene pools (ΔK = 2) with the

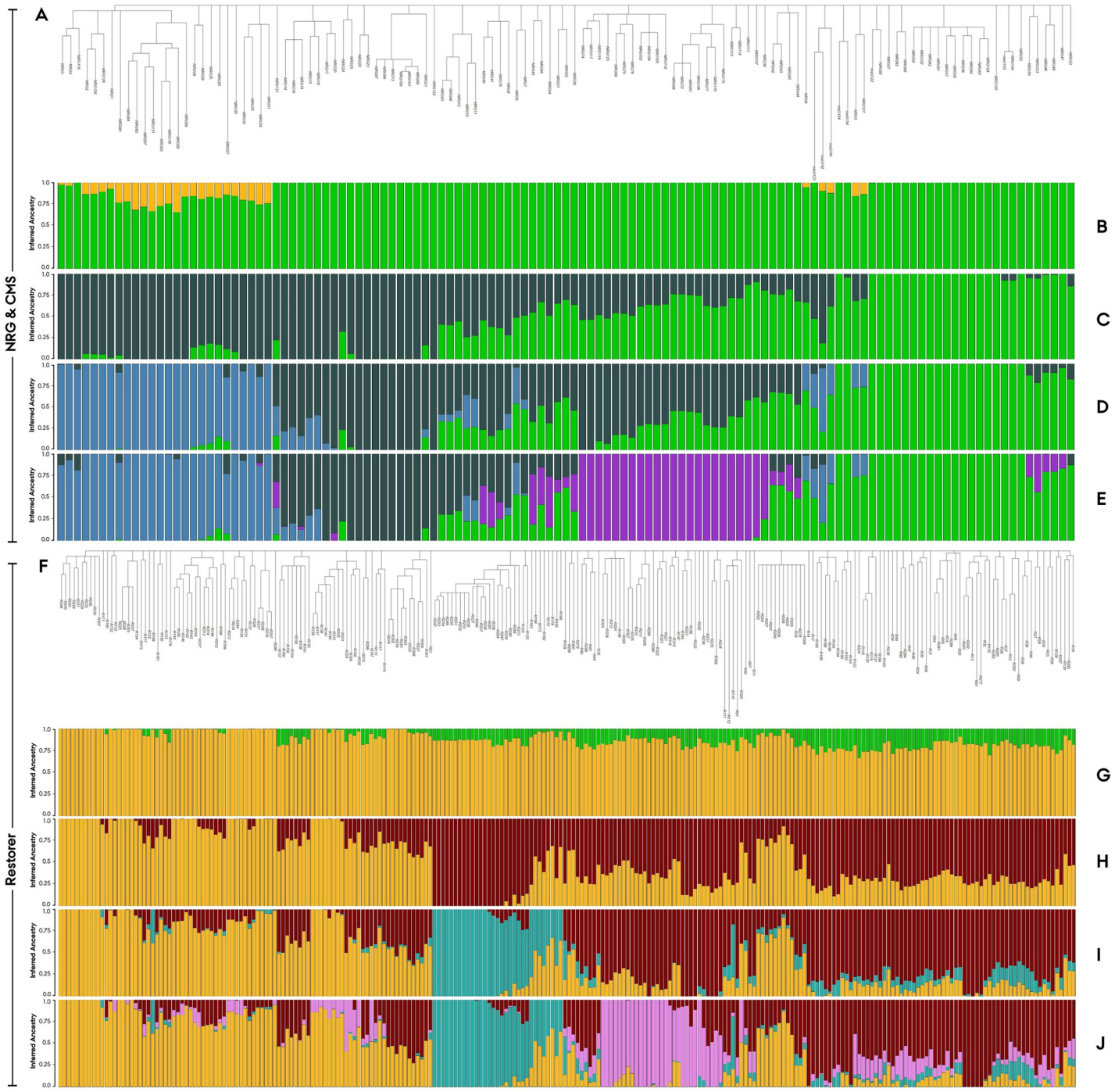


Fig 5. Admixture model of inferred ancestry and neighbor-joining dendrogram of the trimmed Nordic Seed hybrid rye elite breeding germplasm based on 4419 SNP markers. Neighbor-joining dendrogram of the parental populations, (**A**) Seed parent (NRG&CMS, N = 123), and, (**F**) Pollen parent (R, N = 242). Analysis of co-ancestry between parental populations (**B**) NRG&CMS (green) and (**G**) R population (yellow) at K set to 2. Discovery of sub-population structure at K set to 2, 3, 4 respectively for the (**C**,**D**,**E**) NRG&CMS, and (**H**,**I**,**J**) Restorer lines. Colors define fraction of inferred ancestral sub-population.

https://doi.org/10.1371/journal.pone.0239541.g005

Admixture model revealing limited extend of co-ancestry. The NRG&CMS population exhibited a mean co-ancestry of 4.3% shared with the restorer population across a subset of admixed lines (Fig 5B), whereas the restorer population exhibited a mean co-ancestry of 11.6% spread

Table 1. Pairwise fixation indices (F_{ST}), observed heterozygosity (H_o), within subpopulation gene diversity (H_s), and inbreeding coefficient (F_{IS}) of parental subpopulations (SP) within the Nordic Seed hybrid rye elite breeding germplasm (A) seed mother (NRG&CMS, N = 123), and (B) pollen father (Restorer, N = 242). SP constitutes n lines with ≥ 0.95 Admixture inferred ancestry at K inferred subpopulations.

	A	Non-restorer germplasm & Cytoplasmic male-sterile									B	Restorer								
		Pairwise F_{ST}					Genetic characteristics				Pairwise F_{ST}					Genetic characteristics				
Inferred subpopulations (K)	K = 4		SP1	SP2	SP3	SP4	n	H_o	H_s	F_{IS}		SP1	SP2	SP3	SP4	n	H_o	H_s	F_{IS}
		SP1	0				21	0.016	0.054	0.706	SP1	0				15	0.021	0.106	0.801
		SP2	0.347	0			16	0.025	0.133	0.812	SP2	0.694	0			14	0.012	0.033	0.645
		SP3	0.435	0.490	0		22	0.016	0.034	0.524	SP3	0.437	0.594	0		11	0.034	0.198	0.824
		SP4	0.485	0.555	0.489	0	11	0.005	0.046	0.900	SP4	0.420	0.570	0.304	0	7	0.084	0.221	0.622
	K = 32		SP1	SP2	SP3		n	H_o	H_s	F_{IS}		SP1	SP2	SP3		n	H_o	H_s	F_{IS}
		SP1	0				22	0.016	0.055	0.716	SP1	0				26	0.018	0.164	0.893
		SP2	0.492	0			18	0.024	0.126	0.806	SP2	0.524	0			14	0.012	0.033	0.645
		SP3	0.514	0.323	0		13	0.004	0.058	0.931	SP3	0.293	0.484	0		13	0.058	0.261	0.780
	K = 2		SP1	SP2			n	H_o	H_s	F_{IS}		SP1	SP2			n	H_o	H_s	F_{IS}
		SP1	0	0.560			23	0.017	0.055	0.691	SP1	0	0.436			41	0.017	0.185	0.906
		SP2	0.560	0			33	0.020	0.182	0.842	SP2	0.436	0			19	0.014	0.062	0.778

https://doi.org/10.1371/journal.pone.0239541.t001

more homogeneously across the entire panel (Fig 5G). Evanno's DeltaK method was furthermore applied to determine the most likely estimated number of founders for the individual parental populations. Analysis suggested 3 ancestral founders in both the NRG&CMS and restorer population. To examine the underlying sub-population structure, K was set to 2, 3, and 4 within the individual populations (Fig 5C–5E and 5H–5J). For each of the identified subpopulations, H_o, H_e, F_{IS} and pairwise F_{ST} was estimated to dissect their genetic characteristics and differentiation (Table 1). Breeding lines portraying a predominance of a single ancestral subpopulation component (≥ 0.95) was found to cluster according to their position within the neighbor-joining dendrogram. Despite considerable divergence in parental population size the number of lines assigned to subpopulations were predominantly similar, indicative of a more admixed, genetically complex population structure within the restorer.

Linkage disequilibrium & decay

Non-random association of alleles at different loci, *i.e.* linkage disequilibrium (LD) and the decay of linkage are important genetic characteristics of breeding germplasms providing insight into the marker density essential for trait-linked marker discovery by association genetic studies. In our study, the intra- and inter-chromosomal, *i.e.* genome-wise, LD was calculated for the two populations using 3383 mapped markers comprised of 2364 5K-set markers and 1019 600K-set markers (Table 2). Filtration for population-specific monomorphic markers led to the total reduction of 1400 and 85 mapped markers in the NRG&CMS and restorer population, respectively. While the linkage analysis had to be conducted for the individual 5K-set and 600K-set markers due to divergence in the arrays genetic distance measure, the two marker sets led to highly similar estimates of LD (Table 2). The NRG&CMS population exhibited a mean genome-wise LD of 0.394 and 0.364 for the 5K-set and 600K-set markers respectively with a mean of 0.385 across the entire marker panel. Restorer population exhibited a mean genome-wise LD of 0.150 and 0.204 for the 5K-set and 600K-set markers respectively with a mean of 0.166 across the entire marker panel.

Comparison of the intrachromosomal LD between populations led to the finding of a similar lower mean LD on the 4R chromosome, while higher level on the 3R chromosome (Fig 6, Table 2). The restorer population distinguished itself by a ≥15% relative lower and higher LD

Table 2. Mean pairwise intrachromosomal linkage disequilibrium in the Nordic Seed hybrid rye elite breeding component germplasm, comprising a seed mother (NRG&CMS, N = 123) and pollen father (R, N = 242) population using 3383 mapped SNP markers originating from the 5K array and 600K array. Prior to analysis quality control (QC) was done to remove population-specific monomorphic markers.

Chromosome			1R	2R	3R	4R	5R	6R	7R	Mean	Total
Chromosome length (cM)		5K	253	320	247	283	332	277	207	274	1919
		600K	153	167	166	205	209	170	139	172	1209
No. SNP Markers	Total	5K	297	402	313	291	418	302	341	378	2364
		600K	179	158	174	133	153	120	102	146	1019
QC SNP Markers	NRG&CMS	All	235	415	156	274	350	281	271	283	1982
		5K	151	293	108	196	264	199	209	203	1420
		600K	84	122	48	78	86	82	62	80	562
	R	All	476	559	485	423	561	418	441	480	3363
		5K	297	401	312	291	415	298	339	336	2353
		600K	179	158	173	132	146	120	102	144	1010
Linkage disequilibrium + SD	NRG&CMS	All	0.322+0.24	0.367+0.25	0.424+0.29	0.319+0.20	0.389+0.26	0.300+0.19	0.420+0.25	0.385+0.26	-
		5K	0.396+0.30	0.375+0.24	0.546+0.36	0.328+0.21	0.422+0.27	0.326+0.19	0.456+0.25	0.394+0.26	-
		600K	0.322+0.23	0.388+0.26	0.424+0.37	0.385+0.23	0.441+0.28	0.295+0.19	0.391+0.24	0.364+0.26	-
	R	All	0.151+0.17	0.134+0.13	0.191+0.19	0.131+0.13	0.153+0.15	0.190+0.20	0.162+0.14	0.166+0.16	-
		5K	0.146+0.15	0.122+0.11	0.152+0.13	0.129+0.12	0.163+0.15	0.181+0.18	0.167+0.14	0.150+0.14	-
		600K	0.194+0.20	0.175+0.17	0.295+0.25	0.151+0.14	0.184+0.19	0.267+0.27	0.169+0.13	0.204+0.20	-

https://doi.org/10.1371/journal.pone.0239541.t002

on the 2R and 6R chromosome, respectively. The NRG&CMS population was found to exhibit a ≥15% relative lower LD on the 1R and 6R chromosome, while a ≥15% relative higher on the 7R chromosome.

To identify structural differences and patterns between populations, the pairwise intra-chromosomal LD was visualized by a heatmap (S1 Fig). Due to the non-homogeneous distribution of markers, extrapolation of intra-chromosomal distance is not feasible on the heatmap, instead depicting the immediate pairwise relationships. Heatmaps constructed on basis of separate arrays were found to produce comparable intra-chromosomal patterns revealing large LD blocks in the NRG&CMS population with an excess of LD in the peri-centromeric region (Fig 7A, 7C and 7E). The restorer population portrayed a more homogeneously distributed LD

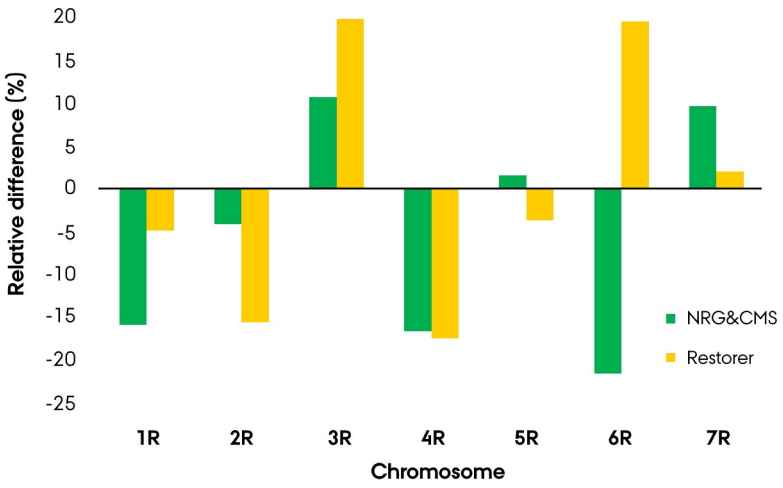


Fig 6. Relative difference in intrachromosomal linkage disequilibrium (%) compared to the genome-wise mean for the seed mother (NRG&CMS) or pollen father (R) population in the Nordic Seed elite hybrid rye breeding germplasm using 3383 SNP markers.

https://doi.org/10.1371/journal.pone.0239541.g006

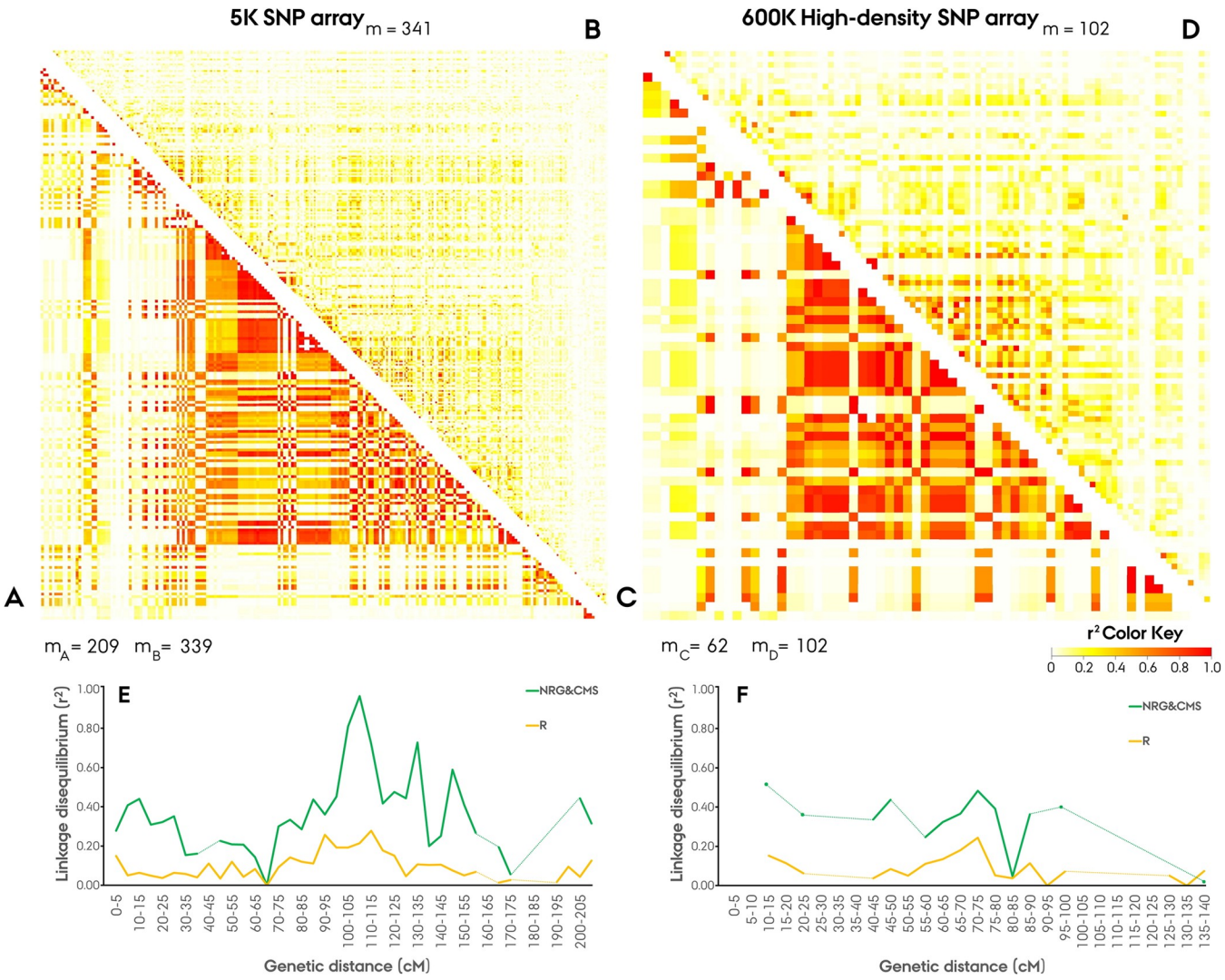


Fig 7. Pairwise linkage disequilibrium (LD, r^2) across the 7R rye chromosome in the Nordic Seed hybrid rye elite breeding populations using population-specific polymorphic SNP markers (m), originating from a 5K (**A,B,E**) and 600K array (**C,D,F**), visualized by heatmap for the, **A/C**) Seed mother (NRG&CMS), **B/D**) Pollen father (R) population, with the genome-wise distribution of LD across arrays visualized in bins of 5 cM. Missing bins is represented by dotted lines.

https://doi.org/10.1371/journal.pone.0239541.g007

with a less distinct pattern of excess LD in the peri-centromeric region and fewer minor LD blocks (Fig 7B, 7D, 7E and 7F). In both populations, the 3R chromosome demonstrated the highest mean LD with this excess readily event in the NRG&CMS population portraying a large LD block spanning from 76 to 108 cM in the 5K array and 76 to 99 cM in the 600K array (Table 2, S1 Fig). In the distal region of the 4RL chromosome, a distinct LD block was present in both populations (S1 Fig). In the NRG&CMS population, the block spanned 12 markers from 184 to 189 cM on the 600K array exhibited a mean pairwise LD of 0.91 and mean marker F_{ST} of 0.54. In the restorer population, the block comprised 5 markers at 189 cM on the 600K array, exhibiting a mean pairwise LD of 0.80 and mean F_{ST} of 0.62. The two populations diverged singularly by more than 40% in relative population-wise mean LD on the 6R chromosome with the increased LD in the restorer population organized in multiple moderate blocks (Fig 6, S1 Fig).

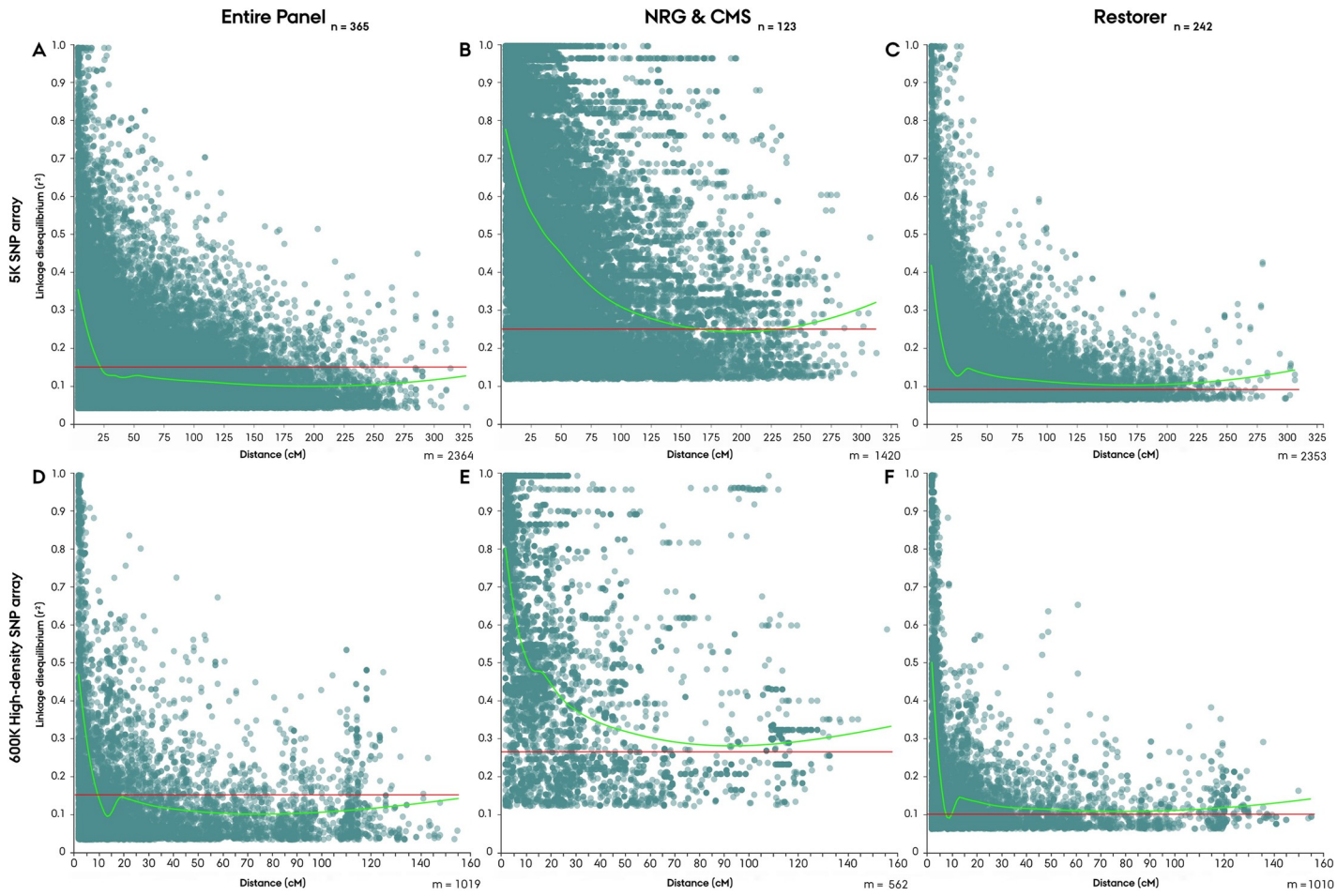


Fig 8. Genome-wise linkage decay in the Nordic Seed hybrid rye elite breeding germplasm (**A,D**), seed mother (NRG&CMS) (**B,E**), and, pollen father (Restorer) (**C,F**) population using population specific polymorphic SNP markers from a 5K (**A-C**) and 600K (**D-F**) array. LOESS smooth fitted curve (green) and baseline of population-specific critical value of r^2 (red).

https://doi.org/10.1371/journal.pone.0239541.g008

The rate of LD decay, determined for the individual arrays, were found to be moderately comparable, with the point of linkage decay determined at 23 cM and 9.5 cM for the 5K-set and 600K-set markers, respectively (Fig 8A and 8D). The rate of decay, however, diverged considerably between the parental populations. In the restorer population, linkage decay could only be determined in the 600K array at 5.5 cM (Fig 8C and 8F), while in the NRG&CMS population it could only be determined in the 5K array at 160 cM (Fig 8B and 8E).

Discussion

Recent years technological and scientific advances have progressively expedited the introduction of genomic-based techniques in modern plant breeding facilitating an unprecedented insight into the genetic features shaping breeding germplasms [46]. Comprehensive understanding of these features is a necessity to efficiently exploit and continuously advance the genetic potential residing in breeding germplasm to ensure genetic gains and trait-diversity to meet future demands [47]. Currently, DArT, SSR and SNP molecular marker systems have been validated for population studies in the *Secale* genus, OPVs, landraces, hybrid cultivars and small scale inbred populations [1, 25, 48]. In this investigation, we present the first

comprehensive population study of an elite germplasm using the Gülzow-type fertilization control system for hybrid breeding comprising 376 inbred component lines based on 4419 SNP markers.

Verifying genetic separation of parental populations

Existence of heterotic parental gene pools constitutes the cornerstone in hybrid breeding programs as the prerequisite for achieving a high heterosis effect in hybrid crosses. Our primary objective was, therefore, to confirm and quantify the genetic separation of parental populations. The immediate necessity of a marker-assisted assessment was further emphasized due to partial pedigree information on the NRG&CMS population.

Initial visualization of the germplasm's genetic architecture through PCA and phylogenetic analysis provided an early indication of a distinct genetic separation (Figs 3 and 4). The admixture model of inferred ancestry likewise separated the two parental populations suggesting the existence of two genetically distinct ancestral founder gene pools with limited degree of co-ancestry (Fig 5B and 5G). This infers that the NRG&CMS population is derived from a gene pool genetically distinct to that of the 'Petkus' and 'Carsten'. Estimation of the fixation indices (F_{ST} = 0.332) quantification the genetic separation providing a measure for comparative interpretation. In an extensive population study by Van Inghelandt *et al.* [49] on 1537 inbred hybrid maize elite breeding lines belonging to four heterotic pools, the F_{ST} ranged from 0.06 to 0.29. Bauer *et al.* [1] conducted a similar small-scale population study in an elite hybrid rye breeding germplasm to validate the 600K high-density SNP array. In their study, they confirmed a 'strong' differentiation between parental populations exhibiting an F_{ST} value of 0.229. In conjunction, our study suggests a comparatively high genetic differentiation between the parental populations in the assayed germplasm validating a heterotic pattern suitable for hybrid development. Furthermore, in maize, the relationship between panmictic midparent heterosis, specific combining ability and yielding characteristics has been reported to be positively correlated with the interparental genetic distance [50, 51]. Genetic distance can hence be readily implemented as a simple predictor of hybrid performance in parental crosses prior to development and implementation of more accurate genomic selection procedures [18, 52]. We have confirmed similar positive correlations between interparental genetic distance and average yield performance of hybrids in the germplasm (P. Sarup, personal communication). However, Frisch *et al.* [53] has demonstrated that transcriptome-based distance measure yields more accurate prediction models of hybrid performance.

Genetic diversity & architectural complexity

Hybrid breeding in rye relies on a three-way cross system involving a genetically distinct pollen parent and a seed parent NRG&CMS. Introgression of foreign released OPVs is a common source for introducing novel genetic variance in otherwise secluded rye breeding germplasm. In practical terms, this involves three-stages, (1) recurrent self-fertilization of OPVs to stabilize their genetic profile, (2) field- and/or greenhouse trials for phenotyping of agronomically important traits, and, (3) evaluation of the genetic compatibility of selected OPVs for introgression in either of the parental gene pools. Restoration of the G-type male sterility (MS) system employed in the CMS population is commonly observed by foreign OPVs, whereas the capacity to maintain the G-type MS, *i.e.* NRG lines, are less frequent [54]. Majority of foreign OPVs introduced to the germplasm has therefore consequently been restorers, hence contributing to an uneven broadening of the germplasms genetic base disfavoring the NRG&CMS population.

As a practical derivative of the three-way cross hybrid system, efforts have primarily been centered around the restorer population relative to the two-component NRG&CMS

population. It is reasonable to presume that these circumstances in conjunction have left a considerable mark on the comparative difference in genetic diversity and architectural complexity between the parental populations. With a sound understanding of the practical breeding history, our study set out to investigate and quantify the derived effect on the genetic features of the elite hybrid rye breeding germplasm. Initial assessment of the breeding lines genetic characteristics led to the finding of a similar low level of residual heterozygosity, high inbreeding coefficient and proportionate effective population size (N_e/N) in the two parental populations suggesting exceedingly fixed inbred individuals. While portraying a similar proportionate effective population size it is important to notice the considerable size difference in the actual effective population (N_e) with only 76 individuals relative to 137 in the NRG&CMS and restorer population, respectively. Owing to a small effective population size the NRG&CMS population is hence exposed to a strong potential effect of genetic drift inferred to facilitate a loss of genetic diversity over time by random fixation of neutral alleles [55].

Estimation of the within population gene diversity term (H_s) confirmed a relative higher level of genetic diversity in the restorer (0.333) relative to the NRG&CMS population (0.250) as expected from the practical breeding history. This divergence in relative genetic diversity could readily be observed in the PCA and phylogenetic analysis providing visualization for a more intuitive interpretation of the estimated genetic characteristics (Figs 3 and 4). In the population study by Bauer *et al*. [1], they identified a similar level of genetic diversity (0.311) within a smaller inbred restorer population, while contrary to our observations a comparatively higher level (0.327) within the NRG&CMS population. In hybrid rye breeding germplasms relying on the more commonly employed 'Pampa' (P) type MS system, introgressed OPVs are contrary to the G type less frequently capable of restoring fertility [14, 56]. As a result, OPVs often exerting the capacity to maintain the P type MS hence contribute to the broadening of the NRG&CMS populations genetic profile. This emphasizes the direct implications of the employed MS system on the introgression of foreign material for broadening the genetic base and the derived effects on the genetic diversity and architectural structure on parental gene pools in elite hybrid rye breeding germplasms. In comparison to elite hybrid breeding germplasm of related cereal species, a similar level of genetic diversity have been reported in maize populations (0.256–0.355) [57] and rice accessions (0.290) [58]. In conjunction, these results indicated a population-wise discrepancy in the genetic diversity defined by a moderate-to-high level of genetic diversity in the restorer population relative to the more genetically narrow NRG&CMS population.

Effects of the practical breeding history could likewise be observed in the parental population's structural architecture. Comparatively, fewer lines could be assigned to secluded subpopulations of a single ancestral component in the restorer population suggesting a more structurally complex admixture which could be explained by the continuous introgression of foreign restorers (Table 1). Interestingly, the Evanno's DeltaK method indicated an additional restorer founder gene pool distinct to the 'Carsten' and 'Petkus' (Fig 5I). Based on available pedigree information, we propose that the more secluded SP2 component is likely derived from the novel ancestral gene pool introduced through introgressed restorers (Table 2, Fig 5I, Teal). The Evanno's DeltaK method, furthermore, suggested that the NRG&CMS population is derived from three, however, unknown ancestral founder populations, portraying a more distinct subpopulation structure (Fig 5D).

Linkage disequilibrium & decay

Existence of non-random association of alleles at different loci, *i.e*. linkage disequilibrium (LD) is a prerequisite of association genetics facilitating the statistical establishment of

causality between marker and a trait of interest [59, 60]. To determine the population-wise LD, the identified 3383 mapped markers were filtered for population-specific monomorphism causing a considerable discrepancy in marker density between the restorer (3298) and NRG&CMS (1983) population (Table 2). The estimation of genome-wide LD led to the discovery of a large divergence in LD between the parental populations exhibiting a mean of 0.381 and 0.166 in the NRG&CMS and restorer population, respectively. Visualization of the intra-chromosomal pairwise LD by heatmap revealed that the excess of LD in the NRG&CMS population was centered around larger LD blocks spanning beyond the low-recombining peri-centromeric region (Fig 7A, 7C and 7E) [61]. This region is shaped by an abundance of repetitive transposable elements and low-recombination frequency, rendering genes largely inaccessible to breeders [61, 62]. In barley, the peri-centromeric region constitutes 48% of the genome harboring an estimated 14–22% of the gene content [63]. Effect of low-recombination frequency could readily be observed in the 5K-set markers with *e.g.* mean pairwise LD in the NRG&CMS population reaching as high as 0.96 in the 105–110 cM bin on the 7R chromosome (Fig 7E). LD blocks, however, residing outside of the apparent centromeric region provides evidence of chromosomal segments targeted by selection in the germplasm [60]. Conserved blocks in both parental populations suggest a common selection scheme of either an agronomic important trait *e.g.* disease resistance or a domestication-related trait *e.g.* grain shattering [64, 65]. Comparative analysis of the population-wise LD patterns led to the discovery of a common excess on the 3R chromosome in both populations (Fig 6). On the short-arm of 3R chromosome we identified a vast LD block spanning in the NRG&CMS population from 76 to 102 cM on the 5K array (S1 Fig). Several agronomic important QTLs controlling genes affecting preharvest sprouting and drought tolerance have been reported on the 3R chromosome in rye [66, 67]. Melz and Adolf [28], furthermore identified a male sterility (ms) minor gene *ms2* on the 3RL chromosome, later denoted as 'restorer of fertility' (*Rf*) 'Gülzow' type *Rfg2* which could likewise be related to the conservation of this segment in the germplasm [28, 68]. While functional analysis in rye is impeded by poor or insufficient gene annotation, certain agronomic traits *e.g.* male-fertility restoration is, however, well annotated, with the major *Rf* genes of the 'Pampa' type, *Rfp1*, *Rfp2*, and, *Rfp3* accurately mapped to the 4RL chromosome [69, 70]. Intriguingly we identified a strongly conserved LD block in both parental populations at the distal region of the 4RL chromosome, with 9/12 markers in the restorer population and 3/5 markers in the NRG&CMS population annotated as '*Rfp3*'. The 'Gülzow' *Rfg1* and 'C' type *Rfc1 Rf* major genes have likewise been mapped to the distal region of the 4RL chromosome and been proposed allelic to the P-type [68, 71]. The novel rye reference genome by Rabanus-Wallace *et al.* [21], however, constitutes a milestone in large-scale functional analysis, *de novo* annotating the near-full complement of 34.441 gene models in the rye genome.

As an outcrossing crop species rye has been reported to exhibit a rapid rate of LD decay [72]. In our analysis the point of LD decay set as the interception point of the LOESS regression curve and the population-specific critical value of r^2 could, however, not be determined for the NRG&CMS population (Fig 8B and 8E). This finding is likely related to the substantial intrachromosomal LD structure observed within the population, impeding the standard decay of linkage as a function of the inter-loci distance [73]. The high LD within the NRG&CMS population putatively derives from a low frequency of detectable recombination events caused by a depletion of rare alleles within the NRG&CMS population with ~50% of the polymorphic markers exhibiting a MAF $\leq$ 0.1 (Fig 1A) [74]. In aggregate, these findings suggest the occurrence of a demographic bottleneck event or intense selection in the NRG&CMS population supported by the observations of a lower genetic diversity, structural complexity and effective population size [75]. It is possible that the low frequency of lines exhibiting the ability to maintain the employed G-type MS system (*i.e.* NRG lines) could have exerted a considerable

selection pressure with introgressed lines exhibiting a high degree of relatedness. In the restorer population, which portrayed a more homogeneous LD structure with minor LD blocks, linkage decay could be determined in the 600K SNP array at 5.5 cM (Figs 7B, 7D and 8F). By comparison, in a study on an elite hybrid maize breeding germplasm by Van Inghelandt *et al.* [76], LD decay was observed in the range from 0.11 to 2.74 cM across four heterotic gene pools. The comparatively rapid rate of LD decay in outcrossing species such as rye promises high resolution for association studies while requiring an equivalently higher marker density [72].

Introgression of novel genetic variation from foreign sources

Genetic diversity constitutes the fabric of which novel cultivars are shaped, portraying a unique combination of alleles superior to the desired purpose. Erosion of genetic diversity, therefore, constitutes a serious threat to modern plant breeding with loss of features imperative to meet the future demands [77]. Findings of our study emphasizes foremost the necessity of addressing the strong population structure and narrow genetic profile of the NRG&CMS population to sustain the genetic gains in the hybrid rye breeding program. The comparatively small effective population size of both parental populations furthermore emphasizes the need for broadening the germplasm, particularly in the NRG&CMS population. In rye, several population genetics studies have recently investigated the accessible genetic reservoir residing in landraces, OPVs, and wild *Secale* species for this immediate purpose [3, 26, 77]. Introgression of landraces and OPVs in elite hybrid breeding germplasms require a comprehensive evaluation of their genetic compatibility in order to preserve the genetic separation of parental gene pools [78]. Marker assisted introgression of OPVs has successfully been demonstrated by Fischer *et al.* [16] in a ´Petkus´ and ´Carsten´ based elite hybrid rye breeding germplasm. Introgression libraries of an exotic Iranian landrace has likewise been developed by Falke *et al.* [79], and demonstrated by marker-assisted backcrossing to significantly enhance baking quality traits relative to the recurrent parent [80].

Performance and informativeness of SNP markers

Stringent assessment of marker properties and informativeness constitutes an imperative step in the successful application of genomic-based breeding techniques in plant breeding. Analysis of polymorphic information content (PIC) is often used as a measure of the informativeness of a genetic marker for linkage studies but has likewise been deployed as a genetic diversity term [81]. Across marker-sets, the 5K-set and 600K-set were highly similar in terms of mean marker PIC (0.268, 0.262) and mean MAF (0.254, 0.250) score. The smaller wheat-set markers (426), however, exhibited a slightly lower PIC (0.225) and MAF (0.201) score. In our study, we determined a mean PIC value of 0.20 and 0.26 across the population-specific polymorphic markers in the NRG&CMS and restorer population, respectively (Fig 1B). These values correspond to PIC values reported in SNP-based population studies on elite hybrid breeding germplasms in maize (0.29) [57] and rice (0.23) [58]. Amongst the 4419 markers, we discovered a considerable divergence in population-specific polymorphic markers with 1870 and 38 monomorphic markers in the NRG&CMS and restorer population, respectively. Initially this divergence was interpreted as a bias introduced in the quality filtration step as a product of the population size difference. Population-wise quality filtration on the entire marker panel, however, only led to the identification of an additional 271 low-polymorphism markers for the NRG&CMS population. This can either be explained by the narrow genetic profile of the NRG&CMS population characterized by an depletion of rare alleles due to higher genetic drift (Fig 1A) and/or by a diverging pattern in SNPs relative to the parental lines utilized for the 5K and 600K SNP array

design more closely resembling the restorer population [1, 20]. If the latter explanation is correct this further supports the distinct genetic separation of the parental populations, characterized by a unique population-dependent pattern in SNPs.

In our investigation of LD structure and rate of decay, the separation of markers into a 5K-set and 600K-set due to diverging array distance measure likewise enabled a comparative assessment of the marker sets performance and informativeness. It is important to emphasize that with an incomplete subset of markers from the 5K an 600K SNP array we cannot speculate on their comparative performance and informativeness. Neither of the assessed genetic characteristics including LD, intrachromosomal organization of LD, and, progression of linkage decay was found to consistently diverge between the marker sets in our analysis (Table 2, Figs 7 and 8). Genetic mapping of the SNPs likewise revealed a similar positional pattern in the arrays, displaying an evident excess of markers in the peri-centromeric region (Fig 2A and 2B). While the non-uniform intrachromosomal distribution of markers clustering in the peri-centromeric region, pose question to the informativeness of the SNP marker panel we observed no apparent evidence of a positional trend on the assessed genetic characteristics. Conclusively, our analysis found no evidence of divergence in performance nor informativeness between the 5K-set and 600K-set markers, hence endorsing the 20K SNP chip construct.

Conclusion

In the present study, we demonstrated the application of a SNP-based marker system for dissecting the genetics of a large elite hybrid rye breeding germplasm. Through a palette of complimenting analysis, we confirmed a strong genetic differentiation of parental populations. These populations were found to diverge in several features with the NRG&CMS population portraying a strong population structure characterized by a narrow genetic profile, small effective population size and high genome-wise LD. We propose that employed MS system putatively constitutes a population determining parameter by influencing the rate of introducing novel genetic variation. Functional analysis of LD blocks for inference of selection on agronomic important traits in the germplasm led to the finding of a conserved segment on the distal 4RL chromosomal region annotated to the *Rfp3* male-fertility restoration gene.

Considering the plethora of diversity preserved in genetic resources it is the emphasis of this study to pursue a marker-assisted broadening of the NRG&CMS population to address the strong population structure and narrow genetic profile. Furthermore, the novel rye reference genome by Rabanus-Wallace *et al*. [21] facilitates physical mapping and annotation of the 5K-set and 600K-set markers for combined linkage study and large-scale functional analysis to further unravel the genetic features of the hybrid rye elite breeding germplasm.

Supporting information

S1 Table. Opposing population ancestry from an admixture model of inferred ancestry at K set to 2 and observed residual heterozygosity (Ho) of 11 discarded Nordic Seed hybrid rye elite breeding lines belonging to the restorer (R) and non-restorer germplasm (NRG) population.
(DOCX)

S1 Fig. Pairwise intra-chromosomal linkage disequilibrium (LD, r2) in the Nordic Seed hybrid rye elite breeding populations using population-specific polymorphic SNP markers originating from a 5K (mNRG&CMS = 1420, mR = 2353) and 600K (mNRG&CMS = 562, mR = 1010) array visualized by heatmap for the seed mother (NRG&CMS) and pollen

father (R) population.
(TIF)

S1 File.
(ZIP)

Acknowledgments

We would like to thank laboratory technician Hanne Svenstrup at Nordic Seed A/S for contribution to the genotypic data collection. The plant material included in this study was provided by Nordic Seed Germany GmbH (Nienstädt, Germany). The high throughput SNP genotyping was performed by Trait Genetics (Gatersleben, Germany).

Author Contributions

Conceptualization: Nikolaj M. Vendelbo, Jihad Orabi, Ahmed Jahoor.

Data curation: Nikolaj M. Vendelbo, Pernille Sarup.

Formal analysis: Nikolaj M. Vendelbo, Pernille Sarup, Jihad Orabi.

Funding acquisition: Jihad Orabi, Ahmed Jahoor.

Investigation: Nikolaj M. Vendelbo, Peter S. Kristensen, Ahmed Jahoor.

Methodology: Nikolaj M. Vendelbo, Ahmed Jahoor.

Project administration: Jihad Orabi, Ahmed Jahoor.

Resources: Jihad Orabi, Peter S. Kristensen.

Software: Nikolaj M. Vendelbo, Pernille Sarup.

Supervision: Pernille Sarup, Jihad Orabi, Ahmed Jahoor.

Validation: Pernille Sarup, Jihad Orabi, Peter S. Kristensen, Ahmed Jahoor.

Visualization: Nikolaj M. Vendelbo.

Writing – original draft: Nikolaj M. Vendelbo.

Writing – review & editing: Nikolaj M. Vendelbo, Pernille Sarup, Jihad Orabi, Peter S. Kristensen, Ahmed Jahoor.

References

1. Bauer E, Schmutzer T, Barilar I, Mascher M, Gundlach H, Martis MM, et al. Towards a whole-genome sequence for rye (Secale cereale L.). The Plant Journal. 2017; 89(5): 853–69. https://doi.org/10.1111/tpj.13436 PMID: 27888547
2. Frederiksen S, Petersen G. A taxonomic revision of Secale (Triticea, Poaceae). Nordic Journal of Botany. 1998; 18(4): 399–420.
3. Maraci O, Ozkan H, Bilgin R. Phylogeny and genetic structure in the genus Secale. PLoS One. 2018; 13(7): 1–21. https://doi.org/10.1371/journal.pone.0200825 PMID: 30024916
4. Schreiber M, Himmelbach A, Börner A, Mascher M. Genetic diversity and relationship between domesticated rye and its wild relatives as revealed through genotyping-by-sequencing. Evolutionary Applications. 2018. https://doi.org/10.1111/eva.12624 PMID: 30622636
5. Hillman G. On the origins of domestic rye—secale cereale: the finds from aceramic Can Hasan III in Turkey. Anatolian Studies. 1978; 28: 157–74. https://doi.org/10.2307/3642748
6. FAO. Food and Agriculture Organization of the United Nations. FAOSTAT statistical database 2018 [cited 2020 25/2]. Available from: http://www.fao.org/faostat.

7. Schlegel R. Hybrid breeding boosted molecular genetics in rye. Russian Journal of Genetics: Applied Research. 2016; 6(5): 569–83. https://doi.org/10.1134/s2079059716050105

8. Geiger HH, Miedaner T. Rye breeding. In: Carena MJ, editor. Cereals (Handbook of Plant Breeding). 3: Vol 3. Springer Science & Business Media, New York, pp 157–181; 2009. p. 157–81.

9. Lundqvist A. Self-incompatability in rye. I. Genetic control in the diploid. Hereditas. 1956; 42: 293–348.

10. Husband BC, Schemske DW. Evolution of the magnitude and timing of inbreeding depression in plants. Evolution. 1996; 50(1): 54–70. https://doi.org/10.1111/j.1558-5646.1996.tb04472.x PMID: 28568860

11. Geiger HH, Miedaner T. Hybrid rye and heterosis. Genetics and Exploitation of Heterosis in crops. 1999: 439–50.

12. Beck DL, Vasal SK, Crossa J. Heterosis and combining ability of cimmyts tropical Heterosis and combining ability of CIMMYT's tropical early and intermediate maturity maize (Zea mays L.) germplasm. Maydica. 1990; 35(3): 279–85.

13. Crossa J, Gardner CO, Mumm RF. Heterosis among populations of maize (Zea mays L.) with different levels of exotic germplasm. Theoretical and Applied Genetics. 1986; 73(3): 445–50.

14. Geiger HH, Schnell FW. Cytoplasmic Male Sterility in Rye (Secale cereale L.). Crop Science. 1970; 10 (5): 590–3.

15. Schlegel RHJ. Rye Genetics, Breeding, and Cultivation: CRC Press; 2013.

16. Fischer S, Melchinger AE, Korzun V, Wilde P, Schmiedchen B, Mohring J, et al. Molecular marker assisted broadening of the Central European heterotic groups in rye with Eastern European germplasm. Theoretical and Applied Genetics. 2010; 120(2): 291–9. https://doi.org/10.1007/s00122-009-1124-0 PMID: 19669632

17. Schulthess AW, Wang Y, Miedaner T, Wilde P, Reif JC, Zhao Y. Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. Theoretical and Applied Genetics. 2016; 129(2): 273–87. https://doi.org/10.1007/s00122-015-2626-6 PMID: 26561306

18. Wang Y, Mette MF, Miedaner T, Gottwald M, Wilde P, Reif JC, et al. The accuracy of prediction of genomic selection in elite hybrid rye populations surpasses the accuracy of marker-assisted selection and is equally augmented by multiple field evaluation locations and test years. BMC Genomics. 2014; 15(1): 556.

19. Heffner EL, Lorenz AJ, Jannink J-L, Sorrells ME. Plant breeding with genomic selection: gain per unit time and cost. Crop Science. 2010; 50(5): 1681–90. https://doi.org/10.2135/cropsci2009.11.0662

20. Haseneyer G, Schmutzer T, Seidel M, Zhou R, Mascher M, Schon CC, et al. From RNA-seq to large-scale genotyping—genomics resources for rye (Secale cereale L.). BMC Plant Biol. 2011; 11(1): 131. https://doi.org/10.1186/1471-2229-11-131 PMID: 21951788

21. Rabanus-Wallace MT, Hackauf B, Mascher M, Lux T, Wicker T, Gundlach H, et al. Chromosome-scale genome assemly provides insight into rye biology, evolution, and agronomic potential. BioRxiv. 2019. https://doi.org/10.1101/2019.12.11.869693

22. Rafalski A. Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology. 2002; 5(2): 94–100. https://doi.org/10.1016/s1369-5266(02)00240-6 PMID: 11856602

23. Gur A, Zamir D. Unused natural variation can lift yield barriers in plant breeding. PLoS Biol. 2004; 2 (10): e245. https://doi.org/10.1371/journal.pbio.0020245 PMID: 15328532

24. Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Are AK, Ortiz R. Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. Trends in Plant Science. 2016; 21(1): 31–42. https://doi.org/10.1016/j.tplants.2015.10.012 PMID: 26559599

25. Targonska M, Bolibok-Bragoszewska H, Rakoczy-Trojanowska M. Assessment of Genetic Diversity in Secale cereale Based on SSR Markers. Plant Molecular Biology Reporter. 2016; 34(1): 37–51. https://doi.org/10.1007/s11105-015-0896-4 PMID: 26843779

26. Bolibok-Bragoszewska H, Targonska M, Bolibok L, Kilian A, Rakoczy-Trojanowska M. Genome-wide characterization of genetic diversity and population structure in secale. BMC Plant Biology. 2014; 14 (1): 184.

27. Persson K, Diaz O, Von Bothmer R. Extent and patterns of RAPD variation in landraces and cultivars of rye (Secale cereale L.) from northern Europe. Hereditas. 2001; 134(3): 237–43. https://doi.org/10.1111/j.1601-5223.2001.00237.x PMID: 11833287

28. Melz G, Adolf K. Genetic analysis of rye (Secale cereale L.) Genetics of male sterility of the G-type. Theoretical and Applied Genetics. 1991; 82(6): 761–4. https://doi.org/10.1007/BF00227322 PMID: 24213452

29. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, et al. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnol Journal. 2014; 12(6): 787–96. https://doi.org/10.1111/pbi.12183 PMID: 24646323

30. Rstudio Team Rstudio: Integrated Development for R. RStudio, Inc., Boston. http://www.rstudio.com 2015.

31. R Core Team. R a language and environment for statistical computing. R foundation for statistical Computing, Vienna, Austria. https://www.R-project.org/. 2020.

32. Granato ISC, Galli G, de Oliveira Couto EG, e Souza MB, Mendonça LF, Fritsche-Neto R. snpReady: a tool to assist breeders in genomic analysis. Molecular Breeding. 2018; 38(8): 102. https://doi.org/10.1007/s11032-018-0844-8

33. Goudet J. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes. 2005; 5(1): 184–6. https://doi.org/10.1111/j.1471-8278

34. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984; 38(6): 1358–70. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x PMID: 28563791

35. Paradis E. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics. 2010; 26(3): 419–20. https://doi.org/10.1093/bioinformatics/btp696 PMID: 20080509

36. Murdoc D, Adler D, others. rgl:3D Visualization Using OpenGL. R package version 010030. 2019.

37. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155(2): 945–59. PMID: 10835412

38. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology. 2005; 14(8): 2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x PMID: 15969739

39. Earl DA. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 2012; 4(2): 359–61.

40. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019; 35(3): 526–8. https://doi.org/10.1093/bioinformatics/bty633 PMID: 30016406

41. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research. 2019; 47(W1): 256–9. https://doi.org/10.1093/nar/gkz239 PMID: 30931475

42. Clayton D. snpStats: SnpMatrix and XSnpMatrix classes and methods. R package version 1360. 2019.

43. Shin JH, Blay S, McNey B, Graham J. LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. Journal of Statistical Software. 2006; 16 (3): 1–10. https://doi.org/10.18637/jss.v000.i00

44. Covarrubias-Pazaran G. Genome-Assisted Prediction of Quantitative Traits Using the R Package sommer. PLoS One. 2016; 11(6): e0156744. https://doi.org/10.1371/journal.pone.0156744 PMID: 27271781

45. Breseghello F, Sorrells ME. Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. Genetics. 2006; 172(2): 1165–77. https://doi.org/10.1534/genetics.105.044586 PMID: 16079235

46. Rasheed A, Hao Y, Xia X, Khan A, Xu Y, Varshney RK, et al. Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives. Molecular Plant. 2017; 10(8): 1047–64. https://doi.org/10.1016/j.molp.2017.06.008 PMID: 28669791

47. Bevan MW, Uauy C, Wulff BB, Zhou J, Krasileva K, Clark MD. Genomic innovation for crop improvement. Nature. 2017; 543(7645): 346–54. https://doi.org/10.1038/nature22011 PMID: 28300107

48. Bolibok-Bragoszewska H, Heller-Uszynska K, Wenzl P, Uszynski G, Kilian A, Rakoczy-Trojanowska M. DArT markers for the rye genome—genetic diversity and mapping. BMC Genomics. 2009; 10(1): 578. https://doi.org/10.1186/1471-2164-10-578 PMID: 19958552

49. Van Inghelandt D, Melchinger AE, Lebreton C, Stich B. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theoretical and Applied Genetics. 2010; 120(7): 1289–99. https://doi.org/10.1007/s00122-009-1256-2 PMID: 20063144

50. Reif JC, Melchinger AE, Xia XC, Warburton ML, Holsington DA, Vasal SK, M. F. Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. Crop Science. 2003; 43(4): 1275–82.

51. Balestre M, Machado JC, Lima JL, Souza JC, Nóbrega-Filho L. Genetic distance estimates among single cross hybrids and correlation with specific combining ability and yield in corn double cross hybrids. Genetic and Molecular Research. 2008; 7(1): 65–73.

52. Auinger HJ, Schonleben M, Lehermeier C, Schmidt M, Korzun V, Geiger HH, et al. Model training across multiple breeding cycles significantly improves genomic prediction accuracy in rye (Secale cereale L.). Theoretical and Applied Genetics. 2016; 129(11): 2043–53. https://doi.org/10.1007/s00122-016-2756-5 PMID: 27480157

53. Frisch M, Thiemann A, Fu J, Schrag TA, Scholten S, Melchinger AE. Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. Theoretical and Applied Genetics. 2010; 120(2): 441–50. https://doi.org/10.1007/s00122-009-1204-1 PMID: 19911157

54. Melz G, Melz G, Hartman F. Genetics of a male-sterile rye of 'G-type' with results of the first F1-hybrids. Plant Breeding and Seed Science. 2003; 47: 47–55.

55. Fu YB. Understanding crop genetic diversity under modern plant breeding. Theoretical and Applied Genetics. 2015; 128(11): 2131–42. https://doi.org/10.1007/s00122-015-2585-y PMID: 26246331

56. Miedaner T, Glass C, Dreyer F, Wilde P, Wortmann H, Geiger HH. Mapping of genes for male-fertility restoration in 'Pampa' CMS winter rye (Secale cereale L.). Theoretical and Applied Genetics. 2000; 101(8): 1226–33.

57. Zhang X, Zhang H, Li L, Lan H, Ren Z, Liu D, et al. Characterizing the population structure and genetic diversity of maize breeding germplasm in Southwest China using genome-wide SNP markers. BMC Genomics. 2016; 17(1): 697. https://doi.org/10.1186/s12864-016-3041-3 PMID: 27581193

58. Zhi-zhou H, Fang-ming X, Li-yun C, Paz MAD. Genetic diversity of tropical hybrid rice germplasm measured by molecular markers. Rice Science. 2012; 19(3): 193–201.

59. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 2013; 9(1): 29.

60. Gupta PK, Rustgi S, Kulwal PL. Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Molecular Biology. 2005; 57(4): 461–85. https://doi.org/10.1007/s11103-005-0257-z PMID: 15821975

61. Baker K, Bayer M, Cook N, Dreissig S, Dhillon T, Russell J, et al. The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. The Plant Journal. 2014; 79(6): 981–92. https://doi.org/10.1111/tpj.12600 PMID: 24947331

62. Ma J, Wing RA, Bennetzen JL, Jackson SA. Plant centromere organization: a dynamic structure with conserved functions. Trends in Genetics. 2007; 23(3): 134–9. https://doi.org/10.1016/j.tig.2007.01.004 PMID: 17275131

63. International Barley Genome Sequencing C. A physical, genetic and functional sequence assembly of the barley genome. Nature. 2012; 491(7426): 711–6. https://doi.org/10.1038/nature11543 PMID: 23075845

64. Whitt SR, Tenaillon LM, Gaut BS, Buckler ES. Genetic diversity and selection in the maize starch pathway. Proceedings in the National Academy of Sciences. 2002; 99(20): 12959–62.

65. Li C, Zhou A, Sang T. Rice domestication by reducing shattering. Science. 2006; 311(5769): 1936–9. https://doi.org/10.1126/science.1123604 PMID: 16527928

66. Masojć P, Milczarski P. Relationship between QTLs for preharvest sprouting and alpha-amylase activity in rye grain. Molecular Breeding. 2008; 23(1): 75–84. https://doi.org/10.1007/s11032-008-9215-1

67. Mohammadi R, Farshadfar E, Aghaee-Sarbarzeh M, Sutka J. Locating QTLs controlling drought tolerance criteria in rye using disomic addition lines. Cereal Research Communications. 2003; 31(3–4).

68. Börner A, Korzun V, Polley A, Malyshev S, Melz G. Genetics and molecular mapping of a male fertility restoration locus (Rfg1) in rye (Secale cereale L.). Theoretical and Applied Genetics. 1998; 97(1–2): 99–102.

69. Stracke S, Schilling AG, Forster J, Weiss C, Glass C, Miedaner T, et al. Development of PCR-based markers linked to dominant genes for male-fertility restoration in Pampa CMS of rye (Secale cereale L.). Theoretical and Applied Genetics. 2003; 106(7): 1184–90. https://doi.org/10.1007/s00122-002-1153-4 PMID: 12748768

70. Hackauf B, Bauer E, Korzun V, Miedaner T. Fine mapping of the restorer gene Rfp3 from an Iranian primitive rye (Secale cereale L.). Theoretical and Applied Genetics. 2017; 130(6): 1179–89. https://doi.org/10.1007/s00122-017-2879-3 PMID: 28315925

71. Stojalowski S, Milczarski P, Hanek M, Tyrka M. The application of GBS markers for extending the dense genetic map of rye (Secale cereale L.) and the localization of the Rfc1 gene restoring male fertility in plants with the C source of sterility-inducing cytoplasm. Journal of Applied Genetics. 2016; 57(4): 439–51. https://doi.org/10.1007/s13353-016-0347-4 PMID: 27085345

72. Li Y, Haseneyer G, Schön C, Ankerst D, Korzun V, Wilde P, et al. High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (secale cereale L.) genes involved in frost response. BMC Plant Biology. 2011; 11(1): 6.

73. Flint-Garcia SA, Thornsberry JM, Buckler ESt. Structure of linkage disequilibrium in plants. Annual Review of Plant Biology. 2003; 54(1): 357–74. https://doi.org/10.1146/annurev.arplant.54.031902.134907 PMID: 14502995

74. Morell PL, Toleno DM, Lundy KE, Clegg MT. Low levels of linkage disequilibrium in wild barley (Hordeum vulgare spp. spontaneum) despite high rates of self-fertilization. Proceedings in the National Academy of Sciences. 2005; 102(7): 2442–7.

75. Gaut BS, Long AD. The lowdown on linkage disequilibrium. The Plant Cell. 2003; 15(7): 1502–6. https://doi.org/10.1105/tpc.150730 PMID: 12837942

76. Van Inghelandt D, Reif JC, Dhillon BS, Flament P, Melchinger AE. Extent and genome-wide distribution of linkage disequilibrium in commercial maize germplasm. Theoretical and Applied Genetics. 2011; 123(1): 11–20. https://doi.org/10.1007/s00122-011-1562-3 PMID: 21404061

77. Monteiro F, Vidigal P, Barros AB, Monteiro A, Oliveira HR, Viegas W. Genetic Distinctiveness of Rye In situ Accessions from Portugal Unveils a New Hotspot of Unexplored Genetic Resources. Frontiers in Plant Science. 2016; 7: 1334. https://doi.org/10.3389/fpls.2016.01334 PMID: 27630658

78. Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, et al. Use of SSRs for establishing heterotic groups in subtropical maize. Theoretical and Applied Genetics. 2003; 107(5): 947–57. https://doi.org/10.1007/s00122-003-1333-x PMID: 12830388

79. Falke KC, Susic Z, Hackauf B, Korzun V, Schondelmaier J, Wilde P, et al. Establishment of introgression libraries in hybrid rye (Secale cereale L.) from an Iranian primitive accession as a new tool for rye breeding and genomics. Theoretical and Applied Genetics. 2008; 117(4): 641–52. https://doi.org/10.1007/s00122-008-0808-1 PMID: 18535814

80. Falke KC, Susic Z, Wilde P, Wortmann H, Mohring J, Piepho HP, et al. Testcross performance of rye introgression lines developed by marker-assisted backcrossing using an Iranian accession as donor. Theoretical and Applied Genetics. 2009; 118(7): 1225–38. https://doi.org/10.1007/s00122-009-0976-7 PMID: 19238350

81. Guo X, Elston R. Linkage information content of polymorphic genetic markers. 1999.

Genetic structure of a germplasm for hybrid breeding in rye (*Secale cereale* L.)

Nikolaj M. Vendelbo, Pernille M. Sarup, Jihad Orabi, Peter S. Kristensen,
and Ahmed Jahoor

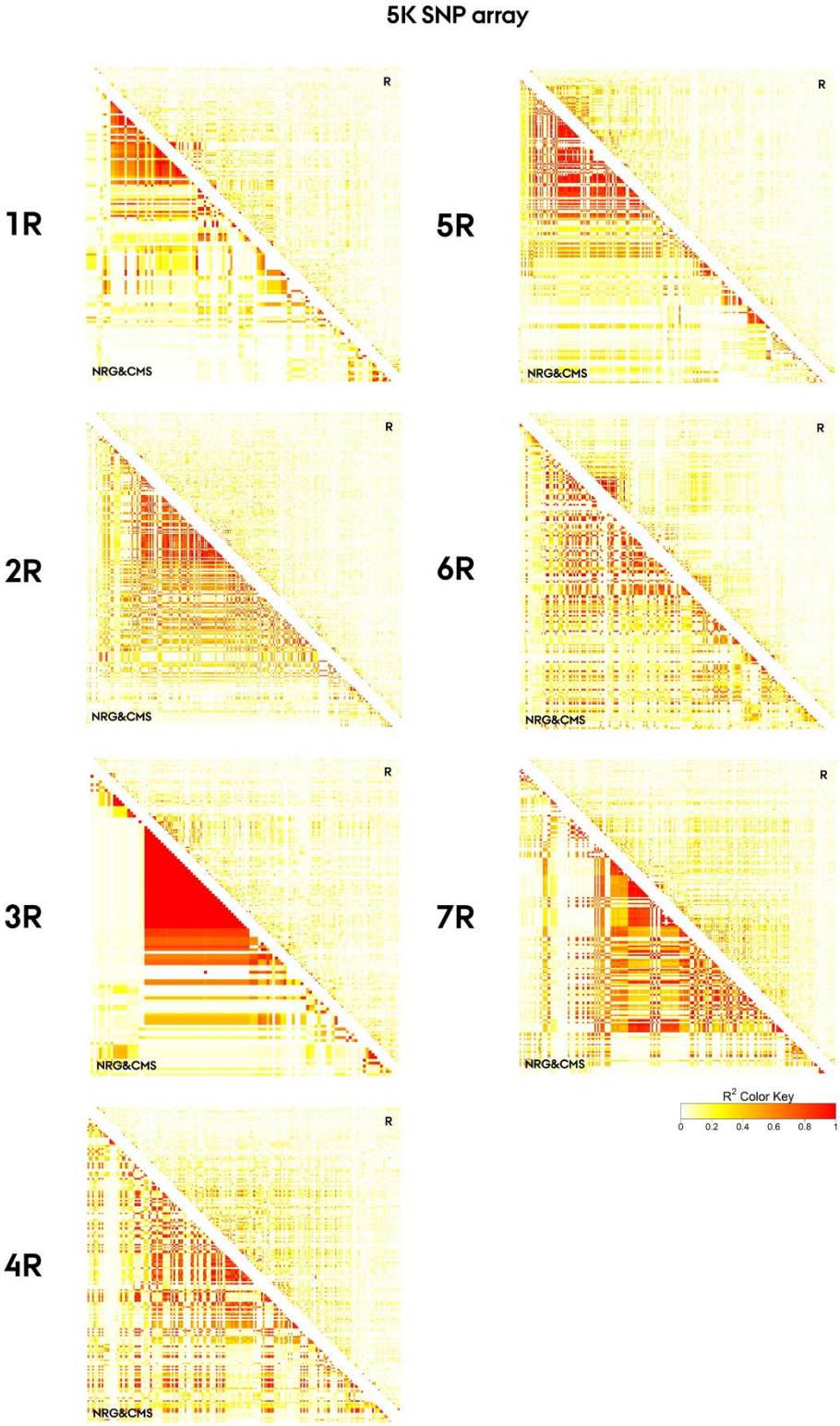
Supplementary material

Caption of all supplementary files has been given below and a subset of supplementary figures and tables included in the dissertation. The extended supplementary materials has been provided at an open-access Zenodo data repository (https://zenodo.org/record/5510272).

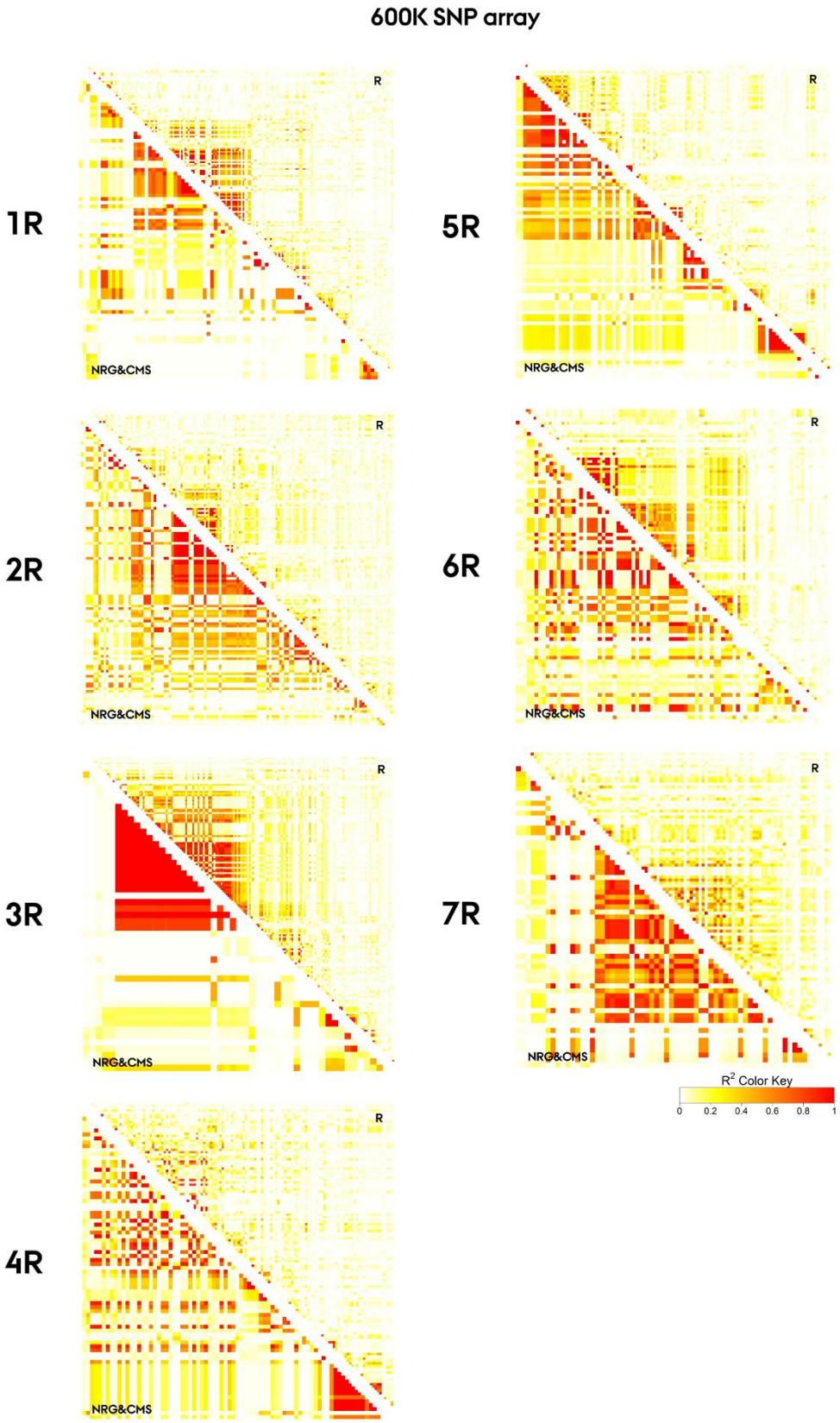
S1 Table. Opposing population ancestry from an admixture model of inferred ancestry at K set to 2 and observed residual heterozygosity (H_o) of 11 discarded Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines belonging to the restorer (R) and non-restorer germplasm (NRG) population.

Line ID	Opposing ancestry	H_o
R10	0.515	0.563
R115	-	0.463
R126	0.398	0.023
R127	0.419	0.004
R310	0.730	0.046
R311	0.497	0.015
R342	0.565	0.353
R343	0.588	0.358
NRG56	0.896	0.056
NRG79	0.410	0.048
NRG129	0.537	0.076

S1 Fig. Pairwise intra-chromosomal linkage disequilibrium (LD, r2) in the Nordic Seed hybrid rye (*Secale cereale* L.) breeding populations using population-specific polymorphic SNP markers originating from a 5K ($m_{NRG\&CMS}$ = 1420, m_R = 2353) and 600K ($m_{NRG}\&_{CMS}$ = 562, m_R = 1010) array visualized by heatmap for the seed mother (NRG&CMS) and pollen father (R) population



S1 Fig. (continued)



Manuscript II

Genomic can of male fertility restoration genes in a ‘Gülzow’ type hybrid breeding system of rye (*Secale cereale* L.)

By

Nikolaj M. Vendelbo*[ǂ1,2], Khalid Mahmood [1ǂ], Pernille M. Sarup [1], Peter S. Kristensen [1], Jihad Orabi [1], and Ahmed Jahoor [1,3]

[1] Nordic Seed A/S, Odder, Denmark

[2] Department of Agroecology, Aarhus University, Slagelse, Denmark

[3] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

[ǂ] These authors contributed equally to this work

Publication details

Published in International Journal of Molecular Science, vol. 22, issue 17, 9277, 2021

International Journal of
Molecular Sciences

MDPI

Article

Genomic Scan of Male Fertility Restoration Genes in a 'Gülzow' Type Hybrid Breeding System of Rye (*Secale cereale* L.)

Nikolaj Meisner Vendelbo [1,2,*,†], Khalid Mahmood [1,†], Pernille Sarup [1], Peter Skov Kristensen [1], Jihad Orabi [1] and Ahmed Jahoor [1,3]

1 Nordic Seed A/S, Grindsnabevej 25, 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.)
2 Department of Agroecology, Faculty of Technology, Aarhus University, Forsøgsvej 1, Flakkebjerg, 4200 Slagelse, Denmark
3 Department of Plant Breeding, The Swedish University of Agricultural Sciences, 23053 Alnarp, Sweden
* Correspondence: nive@nordicseed.com
† These authors contributed equally to this work.

Abstract: Efficient and stable restoration of male fertility (Rf) is a prerequisite for large-scale hybrid seed production but remains an inherent issue in the predominant fertility control system of rye (*Secale cereale* L.). The 'Gülzow' (G)-type cytoplasmic male sterility (CMS) system in hybrid rye breeding exhibits a superior Rf. While having received little scientific attention, one major G-type *Rf* gene has been identified on 4RL (*Rfg1*) and two minor genes on 3R (*Rfg2*) and 6R (*Rfg3*) chromosomes. Here, we report a comprehensive investigation of the genetics underlying restoration of male fertility in a large G-type CMS breeding system using recent advents in rye genomic resources. This includes: (I) genome-wide association studies (GWAS) on G-type germplasm; (II) GWAS on a biparental mapping population; and (III) an RNA sequence study to investigate the expression of genes residing in Rf-associated regions in G-type rye hybrids. Our findings provide compelling evidence of a novel major G-type non-PPR *Rf* gene on the 3RL chromosome belonging to the mitochondrial transcription termination factor gene family. We provisionally denote the identified novel *Rf* gene on 3RL *RfNOS1*. The discovery made in this study is distinct from known P- and C-type systems in rye as well as recognized CMS systems in barley *(Hordeum vulgare* L.) and wheat (*Triticum aestivum* L.). We believe this study constitutes a stepping stone towards understanding the restoration of male fertility in the G-type CMS system and potential resources for addressing the inherent issues of the P-type system.

Keywords: restoration of male fertility (*Rf*); cytoplasmic male sterility (CMS); pentatricopeptide repeat protein (PPR); mitochondrial transcription termination factor (mTERF); 600K SNP array; genome-wide association study (GWAS); RNAseq; chi-square; linkage disequilibrium

Citation: Vendelbo, N.M.; Mahmood, K.; Sarup, P.; Kristensen, P.S.; Orabi, J.; Jahoor, A. Genomic Scan of Male Fertility Restoration Genes in a 'Gülzow' Type Hybrid Breeding System of Rye (*Secale cereale* L.). *Int. J. Mol. Sci.* **2021**, *22*, 9277. https://doi.org/10.3390/ijms22179277

Academic Editor: Pedro Martínez-Gómez

Received: 16 July 2021
Accepted: 24 August 2021
Published: 27 August 2021

1. Introduction

In recent years, hybrids have become the predominant class of cultivated winter rye (*Secale cereale* L.) in Northern Europe [1]. Outperforming population-based cultivars, hybrids in rye demonstrate strong heterotic effects on all developmental and yield characteristics [2,3]. Breeding of hybrids relies on the existence of cytoplasmic male sterility (CMS) and restoration of male fertility (Rf) genes that reside in genetically distinct parental populations [3,4]. This system efficiently enables control of parental crossing in the field, which is a prerequisite for large-scale hybrid seed production [5]. In hybrid rye, numerous CMS systems exist, of which the most predominant is the Pampa (P) type [6]. In this system, five major P-type *Rf* genes have been identified on 1RS, 4RL (*Rfp1*, *Rfp2*, *Rfp3*), and 6R (dominant modifier) chromosomes, and three minor genes on 3RL, 4RL, and 5R chromosomes [7–10]. The less prevalent CMS systems include the 'Gülzow' (G) type

originating from the population of rye variety 'Schlägler alt' [11], R-type originating from a Russian population [12], and C- [13] and S- [14] types originating from the old Polish population of rye variety 'Smolickie'. In the G-type CMS system, one major gene has been identified on 4RL (*Rfg1*) and two minor genes on 3R (*Rfg2*) and 6R (*Rfg3*) chromosomes [15]. In the C-type CMS system, two major *Rf* genes have been identified on 4RL (*Rfc1*) and 6RS (*Rfc2*) [16,17]. Intriguingly, Stojalowski et al. [18] observed a linkage between major *Rf* genes on 4RL for all three CMS systems, C-type (*Rfc1*), G-type (*Rfg1*), and P-type (*Rfp1*, *Rfp2*, *Rfp3*) to the same marker loci. This finding accentuates the pivotal importance of 4RL across CMS systems in hybrid rye breeding.

Restoration of male fertility in hybrids derived from the predominant P-type cytoplasm is frequently incomplete and highly environmentally unstable [9,19–21]. In addition to a potential loss in grain yield, partial pollination renders the cultivar susceptible to fungal infection by ergot (*Claviceps purpurea* (Fr.) Tul.) which can contaminate the rye grains with toxic sclerotia [22–24]. The P-type system is inherently shaped by the low frequency of restorer gametes in European populations in which the predominance exhibits unsatisfactory restoration [19,20]. In 1991, several non-adapted Argentinian and Iranian rye populations with a high frequency of restorer gametes were identified [25]. Crossing an elite maternal line with one of these non-adapted exotics led to observations of significantly higher restoration levels and environmental stability [26,27]. In order to steer the introgression of novel superior exotic *Rf* genes through marker-assisted selection, molecular markers were developed for *Rfp1*, *Rfp2*, and *Rfp3* [8,9]. Hybrids carrying an exotic *Rf* gene were, however, found to exhibit a significant reduction in grain yield by 4.4% to 9.4% caused by linkage drag effects or epistatic interactions associated with the exotic *Rf* gene [28]. Despite these deleterious effects, hybrid cultivars carrying the exotic *Rfp1* have been introduced to the Northern European market by a patented brand, PollenPlus® [29]. In contrast, hybrids derived from the less prevalent G-type cytoplasm are characterized by a complete and environmental stable restoration of male fertility [30]. However, having received little scientific attention, the underlying genetics of the G-type CMS system remain largely unexplored [15].

The male sterility factors in CMS lines are encoded by mitochondrial genes that cause a defect in the production of viable pollen [31]. While multiple gene families have been linked with male fertility restoration, a distinct clade of the pentatricopeptide repeat (PPR) RNA-binding factor family referred to as *Rf*-like PPR (RFL-PPR) constitutes the predominant class of isolated *Rf* genes [5]. Proteins of the PPR superfamily are characterized by up to 30 tandem repeats of a canonical 35-amino-acid motif [32]. Based on motif composition, PPR genes are divided into two subclasses: the P class solely containing the canonical motif and the PLS class containing triplets of P, L ('long', ≈36 aa), and S ('short', ≈31 aa) motifs [33,34]. While the P class has predominantly been associated with RFL-PPR genes, instances of the PLS class have also been identified [35]. In rye, 591 PPR genes have been identified, out of which 83 belong to the RFL-PPR clade [36] (Table S8a). PPR proteins target mitochondrial or chloroplast mRNA, participating in a range of post-transcriptional processes (RNA editing, splicing, cleavage, and translation) with profound effects on organelle biogenesis and function [33,37,38]. In wheat, the RFL-PPR genes *Rf1* and *Rf3* have been shown to bind to a mitochondrial *orf279* transcript, induce cleavage, and prevent the expression of the sterilizing factor [39]. Within grasses, several isolated *Rf* genes have been characterized as RFL-PPRs, including *Rfm1* in barley [35], *Rf1* in sorghum (*Sorghum bicolore* L.) [40], *Rf5* in maize [41], and *Rf4*, *Rf5*, and *Rf6* in rice (*Oryza sativa* L.) [42–44]. In the C-type CMS system of rye, the *Rfc1* locus has been found to reside in close proximity to a cluster of RFL-PPR genes on 4RL [36].

Another gene family less prevalently associated with Rf is the mitochondrial transcription termination factors (mTERF) [45]. In rye, 131 mTERF genes have been identified [36] (Table S8b). Similar to the PPR, mTERF genes encode helical repeat proteins that target mitochondrial DNA, regulating the expression of mitochondrial genes [46]. Within grasses, mTERF genes have been associated with restoration of male fertility in barley *Rfm3* on

the 6HS chromosome [47] and wheat *Rf9* on the 6BS chromosome [48]. In the P-type CMS system of rye, the *Rfp1* locus has been found to reside in close proximity to four mTERF genes [49,50]. In consistence, hotspots of RFL-PPR and mTERF genes have been identified in regions harboring known *Rf* genes in rye [36].

In this paper, we report an investigation of the genetics underlying male fertility restoration in G-type CMS-based hybrid rye breeding systems. The objective of this study was to identify major and minor G-type *Rf* genes. This was approached through: (I) genome -wide association study (GWAS) on a G-type CMS hybrid rye breeding germplasm; (II) GWAS on a biparental mapping population for studying the inheritance of male fertility restoration; and (III) gene expression analysis of PPR, RFL-PPR, and mTERF genes residing in Rf-associated blocks in two G-type hybrid cultivars for the identification of causative genes. This knowledge will serve as a stepping stone towards developing novel hybrid cultivars exhibiting superior and environmentally stable restoration of male fertility to maximize grain yield and enhance ergot resistance.

2. Results

2.1. Analysis of Genotyping Data

Prior to bioinformatic analysis using the single nucleotide polymorphism (SNP) array genotype data, a quality filtration was conducted to remove monomorphic, non-informative markers. Polymorphism information content (PIC) was calculated as a measure of the identified marker's informativeness, with a mean PIC of 0.26 for the 20K platform (n = 365), 0.34 for the 30K platform (n = 181), and 0.23 for the 600K platform (n = 180). All SNP arrays portrayed a uniform distribution of markers across the rye genome (Table S1). In total, 4419 informative markers were identified in the 20K array on the entire germplasm and can be found thoroughly characterized in a recent study by Vendelbo et al. (2020). A subset of this germplasm was genotyped on the recent rye 600K array, yielding 261406 informative markers. In the F_2 mapping population composed of 181 plants, 3493 informative markers were identified, out of which 1088 were derived from the 5K rye array, 808 from the 600K rye array, and 1597 from the 90K wheat array.

2.2. Genome-Wide Association Study—Case Control

Genome-wide association study (GWAS) was conducted using population origin as phenotypic input in a case control analysis for an initial 'crude' identification of potential restoration of male fertility (Rf) genes in the germplasm. The 20K GWAS analysis produced a distinct peak in the Manhattan plot at 724 to 745 Mbp on 3RL, with the highest associated marker ($-\log_{10}(p)$ = 19.1) located at 745 Mbp (Figure 1A, Table S2). In the successive 600K GWAS analysis, a similar peak was identified at 710–747 Mbp with the highest associated markers ($-\log_{10}(p)$ = 27.07) located between 729 and 730 Mbp (Figure 1B–D, Table S3).

The identified Rf-associated region harbored five PLS-class PPR genes and one mTERF gene in the 'Lo7' reference genome (Figure 1D, Table S7a). In addition, a unique peak portraying a similarly strong association was found on 1RS at 49.3–58.5 Mbp in the 600K GWAS analysis (Figure 1C). The region harbored 22 RFL-PPR genes organized in four clusters, 2 P-class PPR genes, and 1 mTERF gene (Figure 1C, Table S7b). Out of 20 significant associated markers residing at the site, 18 mapped to a narrow peak from 58.02 to 58.47 Mbp on 1 RS (Table S3).

2.3. Biparental Population

A biparental F_2 population consisting of 181 plants was developed from the 'Gülzow' (G)-type hybrid cv. Stannos. The population was phenotyped for six Rf-associated traits as well as traits related to restoration in order to obtain a comprehensive dataset on the inheritance of G-type *Rf* genes. Seed number and pollen production were found, on the basis of our observations, to be the most representative Rf-associated traits (Figure 2A,B).

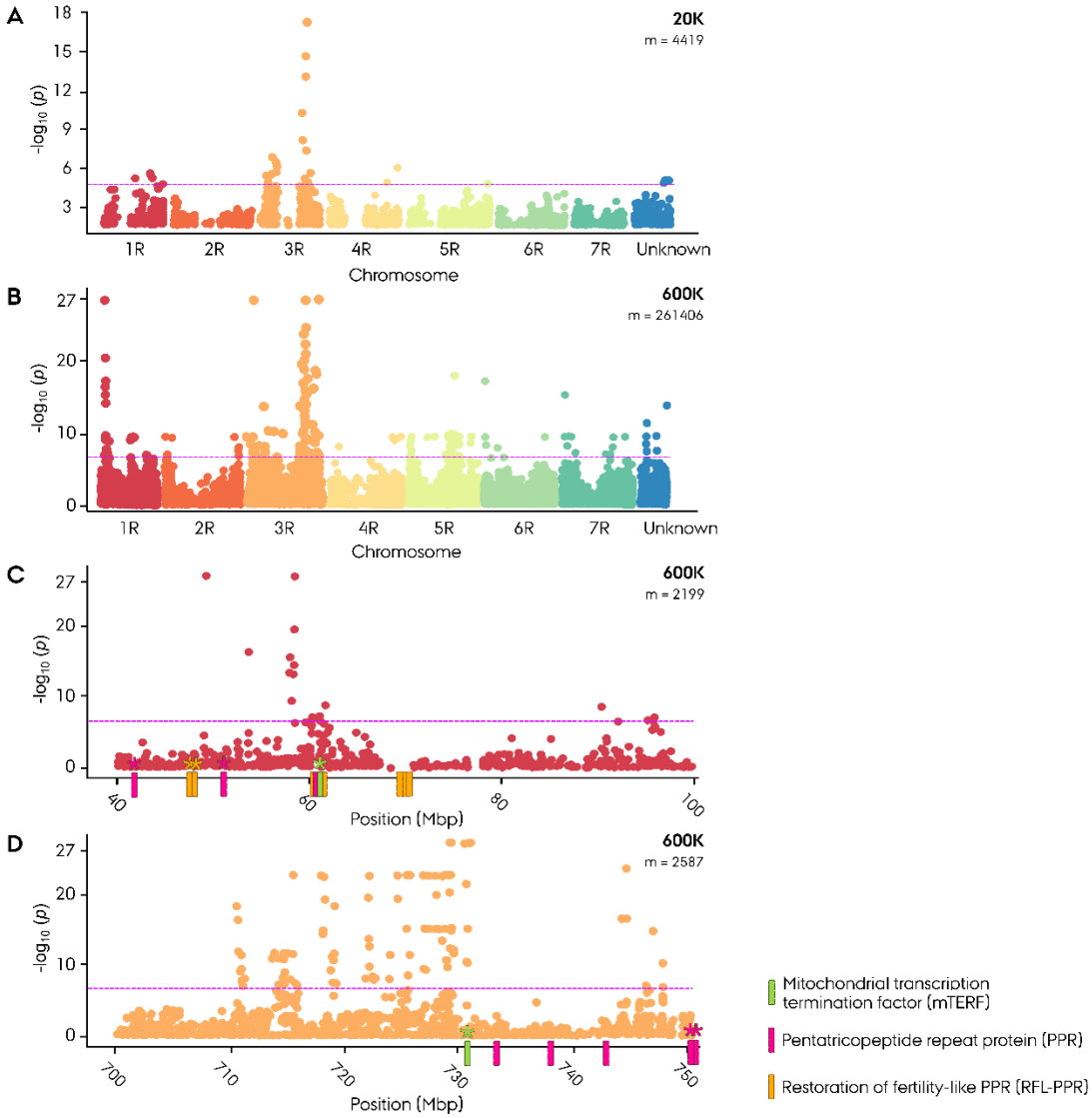


Figure 1. Manhattan plot for genome-wide association study (GWAS) on population origin ('case control') in Nordic Seed elite hybrid rye (*Secale cereale* L.) breeding germplasm. (**A**) Genome-wise Manhattan plot of 20K SNP array GWAS (n = 365). (**B**) Genome-wise Manhattan plot of 600K SNP array GWAS (n = 180). (**C**) Manhattan plot of the 600K SNP array 1RS region. (**D**) Manhattan plot of the 600K SNP array 3RL region. Position of major restoration of male fertility-associated genes has been included with genes expressed in G-type hybrids marked with an asterisk. Significant association was identified using criterion of $-\log_{10}(p) > 4.95$ in 20K and >6.72 in 600K depicted as a magenta line.

The observed segregation ratio of sterile and fertile F_2 plants was tested for goodness of fit to the expected Mendelian ratio at the scenario of one, two, and three major *Rf* genes using an χ^2 test. Intriguingly, the observed segregation ratios were in accordance with a monogenic dominant inheritance of male fertility restoration with χ^2 (1, $n_{,\text{infertile}}$ = 38, $n_{,\text{fertile}}$ = 143) = 2.26, p = 0.13 for seed number and χ^2 (1, $n_{\text{infertile}}$ = 43, n_{fertile} = 138) = 1.11, p = 0.29 for pollen production (Table S4). GWAS led to the identification of 16 Rf-associated SNP markers, of which 5 markers showed a significant association with $-\log_{10}(p) > 5.2$ (Figure 3). On 3RL, a twin peak was identified in the GWAS. The first peak spanning from 627 to 769 Mbp with the highest associated marker derived from the 90K wheat array ($-\log_{10}(p) = 6.66$) was localized at 627 Mbp. The majority of the Rf-associated markers were located around 745 Mbp (Figure 3, Table S4). The second peak, comprising four markers, spanned from 807.1 to 808.7 Mbp with its highest associated marker ($-\log_{10}(p) = 4.68$) just below the significance threshold of 4.85.

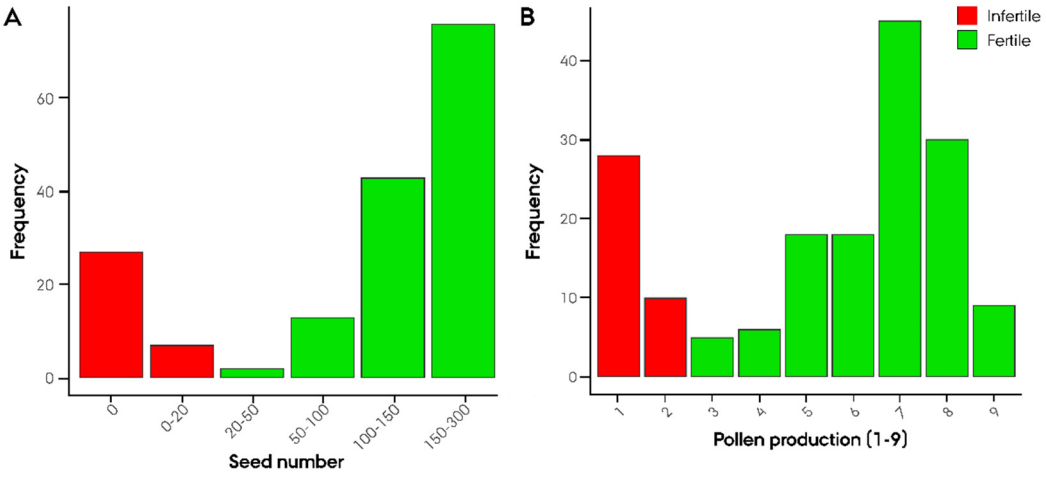


Figure 2. Phenotypic distribution of restoration of male-fertility-related traits. (**A**) Seed number and (**B**) pollen production in 181 F_2 plants derived from a 'Gülzow' type hybrid rye (*Secale cereale* L.) cv. Stannos.

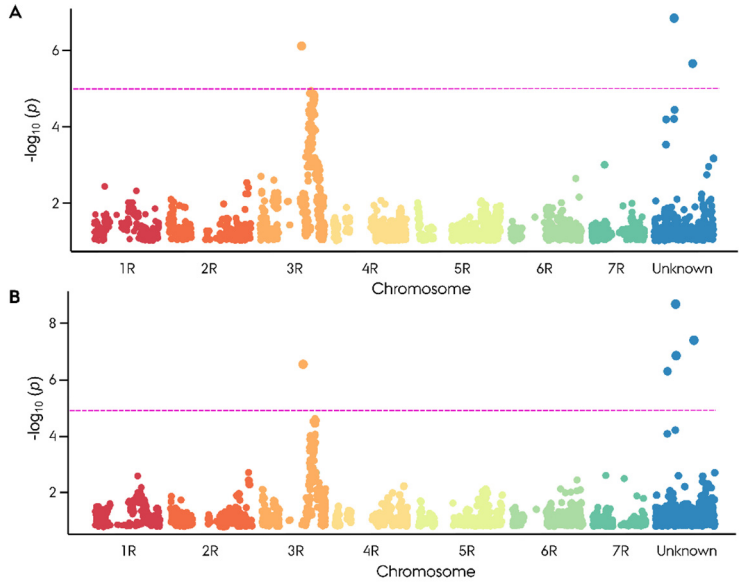


Figure 3. Manhattan plot for genome-wide association study on the restoration of male-fertility-related phenotypic scores. (**A**) Seed number and (**B**) pollen production (1–9) in an F_2 biparental population composed of 181 plants derived from the hybrid rye cv. Stannos. In total, 3494 informative SNP markers were used. Significant association was identified using criterion of $-\log_{10}(p) > 4.85$ depicted as a magenta line.

The remaining four significantly associated SNP markers were derived from the 90K wheat SNP array. None of the four markers could successfully be mapped to the 'Lo7' rye reference genome. Two of these markers mapped to the short arm of the wheat 3B chromosome, including the highest Rf-associated ($-\log_{10}(p)$ = 9.12) wheat marker AX_158558079. One of the remaining moderately Rf-associated markers mapped to the short arm of the wheat 1A chromosome, while the last had no available mapping position in wheat. With no mapping position, genome-wide pairwise linkage disequilibrium was calculated for each of the four highly Rf-associated wheat-derived SNP markers to determine their position based on linkage to mapped markers (Table S5). All four wheat markers exhibited a singular peak on 3RL with a top LD ranging from 0.43 to 0.97 in the region spanning 701 to 747 Mbp (Figure S2). The top Rf-associated wheat marker AX_158558079 exhibited a max LD of 0.85 at 747 Mbp (Figure 4A,B). None of the

Rf-associated wheat-derived markers showed linkage towards the Rf-associated marker cluster at 807.1–808.7 Mbp (Table S9).

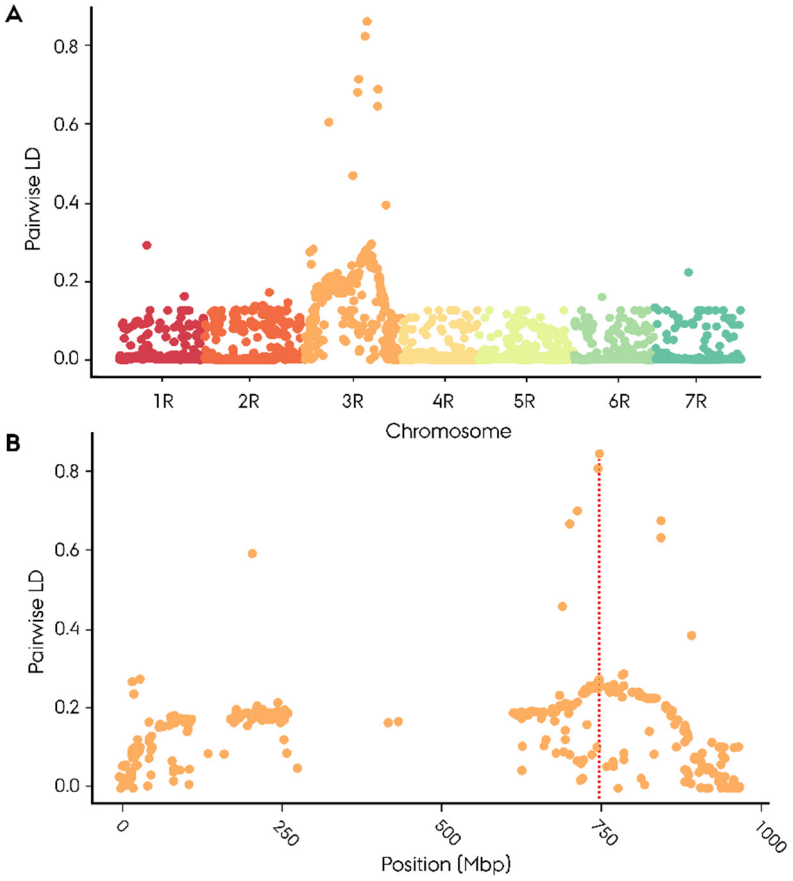


Figure 4. Mapping of an SNP marker derived from 90K wheat array highly associated with restoration of male fertility in 181 (*Secale cereale* L.) F_2 plants by genome-wide pairwise linkage disequilibrium (LD) towards 2448 informative SNP markers with accurate mapping position on 'Lo7' rye reference genome. (**A**) Genome-wide distribution of LD, (**B**) LD distribution on 3R chromosome.

2.4. Genomic Scan and Expression of Genes Residing within Restoration of Male-Fertility-Associated Regions of G-Type Hybrids

To identify a likely candidate *Rf* gene, a genomic scan of identified Rf-associated regions was performed and gene expression was investigated in two G-type hybrids. As the RNA-seq data were obtained from spikes at flowering, active expression of *Rf* genes would be expected in the fertile hybrids. The Rf-associated region spanning 710 to 760 Mbp was found to harbor 448 genes in the 'Lo7' reference genome, out of which 272 were expressed in cv. Helltop and 266 in cv. Stannis (Table S7a). Among the expressed genes, 251 were co-expressed in both hybrids. Amongst the panel of PPR, RFL-PPR and mTERF genes residing in the Rf-associated regions, and two out of five PLS-class PPR genes, at 752.1 and 759.1 Mbp, were co-expressed in both G-type hybrids (Figure 1D, Table 1). The site also harbored a single mTERF gene at 731.7 Mbp, likewise co-expressed in both G-type hybrids. The mTERF gene co-localized with the top associated marker in the 600K case control GWAS with less than 1 Mbp distance (Figure 1D, Table S3).

Table 1. Pentatricopeptide repeat protein (PPR) and mitochondrial transcription termination factor (mTERF) genes in restoration of male-fertility-associated regions identified on 1RS (40–70 Mbp) and 3RL (710–760 Mbp) chromosome in 'Lo7' rye (*Secale cereale* L.) reference genome. Genes expressed in 'Gülzow' type hybrid cultivars at flowering are marked as green, and genes not expressed as red.

Chromosome	Position (Mbp)	Gene Length (bp)	Annotation	Expression	
				cv. Helltop	cv. Stannos
1R	42.9	822	PPR (P Type)		
	46.2	2091	RFL-PPR (P Type)		
	46.2	2613	RFL-PPR (P Type)		
	46.3	2601	RFL-PPR (P Type)		
	46.9	2487	RFL-PPR (P Type)		
	47.1	2490	RFL-PPR (P Type)		
	47.1	1833	RFL-PPR (P Type)		
	51.9	1659	PPR (P Type)		
	61.1	417	RFL-PPR (P Type)		
	61.3	287	PPR (P Type)		
	61.3	213	PPR (P Type)		
	61.5	2499	RFL-PPR (P Type)		
	61.5	1866	mTERF		
	61.8	2472	RFL-PPR (P Type)		
	61.9	2448	RFL-PPR (P Type)		
	61.9	1460	RFL-PPR (P Type)		
	61.9	2451	RFL-PPR (P Type)		
	62.0	2499	RFL-PPR (P Type)		
	70.8	510	mTERF-like		
	70.8	540	mTERF-like		
3R	731.7	462	mTERF		
	733.4	1986	PPR (PLS Type)		
	738.4	1743	PPR (PLS Type)		
	743.7	1461	PPR (PLS Type)		
	751.6	1356	PPR (PLS Type)		
	752.1	1548	PPR (PLS Type)		
	759.7	825	PPR (PLS Type)		

The Rf-associated site on 1RS spanning from 40 to 70 Mbp was found to harbor 380 genes in the 'Lo7' reference genome, out of which 166 were expressed in cv. Helltop and 160 in cv. (Table S7b) Stannos. Among the expressed genes, 144 were co-expressed in both G-type hybrids. Out of 13 RFL-PPR genes residing in the region, three situated at 46.2–47.1 Mbp were co-expressed in both G-type hybrids (Figure 1C, Table 1). The site also harbored two P-class PPR genes, at 42.9 and 51.9 Mbp, and a single mTERF gene, at 61.5 Mbp, likewise co-expressed in both G-type hybrids. The mTERF gene resided 3.2 Mbp from the top associated marker on 1RS.

3. Discussion

While the less common 'Gülzow' (G)-type system demonstrates superior restoration of male fertility, it has received little scientific attention in the past. This is the first study since Melz et al. [30] in 2003 to investigate the genetics underlying male fertility restoration in G-type CMS hybrid rye breeding systems. Until now, only three G-type restoration of male fertility (Rf) genes have been reported, a major gene located on 4RL (*Rfg1*) and two modifying genes on 3R (*Rfg2*) and 6R (*Rfg3*) [15,51]. By exploiting recent advents in rye genomic resources, we succeeded in identifying a novel major G-type *Rf* gene on 3RL, in addition to further evidence of a major gene on 1RS and a modifying gene on 3RL

chromosome. Our findings provide a novel insight into the differentiation of the G-type fertilization control system from the predominant P-type.

3.1. Indications of a Major Restoration of Male-Fertility-Like Pentatricopeptide Repeat Gene on 1RS

While case control genome-wide association study (GWAS) is a useful tool for providing an insight into the genetics differentiating the parental gene pools, it has several limitations. Using population origin of lines as 'phenotypic' input, statistically associated markers in case control GWAS, can either be a population-defining trait such as an *Rf* QTL, or a product of population structure.

In a recent population study by Vendelbo et al. [52] on the entire G-type hybrid rye elite breeding germplasm, the maternal NRG & CMS population was found to exhibit considerable population structure and vast LD blocks. Unequal relatedness among individuals and population structure introduces a confounding effect that might cause spurious marker associations and introduce a risk of false positives [53,54]. To moderate the effect of these confounding factors, the GAPIT software used to conduct the GWAS therefore utilizes a compressed mixed linear model [55,56]. Large LD blocks, on the other hand, introduce an uncorrectable confounding factor. Long-distance LD complicates the disentanglement of actual causal variants from linked neutral markers, which can in turn lead to spurious associations [54].

In the 600K case control GWAS, a unique strong peak was identified on 1RS (Figure 1B,C). While the evidential significance of case control GWAS is insufficient to draw definitive conclusions, it provides an insight into pivotal genomic sites differentiating the parental populations. Intriguingly, the region was found to harbor 22 out of the 83 annotated *Rf*-like PPR (RFL-PPR) genes in the 'Lo7' reference genome situated in three large clusters (Figure 1C) [36]. In wheat, the syntenic segment has been reported to house two major *Rf* genes, *Rf1* (1AS) and *Rf3* (1BS), both belonging to the RFL-PPR gene family [39,57,58]. In rye, a major P-type *Rf* gene has likewise been identified on 1RS in a German inbred rye line 'L18' [7]. This *Rf* gene has, however, received little attention and it remains unknown whether the underlying *Rf* gene belongs to the RFL-PPR family.

Examining the RNA-seq data, none of the eight RFL-PPR genes residing in close proximity to the top associated marker peak were, however, found to be co-expressed at the flowering stage in the G-type hybrids (Figure 1C, Table S7b). Instead, a mitochondrial transcription termination factor (mTERF) gene residing 3.2 Mbp from the peak was found to be co-expressed.

While these findings suggest that the germplasm houses an additional major G-type *Rf* gene on 1RS, we did not observe any Rf-associated QTLs on 1RS in the mapping population GWAS (Figures 3 and 4). This can either be due to the absence of the major *Rf* gene on 1RS in the pollen father of cv. Stannos, or that the region harbors a population-defining trait other than an *Rf* gene. Given that the locus aligns with a region containing numerous known Rf genes in both wheat and rye, as well as an mTERF gene co-expressed in both G-type hybrids, it is not unlikely that a large G-type Rf gene occurs on 1RS in the germplasm.

3.2. Modifying G-Type Restoration of Male Fertility Genes

While incapable of restoring male fertility, minor *Rf* genes are believed to enhance ('modify') the effect of major *Rf* genes [7]. At present, two minor *Rf* genes have been identified in the G-type CMS hybrid rye breeding systems on the 3R (*Rfg2*) and 6R (*Rfg3*) chromosomes [15]. In the mapping population, we identified an *Rf* QTL at 807.1–808.7 Mbp, indicative of a minor *Rf* gene (Table S4). While no positional information exists on the *Rfg2* gene, a minor P-type *Rf* gene has been identified in the German inbred line 'L18' with associated markers mapping to 806.1 and 869.5 Mbp [7]. Calculation of pair-wise LD ruled out the presence of spurious association between the potential minor *Rf* gene and the downstream major *Rf* gene, confirming an independent QTL (Table S9). It remains uncertain whether the potential minor *Rf* gene represents a unique gene or the previously described *Rfg2*.

3.3. Decisive Role of 3R in the G-Type CMS Breeding System

The initial genome scan for population-differentiating traits using case control GWAS indicated that the 3R chromosome played a unique role in the G-type CMS system (Figure 1A,B). This is consistent with previous population studies on the assayed germplasm, identifying a singular enrichment of interchromosomal LD for both parental populations on the 3R chromosome, suggesting a conservation of a population-defining trait or traits [52].

To investigate whether the population-differentiating region on 3RL harbored a G-type *Rf* gene, a biparental mapping population was developed. In contrast to the case control GWAS, the biparental mapping population is not subject to confounding issues related to population structure. The segregation ratio of Rf-associated traits in the mapping population was found to be in accordance with a monogenic dominant inheritance of an *Rf* gene by χ^2 test consistent with the singular peak identified in the case control GWAS (Figure 1A,B). Since minor *Rf* genes are incapable of restoring male fertility on their own, they do not influence the segregation ratio of infertile/fertile F_2 progeny and are hence not 'caught' in the χ^2 test. Utilizing the phenotypic dataset on Rf-associated traits from the mapping population, we identified an Rf-associated region on 3RL consistent with findings in the initial case control GWAS (Figure 3A,B). The precise position of the causative *Rf* gene was, however, initially obscured by the finding that the four most associated SNP markers, deriving from the 90K wheat array, could not be accurately mapped to the rye reference genome 'Lo7' [36,59]. Instead, chromosome-wide LD mapping of each Rf-associated wheat marker was conducted, with the top associated marker mapping to 747 Mbp (Figure 4, Table S5). Intriguingly, while the 4R chromosome plays a pivotal role in the P- and C-type CMS systems in rye, housing *Rfp1*, *Rfp2*, *Rfp3*, and *Rfc1*, we found no evidence of G-type genes on 4R in either case control or mapping population GWAS (Figures 1 and 3) [8,9,28,60]. Instead, our findings suggest that the 3R chromosome plays a unique role in the G-type CMS system, possibly harboring both a minor and major *Rf* gene.

3.4. Novel Major Restoration of Male Fertility Gene Unique to the G-Type CMS Breeding System

While no major *Rf* gene on the 3RL chromosome has to our knowledge been identified in either of the known CMS systems in rye, minor genes have been identified in both the G-type and P-type [7,15]. In the case of *Rfg2*, ambiguous segregation ratios of primary trisomics of rye 3R led to the assumption that 3R likely housed a minor G-type *Rf* gene. While Melz and Adolf [15] cautiously interpreted this anomaly as a product of a modifying gene, their findings suggest that something of significance is occurring on 3R in the G-type CMS system. It therefore remains open whether the identified major G-type gene is *Rfg2*, previously misclassified as a minor gene, or an unreported *Rf* gene on 3RL. We propose to denote the major G-type *Rf* gene on 3RL *RfNOS1.*

To our knowledge, no *Rf* gene has been reported on chromosome segments orthologous to rye 3RL in any of the domesticated species residing within the botanical tribe Triticeae. In wheat, major *Rf* genes have been identified on 1AS (*Rf1*), 1BS (*Rf3*), 6AS (*Rf9*), 6BS (*Rf4*, *Rf6*), 6D (*Rf5*), and 7D (*Rf2*) chromosomes [48,57,58,61–63]. In barley (*Hordeum vulgare* L.), two major *Rf* genes have been identified on 6HS (*Rfm1*, *Rfm3*) [47,64]. Intriguingly, Martis et al. [65] discovered that the distal regions of 3RL and 4RL conserved syntenic segments of an ancestral Triticeae chromosome a6. In a comparative analysis, they found that the segment on 3RL portrayed distinctly less collinearity than all other syntenic segments, suggesting a differential evolution of 3RL during rye speciation. On the contrary, the syntenic segment on 4RL was found to be highly conserved in *Brachypodium distachyon* L., rice (*Oryza sativa* L.), sorghum (*Sorghum bicolore* L.), and barley. Rye 4RL, a region housing three major P-type and one C-type *Rf* genes, was found to be syntenic to barley 6HS [60,65]. These results are consistent with the findings of Hackauf et al. [50], who reported that the segment housing *Rfp1* on 4RL exhibits an ortholog on wheat 6DS and barley 6H. In a subsequent study, Hackauf et al. [8], furthermore, proposed that *Rfp3* on 4RL likely maps to an orthologous segment housing *Rf6* on wheat 6BS and

Rfm1 on barley 6HS [62]. These findings provide further evidence of a conserved synteny between these chromosomal segments and the conservation of *Rf* genes pivotal to the fertilization control systems across rye (P- and C-type), barley, and wheat [51]. Together, this accentuates the novelty of our discovery of a major *Rf* gene, *RfNOS1*, on 3RL in the G-type CMS system in rye, with no known *Rf* genes on orthologous chromosome segments in other Triticeae species.

3.5. Non-Pentatricopeptide Repeat Restoration of Male Fertility Gene on 3RL

The majority of characterized *Rf* genes have been assigned to the PPR superfamily, denoted as *Rf*-like PPR genes or RFL-PPRs. In domesticated Poaceae species, this includes wheat *Rf1* and *Rf3* [39], barley *Rfm1* [35], sorghum *Rf1* [40], maize *Rf5* [41], and rice *Rf4*, *Rf5*, and *Rf6* [42–44]. A genome scan for Rf-associated genes at the site on 3RL harboring the major G-type *Rf* gene *RfNOS1* in the 'Lo7' reference genome revealed no RFL-PPR genes. Instead, the region harbored five PLS-class PPR genes either misclassified, not resembling known PLS-type RFL-PPR genes, or pointing towards a non-PPR *Rf* gene on 3RL (Figure 1D, Table S8a). Out of 83 RFL-PPRs annotated in the 'Lo7' reference genome, none of them belong to the PLS class. Whereas a less prevalent group, PLS-type RFL-PPR genes exists, as observed in the case of *Rfm1* in barley [35]. It, therefore, cannot be ruled out that the RFL-PPR annotation of 'Lo7' is incomplete, inclined to annotate the prevalent P-class as RFL-PPR genes, while disregarding the less well-represented PLS class. However, both of the PLS-class PPR genes residing in the Rf-associated region on 3RL that were co-expressed in the two G-type hybrids resided more than 20 Mbp from the top associated peak in the 600K case control GWAS (Figure 1D, Table S8b). In conjunction, this suggests that the major *Rf* gene on 3RL belongs to the non-PPR family.

A growing body of *Rf* genes is now being characterized as non-PPR *Rf* genes, adding to the complexity of male fertility restoration. Until now, this includes glycine-rich proteins (Rf2, in rice; [66]), acyl-carrier protein synthase (Rf17, in rice; [67]), aldehyde dehydrogenase (Rf2, in maize; [68]), bHLH transcription factor (Rf4, in maize; [69]), and *Rf1*, a peptidase, in sugar beet (*Beta vulgaris* L.) [70]. The region on 3RL harbored several genes related to plant fertility that were co-expressed in both G-type hybrids at anthesis (Table S7a). Amongst these were two KATANIN genes located at 730.3 and 759.9 Mbp. In *Arabidopsis*, KATANIN is required for fertility, embryo development, and seed production [71]. Genes involved in flower development, including MADS-box transcription factor [72–74], BZIP transcription factor [75], and basic helix-loop-helix (bHLH) DNA-binding superfamily protein [69] were likewise found co-expressed within the region.

Intriguingly, the region was found to harbor a single mTERF gene located at 731 Mbp, coinciding with the top-associated peak in the 600K case control GWAS (Figure 1D, Table S8b). In rye, Hackauf et al. [8] reported a close linkage between *Rfp1* and *Rfp3* on 4RL to mTERF genes. A similar observation was made by Bernhard et al. [47], who identified two mTERF genes closely linked to *Rfm3* on barley 6HS, syntenic to rye 4RL. Pan et al. [76] successively observed the role of mTERF genes in kernel development in maize, connecting the gene family to the reproductive system of plants. This newly identified mTERF gene, which is expressed at flowering in both G-type hybrids, is a potential candidate for the novel major G-type *Rf* gene on 3RL. We denote the gene as *RfNOS1*, belonging to the expanding class of non-PPR *Rf* genes.

3.6. CMS Systems in Hybrid Rye Breeding

On the basis of male fertility restoration requirements and genetic similarity of sterilizing cytoplasms, G-, C-, and R-type CMS systems have been proposed to belong to the larger Vavilovii (V) type [6,77,78]. In a comprehensive study by Lapinski and Stojalowski [6] on 50 rye populations from 23 countries, the vast majority of male sterility sources were found to belong to the V-type. Populations with European descent were predominantly found to carry the V-type sterility-inducing cytoplasm, while the P-type was exclusively observed in lines descending from South America. Nonetheless, with no previous report of a major

Rf gene on 3RL in either R, S, or C-type CMS system, such a unilateral grouping as V-type seems premature. From our observations, the G-type CMS system distinguishes itself from the other CMS systems by a less pivotal role of *Rf* genes on 4RL.

4. Methods

4.1. Plant Material

In total, 365 Nordic Seed Germany GmbH inbred hybrid rye (*Secale cereale* L.) elite breeding lines were selected for this study, comprising 242 restorers, 116 non-restorer germplasm (NRG), and 7 cytoplasmic male sterile (CMS) lines. The CMS male sterility is based on the 'Gülzow' (G)-type cytoplasm originating from the population of the rye variety 'Schlägler alt' [11,15]. Population structure and information on genetic characteristics of the germplasm are presented in a recent study by Vendelbo et al. [52]. A biparental mapping population was developed from a hybrid rye cv. Stannos. DNA extraction was performed using an adapted SDS-based method according to USDA [79] after Pallotta et al. [80] on an equivalent of 75 mg of plant material collected from the primary leaves of two seven-day-old, seedlings per line. DNA concentration and 260/280 nm ratio of samples were measured using an Epoch^{TM} microplate spectrophotometer (Biotek® Vermont, Winooski, VT, USA) and evidence of fragmentation by size visualization on a 1.2% agarose gel.

4.2. Biparental Mapping Population

To investigate the inheritance of male fertility restoration in the G-type CMS-based Nordic Seed breeding system, a biparental mapping population was developed. The population was phenotyped for restoration of male fertility, including traits associated with restoration. Seeds of the hybrid cv. Stannos (F_1) were sown in pots containing a course-grain sphagnum substrate at Nordic Seed Germany GmbH greenhouse facilities. The seedlings were cultivated under a 16 h light regime with night temperatures of 14–16 °C and day temperatures of 18–24 °C. Seven days after sowing, at the 2-leaf stage, seedlings were set to vernalize in a climate chamber under 16 h of light at 8 °C for a week and then 3 °C for the following seven weeks. After vernalization, the pots were transferred to the greenhouse. Prior to anther-protrusion, cellophane bags were put on the spikes to prevent cross-fertilization. At maturity, seeds of a single F_1 plant were harvested and the procedure was repeated to generate an F_2 biparental mapping population. At four timepoints, a total of 181 F2 plants were phenotyped for pollen production using a customized visual 1–9 scale (1: no pollen, 9: large amount of pollen). The plants were, furthermore, scored for number of spikes per plant, total seed number, seeds per spike, total grain weight, and thousand kernel weight in order to obtain a comprehensive phenotypic dataset on the inheritance of male fertility restoration in the population. The segregation ratio of infertile and fertile F_2 plants was tested for goodness of fit to the expected Mendelian ratio at the scenario of one, two, and three major restorations of male fertility (*Rf*) genes using an χ^2 test [81]. An F_2 plant was considered 'sterile' if it either yielded less than 20 seeds or scored ≤ 2 in pollen production.

4.3. Molecular Markers

All rye lines included in this investigation were genotyped using a custom Illumina Infinium $15K_{wheat}$ [59] and $5K_{Rye}$ [82,83] single nucleotide polymorphism (SNP) array, denoted 20K, as described by Vendelbo et al. (2020). In addition, 180 lines comprising 88 NRG and 92 restorer lines were also genotyped using the state-of-the-art 600K high-density rye array by Bauer et al. [83]. The F_2 biparental mapping population was genotyped on a custom Illumina Infinium $25K_{wheat}$ and $5K_{Rye}$ SNP array, denoted 30K, enriched with an additional 10K wheat markers, compared to the previous 20K array, deriving from the 90K wheat SNP array by Wang et al. [59]. The mapping position of SNP markers derived from the 90K wheat array was found by mapping the marker sequences to the 'Lo7' reference genome using the NCBI blastn (v. 2.9.0+, ML, USA) function at a significance threshold of the e-value at 10^{-5}, selecting the physical position of the top hit [36,84]. The

position of *Rf*-associated markers without accurate mapping position in the 'Lo7' reference genome was determined by calculation of pairwise LD across the entire mapped marker panel. LD was calculated using snpStats (v. 1.36.0, Cambridge, UK) R package with LD set as the coefficient of determination (r^2) [85].

4.4. Data Analysis

Genetic analysis of SNP marker data was performed in R studio (v. 1.1.463, Boston, MA, USA) interface in R statistical software (v. 3.6.3) by application of various predesigned packages [86,87]. Prior to analysis, markers were filtered for marker allele frequency $\geq$0.005, missing individual score $\leq$ 0.2 and missing marker score $\leq$ 0.1 to identify informative markers. The polymorphism information content (PIC) was calculated as an estimate of marker informativeness using SnpReady (v. 0.9.6, Laguna, Philippines) [88].

4.5. Genome-Wide Association Study

The discovery of *Rf*-associated regions was made by a genome-wide association study (GWAS) using the genomic association and prediction integration tool (GAPIT) (v.3) package in R [55]. Phenotypic input for GWAS included all recordings of the biparental F_2 population, and a binary case control for the entire population relative to their population origin using the 20K SNP array and 600K high-density SNP array, respectively. A standard Bonferroni-corrected threshold of $\alpha = 0.05$ was used as the significance threshold

4.6. RNA-Seq Data Expression Analysis of PPR and mTERF Genes Residing in Rf-Associated Region in G-Type Hybrids of Rye

Annotated major *Rf*-associated genes, pentatricopeptide repeat proteins (PPR), Rf-like PPR, and mitochondrial transcription termination factors (mTERF), residing in regions associated with Rf were identified in the 'Lo7' reference genome [36]. To ascertain the potential causative Rf gene or genes amongst the identified panel, gene expression was investigated in two G-type Nordic Seed hybrid cv. Helltop and cv. Stannos de novo transcriptome assemblies. The transcript data were obtained from spikes of the G-type hybrids at flowering. High-quality de novo transcriptome assembly of the two hybrids has recently been published by Mahmood et al. [89] and raw reads from the transcript library have been made accessible in a sequence read archive repository at (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA612415, accessed on 12 March 2021). Coding and protein sequences of the Rf-associated gene panel were extracted from the 'Lo7' reference genome repository. The de novo assembled transcriptome assemblies were likewise translated to coding protein sequences with TransDecoder (v. 5.5.0, Valencia, Spain). The NCBI blastx and blastp (v. 2.9.0+, Maryland, USA) functions were used to blast coding and protein sequences of Rf-associated genes in the panel towards the generated protein transcriptome database at a significance threshold of the e-value at 10^{-5} to ascertain genes expressed in the two G-type hybrid cv. at flowering stage [84].

4.7. Graphical Editing

Graphs and figures were outputted from R in .svg format and manually curated using Inkscape (v. 1.1) program (https://inkscape.org/, accessed on 7 May 2021).

5. Conclusions

In this study, we exploited recent advents in rye genomic resources to dissect the genetics underlying restoration of male fertility in a G-type CMS system. Our findings provide compelling evidence of a novel major G-type *Rf* gene on 3RL with no known orthologues in either barley or wheat. Gene mining of the Rf-associated region on 3RL led to the identification of an mTERF gene co-expressed in two G-type hybrids as a candidate gene for restoration. We propose to denote the novel G-type *Rf* gene as *RfNOS1*. Conclusively, our investigation provides a novel insight into the genetics of male fertility restoration in

a G-type CMS system and its differentiation to rye P- and C-types in addition to known CMS systems in barley and wheat.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms22179277/s1.

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

Funding: This research was funded by Innovation Fund Denmark (8053-00085B and 7039-00016B) and Pajbjerg Foundation.

Data Availability Statement: The data presented in this study are available in the Supplementary Material.

Acknowledgments: We would like to thank laboratory technician Hanne Svenstrup at Nordic Seed A/S (Dyngby, Denmark) for contribution to the genotypic data collection and Anette Deterding and Marlene Walbrodt at Nordic Seed Germany GmbH (Nienstädt, Germany) for phenotyping the F_2 population. The plant material included in this study was provided by Nordic Seed Germany GmbH (Nienstädt, Germany). The 20K SNP genotyping was performed by Trait Genetics (Gatersleben, Germany). The 600K SNP genotyping was performed by Eurofins Genomics (Skejby, Denmark).

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

References

1. Miedaner, T.; Hübner, M. Quality demands for different uses of hybrid rye. *Tag. Ver. Pflanz. Saatgutkaufleute Osterr.* **2011**, *61*, 45–49.
2. Laidig, F.; Piepho, H.P.; Rentel, D.; Drobek, T.; Meyer, U.; Huesken, A. Breeding progress, variation, and correlation of grain and quality traits in winter rye hybrid and population varieties and national on-farm progress in Germany over 26 years. *Theor. Appl. Genet.* **2017**, *130*, 981–998. [CrossRef] [PubMed]
3. Geiger, H.H.; Miedaner, T. Hybrid rye and heterosis. In *Genetics and Exploitation of Heterosis in Crops*; American Society of Agronomy-Crop Science Society of America: Madison, WI, USA, 1999; pp. 439–450.
4. Chen, L.; Liu, Y.G. Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* **2014**, *65*, 579–606. [CrossRef] [PubMed]
5. Kim, Y.J.; Zhang, D. Molecular control of male fertility for crop hybrid breeding. *Trends Plant Sci.* **2018**, *23*, 53–65. [CrossRef] [PubMed]
6. Lapinski, M.; Stojalowski, S. Occurence and genetic identity of male sterility-inducing cytoplasm in rye. *Plant Breed. Seed Sci.* **2003**, *48*, 7–23.
7. Miedaner, T.; Glass, C.; Dreyer, F.; Wilde, P.; Wortmann, H.; Geiger, H.H. Mapping of genes for male-fertility restoration in 'Pampa' CMS winter rye (*Secale cereale* L.). *Theor. Appl. Genet.* **2000**, *101*, 1226–1233. [CrossRef]
8. Hackauf, B.; Bauer, E.; Korzun, V.; Miedaner, T. Fine mapping of the restorer gene Rfp3 from an Iranian primitive rye (*Secale cereale* L.). *Theor. Appl. Genet.* **2017**, *130*, 1179–1189. [CrossRef]
9. Stracke, S.; Schilling, A.G.; Forster, J.; Weiss, C.; Glass, C.; Miedaner, T.; Geiger, H.H. Development of PCR-based markers linked to dominant genes for male-fertility restoration in Pampa CMS of rye (*Secale cereale* L.). *Theor. Appl. Genet.* **2003**, *106*, 1184–1190. [CrossRef]
10. Niedziela, A.; Brukwinski, W.; Bednarek, P.T. Genetic mapping of pollen fertility restoration QTLs in rye (*Secale cereale* L.) with CMS Pampa. *J. Appl. Genet.* **2021**, *62*, 1–14. [CrossRef]
11. Adolf, K. A new source of spontaneous sterility in winter rye-preliminary results. *Proc. EUCARPIA Meet. Cereal Sect. Rye* **1986**, *1985*, 11–13.
12. Kobylianskii, V.D. Production of sterile analogues of winter rye varieties, sterile mantainers and fertile restorers. *Trudy Po Prikladnoi Botanike Genetike Selektsii* **1971**, 76–85.
13. Lapinski, M. Cytoplasmic-genic type of male sterility in Secale montanum Guss. *Wheat Inf. Serv.* **1972**, *35*, 25–28.
14. Madej, L. Research on male sterility in rye. *Hodowla Rosl Aklim Nasienn* **1975**, *10*, 590–593.
15. Melz, G.; Adolf, K. Genetic analysis of rye (*Secale cereale* L.) genetics of male sterility of the G-type. *Theor. Appl. Genet.* **1991**, *82*, 761–764. [CrossRef]

16. Milczarski, P.; Hanek, M.; Tyrka, M.; Stojalowski, S. The application of GBS markers for extending the dense genetic map of rye (*Secale cereale* L.) and the localization of the Rfc1 gene restoring male fertility in plants with the C source of sterility-inducing cytoplasm. *J. Appl. Genet.* **2016**, *57*, 439–451. [CrossRef] [PubMed]
17. Stojalowski, S.; Apińaski, M.; Masojć, P. RAPD markers linked with restorer genes for the C-sources of cytoplasmic male sterility in rye (*Secale cereale* L.). *Plant Breed.* **2004**, *123*, 428–433. [CrossRef]
18. Stojalowski, S.; Jaciubek, M.; Masojć, P. Rye SCAR markers for male fertility restoration in the P cytoplasm are also applicable to marker-assisted selection in C cytoplasn. *J. Appl. Genet.* **2005**, *46*, 371–373. [PubMed]
19. Yuan, Y. *Umweltstabilität der Cytoplasmitsch-Genisch Vererbten Männlichen Sterilität (CMS) bei Roggen (Secale cereale L.)*; Verlag UE Grauer: Stuttgart, Germany, 1995.
20. Geiger, H.H.; Yuan, Y.; Miedaner, T.; Wilde, P. Environmental sensitivity of cytoplasmic genic male sterility (CMS) in *Secale cereale* L. *Fortschr. Pflanz.* **1995**, *18*, 7–18.
21. Kodisch, A.; Wilde, P.; Schmiedchen, B.; Fromme, F.-J.; Rodemann, B.; Tratwal, A.; Oberforster, M.; Wieser, F.; Schiemann, A.; Jørgensen, L.N.; et al. Ergot infection in winter rye hybrids shows differential contribution of male and female genotypes and environment. *Euphytica* **2020**, *216*, 1–14. [CrossRef]
22. Klotz, J.L. Activities and effects of ergot alkaloids on livestock physiology and production. *Toxins* **2015**, *7*, 2801–2821. [CrossRef] [PubMed]
23. Blaney, B.J.; Molloy, J.B.; Brock, I.J. Alkaloids in Australian rye ergot (Claviceps purpurea) sclerotia: Implication for food and stockfeed regulations. *Anim. Prod. Sci.* **2009**, *49*, 975–982. [CrossRef]
24. Miedaner, T.; Mirdita, V.; Rodemann, B.; Drobeck, T.; Rentel, D. Genetic variation of winter rye cultivars for their ergot (Claviceps purpurea) reaction tested in a field design with minimized interplot interference. *Plant Breed.* **2010**, *129*, 58–62. [CrossRef]
25. Geiger, H.H.; Miedaner, T. Genetic basis and phenotypic stability of male-fertility restoration in rye. *Vorträge Pflanz.* **1996**, 27–38.
26. Miedaner, T.; Wilde, P.; Wortmann, H. Combining ability of non-adapted sources for male-fertility restoration in Pampa CMS of hybrid rye. *Plant Breed.* **2004**, *124*, 39–43. [CrossRef]
27. Falke, K.C.; Wilde, P.; Miedaner, T. Rye introgression lines as source of alleles for pollen-fertility restoration in pampa cms. *Plant Breed.* **2009**, *128*, 528–531. [CrossRef]
28. Miedaner, T.; Herter, C.P.; Goßlau, H.; Wilde, P.; Hackauf, B.; Pillen, K. Correlated effects of exotic pollen-fertility restorer genes on agronomic and quality traits of hybrid rye. *Plant Breed.* **2017**, *136*, 224–229. [CrossRef]
29. KWS. PollenPlus. Available online: https://www.kws.com/gb/en/products/cereals/hybrid-rye/pollenplus-kws-files-for-ergot-patent-in-hybrid-rye/ (accessed on 30 June 2021).
30. Melz, G.; Melz, G.; Hartman, F. Genetics of a male-sterile rye of 'G-type' with results of the first F1-hybrids. *Plant Breed. Seed Sci.* **2003**, *47*, 47–55.
31. Gaborieau, L.; Brown, G.G.; Mireau, H. The propensity of pentatricopeptide repeat genes to evolve into restorers of cytoplasmic male sterility. *Front. Plant Sci.* **2016**, *7*, 1816. [CrossRef]
32. Gully, B.S.; Cowieson, N.; Stanley, W.A.; Shearston, K.; Small, I.D.; Barkan, A.; Bond, C.S. The solution structure of the pentatricopeptide repeat protein PPR10 upon binding atpH RNA. *Nucleic Acids Res.* **2015**, *43*, 1918–1926. [CrossRef]
33. Schmitz-Linneweber, C.; Small, I. Pentatricopeptide repeat proteins: A socket set for organelle gene expression. *Trends Plant Sci.* **2008**, *13*, 663–670. [CrossRef]
34. Ban, T.; Ke, J.; Chen, R.; Gu, X.; Tan, M.H.; Zhou, X.E.; Kang, Y.; Melcher, K.; Zhu, J.K.; Xu, H.E. Structure of a PLS-class pentatricopeptide repeat protein provides insights into mechanism of RNA recognition. *J. Biol. Chem.* **2013**, *288*, 31540–31548. [CrossRef] [PubMed]
35. Rizzolatti, C.; Bury, P.; Tatara, E.; Pin, P.A.; Rodde, N.; Bergès, H.; Budar, F.; Mireau, H.; Gielen, J.J.L. Map-based cloning of the fertility restoration locus Rfm1 in cultivated barley (Hordeum vulgare). *Euphytica* **2017**, *213*. [CrossRef]
36. 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]
37. Qi, W.; Yang, Y.; Feng, X.; Zhang, M.; Song, R. Mitochondrial function and maize kernel development requires Dek2, a pentatricopeptide repeat protein involved in nad1 mRNA splicing. *Genetics* **2017**, *205*, 239–249. [CrossRef] [PubMed]
38. Liu, Y.J.; Liu, X.; Chen, H.; Zheng, P.; Wang, W.; Wang, L.; Zhang, J.; Tu, J. A plastid-localized pentatricopeptide repeat protein is required for both pollen development and plant growth in rice. *Sci. Rep.* **2017**, *7*, 1–12. [CrossRef]
39. Melonek, J.; Duarte, J.; Martin, J.; Beuf, L.; Murigneux, A.; Varenne, P.; Comadran, J.; Specel, S.; Levadoux, S.; Bernath-Levin, K.; et al. The genetic basis of cytoplasmic male sterility and fertility restoration in wheat. *Nat. Commun.* **2021**, *12*, 1036. [CrossRef]
40. Klein, R.R.; Klein, P.E.; Mullet, J.E.; Minx, P.; Rooney, W.L.; Schertz, K.F. Fertility restorer locus Rf1 [corrected] of sorghum (Sorghum bicolor L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theor. Appl. Genet.* **2005**, *111*, 994–1012. [CrossRef]
41. Beick, S.; Schmitz-Linneweber, C.; Williams-Carrier, R.; Jensen, B.; Barkan, A. The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol. Cell. Biol.* **2008**, *28*, 5337–5547. [CrossRef] [PubMed]
42. Kazama, T.; Toriyama, K. A fertility restorer gene, Rf4, widely used for hybrid rice breeding encodes a pentatricopeptide repeat protein. *Rice* **2014**, *7*, 1–5. [CrossRef]

43. Hu, J.; Wang, K.; Huang, W.; Liu, G.; Gao, Y.; Wang, J.; Huang, Q.; Ji, Y.; Qin, X.; Wan, L.; et al. The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. *Plant Cell* **2012**, *24*, 109–122. [CrossRef]
44. Huang, W.; Yu, C.; Hu, J.; Wang, L.; Dan, Z.; Zhou, W.; He, C.; Zeng, Y.; Yao, G.; Qi, J.; et al. Pentatricopeptide-repeat family protein RF6 functions with hexokinase 6 to rescue rice cytoplasmic male sterility. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14984–14989. [CrossRef]
45. Barkan, A.; Small, I. Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **2014**, *65*, 415–442. [CrossRef]
46. Robles, P.; Quesada, V. Research progress in the molecular functions of plant mTERF proteins. *Cells* **2021**, *10*, 205. [CrossRef]
47. Bernhard, T.; Koch, M.; Snowdon, R.J.; Friedt, W.; Wittkop, B. Undesired fertility restoration in msm1 barley associates with two mTERF genes. *Theor. Appl. Genet.* **2019**, *132*, 1335–1350. [CrossRef]
48. Shahinnia, F.; Geyer, M.; Block, A.; Mohler, V.; Hartl, L. Identification of Rf9, a gene contributing to the genetic complexity of fertility restoration in hybrid wheat. *Front. Plant Sci.* **2020**, *11*, 1720. [CrossRef] [PubMed]
49. Wilde, P.; Korzun, V.; Menzel, J.; Ruonan, Z.; Stein, N.; Hackauf, B. Restorer Plants. KWS SAAT SE. U.S. Patent App. 16/064304, 9 May 2019.
50. Hackauf, B.; Korzun, V.; Wortmann, H.; Wilde, P.; Wehling, P. Development of conserved ortholog set markers linked to the restorer gene Rfp1 in rye. *Mol. Breed.* **2012**, *30*, 1507–1518. [CrossRef]
51. Börner, A.; Korzun, V.; Polley, A.; Malyshev, S.; Melz, G. Genetics and molecular mapping of a male fertility restoration locus (Rfg1) in rye (*Secale cereale* L.). *Theor. Appl. Genet.* **1998**, *97*, 99–102. [CrossRef]
52. 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]
53. Vilhjalmsson, B.J.; Nordborg, M. The nature of confounding in genome-wide association studies. *Nat. Rev. Genet.* **2013**, *14*, 1–2. [CrossRef]
54. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **2013**, *9*, 1–9. [CrossRef]
55. 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]
56. Zhang, Z.; Ersoz, E.; Lai, C.Q.; Todhunter, R.J.; Tiwari, H.K.; Gore, M.A.; Bradbury, P.J.; Yu, J.; Arnett, D.K.; Ordovas, J.M.; et al. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **2010**, *42*, 355–360. [CrossRef] [PubMed]
57. Geyer, M.; Albrecht, T.; Hartl, L.; Mohler, V. Exploring the genetics of fertility restoration controlled by Rf1 in common wheat (Triticum aestivum L.) using high-density linkage maps. *Mol. Genet. Genom.* **2018**, *293*, 451–462. [CrossRef]
58. Wurschum, T.; Leiser, W.L.; Weissmann, S.; Maurer, H.P. Genetic architecture of male fertility restoration of Triticum timopheevii cytoplasm and fine-mapping of the major restorer locus Rf3 on chromosome 1B. *Theor. Appl. Genet.* **2017**, *130*, 1253–1266. [CrossRef] [PubMed]
59. Wang, S.; Wong, D.; Forrest, K.; Allen, A.; Chao, S.; Huang, B.E.; Maccaferri, M.; Salvi, S.; Milner, S.G.; Cattivelli, L.; et al. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **2014**, *12*, 787–796. [CrossRef] [PubMed]
60. Stojalowski, S.A.; Milczarski, P.; Hanek, M.; Bolibok-Bragoszewska, H.; Myskow, B.; Kilian, A.; Rakoczy-Trojanowska, M. DArT markers tightly linked with the Rfc1 gene controlling restoration of male fertility in the CMS-C system in cultivated rye (*Secale cereale* L.). *J. Appl. Genet.* **2011**, *52*, 313–318. [CrossRef]
61. Maan, S.S.; Luchen, K.A.; Bravo, J.M. Genetic analyses of male-fertility restoration in wheat. I. Chromosomal location of Rf genes 1. *Crop. Sci.* **1984**, *24*, 17–20. [CrossRef]
62. Ma, Z.Q.; Zhao, Y.H.; Sorrells, M.E. Inheritance and chromosomal locations of male fertility restoring gene transferred from Aegilops umbellulata Zhuk. to *Triticum aestivum* L. *Mol. Gen. Genet. MGG* **1995**, *247*, 351–357. [CrossRef]
63. Du, H.; Maan, S.S.; Hammond, J.J. Genetic analysis of male-fertility restoration in wheat: III. Effects of aneuploidy. *Crop. Sci.* **1991**, *31*, 319–322. [CrossRef]
64. Ui, H.; Sameri, M.; Pourkheirandish, M.; Chang, M.C.; Shimada, H.; Stein, N.; Komatsuda, T.; Handa, H. High-resolution genetic mapping and physical map construction for the fertility restorer Rfm1 locus in barley. *Theor. Appl. Genet.* **2015**, *128*, 283–290. [CrossRef]
65. Martis, M.M.; Zhou, R.; Haseneyer, G.; Schmutzer, T.; Vrana, J.; Kubalakova, M.; Konig, S.; Kugler, K.G.; Scholz, U.; Hackauf, B.; et al. Reticulate evolution of the rye genome. *Plant Cell* **2013**, *25*, 3685–3698. [CrossRef]
66. Itabashi, E.; Iwata, N.; Fujii, S.; Kazama, T.; Toriyama, K. The fertility restorer gene, Rf2, for lead rice-type cytoplasmic male sterility of rice encodes a mitochondrial glycine-rich protein. *Plant J.* **2011**, *65*, 359–367. [CrossRef]
67. Fujii, S.; Toriyama, K. Suppressed expression of retrogrede-regulated male sterility restores pollen fertility in cytoplasmic male sterile rice plants. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9513–9518. [CrossRef]
68. Liu, F.; Schnable, P.S. Functional specialization of maize mitochondrial aldehyde dehydrogenases. *Plant Physiol.* **2002**, *130*, 1657–1674. [CrossRef]
69. Jaqueth, J.S.; Hou, Z.; Zheng, P.; Ren, R.; Nagel, B.A.; Cutter, G.; Niu, X.; Vollbrecht, E.; Greene, T.W.; Kumpatla, S.P. Fertility restoration of maize CMS-C altered by a single amino acid substitution within the Rf4 bHLH transcription factor. *Plant J.* **2020**, *101*, 101–111. [CrossRef]

70. Kitazaki, K.; Arakawa, T.; Matsunaga, M.; Yui-Kurino, R.; Matsuhira, H.; Mikami, T.; Kubo, T. Post-translational mechanisms are associated with fertility restoration of cytoplasmic male sterility in sugar beet (*Beta vulgaris*). *Plant J.* **2015**, *83*, 290–299. [CrossRef] [PubMed]
71. Luptovčiak, I.; Samakovli, D.; Komis, G.; Šamaj, J. KATANIN1 is essential for embryogenesis and seed formation in Arabidopsis. *Front. Plant Sci.* **2017**, *8*, 728. [CrossRef] [PubMed]
72. Schreiber, D.N.; Bantin, J.; Dresselhaus, T. The MADS box transcription factor ZmMADS2 is required for anther and pollen maturation in maize and accumulates in apoptotic bodies during anther dehiscence. *Plant Physiol.* **2004**, *134*, 1069–1079. [CrossRef] [PubMed]
73. Zhou, L. A MADS-box transcription factor related to fertility conversion in male sterile wheat lines. *Acta Agron. Sin.* **2008**, *34*, 598. [CrossRef]
74. Linke, B.; Nothnagel, T.; Börner, T. Flower development in carrot CMS plants: Mitochondria affect the expression of MADS box genes homologous to GLOBOSA and DEFICIENS. *Plant J.* **2003**, *34*, 27–37. [CrossRef]
75. Zou, M.; Guan, Y.; Ren, H.; Zhang, F.; Chen, F. A bZIP transcription factor, OsABI5, is involved in rice fertility and stress tolerance. *Plant Mol. Biol.* **2008**, *66*, 675–683. [CrossRef]
76. Pan, Z.; Ren, X.; Zhao, H.; Liu, L.; Tan, Z.; Qiu, F. A mitochondrial transcription termination factor, ZmSmk3, is required for nad1 intron4 and nad4 intron1 splicing and kernel development in maize. *G3 Genes Genomes Genet.* **2019**, *9*, 2677–2686. [CrossRef]
77. Lapinski, M.; Stojalowski, S. The C-source of sterility-inducing cytoplasm in rye: Origin, identity and occurrence. *Vorträge Pflanz.* **1996**, *35*, 51–60.
78. Warsecha, R.; Salak-Warzecha, K. Comparative studies on CMS sources in rye. *Vorträge Pflanz.* **1996**, *35*, 39–49.
79. 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 (accessed on 7 July 2021).
80. 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. *Proc. Tenth Int. Wheat Genet. Symp.* **2003**, 789–791.
81. McHugh, M.L. The chi-square test of independence. *Biochem. Med.* **2013**, *23*, 143–149. [CrossRef]
82. Haseneyer, G.; Schmutzer, T.; Seidel, M.; Zhou, R.; Mascher, M.; Schon, C.C.; Taudien, S.; Scholz, U.; Stein, N.; Mayer, K.F.; et al. From RNA-seq to large-scale genotyping-genomics resources for rye (*Secale cereale* L.). *BMC Plant Biol.* **2011**, *11*, 131. [CrossRef] [PubMed]
83. 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]
84. NCBI. National Center for Biotechnology Information. Available online: https://www.ncbi.nlm.nih.gov (accessed on 13 June 2021).
85. Clayton, D. snpStats: SnpMatrix and XSnpMatrix Classes and Methods. R Package Version 1.36.0. 2019. Available online: https://rdrr.io/bioc/snpStats/ (accessed on 7 July 2021).
86. Rstudio Team. Rstudio: Integrated development for R. RStudio, Inc., Boston. 2015. Available online: http://www.rstudio.com (accessed on 15 June 2021).
87. R Core Team. *R a Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: https://www.R-project.org/ (accessed on 13 June 2021).
88. Granato, I.S.C.; Galli, G.; de Oliveira Couto, E.G.; e Souza, M.B.; Mendonça, L.F.; Fritsche-Neto, R. snpReady: A tool to assist breeders in genomic analysis. *Mol. Breed.* **2018**, *38*, 102. [CrossRef]
89. Mahmood, K.; Orabi, J.; Kristensen, P.S.; Sarup, P.; Jorgensen, L.N.; Jahoor, A. De novo transcriptome assembly, functional annotation, and expression profiling of rye (*Secale cereale* L.) hybrids inoculated with ergot (*Claviceps purpurea*). *Sci. Rep.* **2020**, *10*, 13475. [CrossRef] [PubMed]

Genomic can of male fertility restoration genes in a 'Gülzow' type hybrid breeding system of rye (*Secale cereale* L.)

Nikolaj M. Vendelbo, Khalid Mahmood , Pernille M. Sarup,
Peter S. Kristensen, Jihad Orabi, and Ahmed Jahoor

Supplementary material

Caption of all supplementary files has been given below and a subset of supplementary figures and tables included in the dissertation. The extended supplementary materials has been provided at an open-access Zenodo data repository (https://zenodo.org/record/5510286).

Table S1. Chromosomal distribution and polymorphism information content (PIC) of informative single nucleotide polymorphism (SNP) markers deriving from three genotyping platforms on Nordic Seed hybrid rye (*Secale cereale* L.) breeding germplasm (20K, 600K) and F_2 biparental mapping population (30K).

Genotyping platform	20K	30K	600K
lines	365	181	180
Chromosome	Markers		
1R	590	332	33854
2R	711	395	33698
3R	631	386	31493
4R	515	299	32555
5R	669	374	37073
6R	516	322	36872
7R	528	340	38918
Unknown	259	1045	16943
Total	4419	3493	261406
PIC	0.26	0.34	0.23

Table S2: Top-most associated SNP markers in case control GWAS on Nordic Seed hybrid rye (*Secale cereale* L.) breeding germplasm (n = 365) using 4419 informative SNP markers from a 20K SNP array

marker_id	Chromosome	Position (Mbp)	-log10(p)	F_st	-log10(p) Biparental Mapping Population (30K, n = 181)	-log10(p) Case Control (600K, n = 180)
c9229_3272	3R	656,97	7,96	0,9	1,68	3,87
c9041_959	3R	659,47	10,59	1,0	0,94	13,13
c9041_915	3R	659,47	10,59	1,0	0,94	13,13
AX_99687383	3R	724,45	15,87	1,0	2,34	22,57
AX_99754750	3R	724,46	15,87	1,0	1,70	22,57
c10717_796	3R	726,61	13,96	1,0	2,34	22,57
c12145_1411	3R	728,04	6,91	1,0	1,88	NA
c11210_1125	3R	745,44	19,08	1,0	2,13	4,96
AX_99878275	4R	884,68	5,43	1,0	0,48	NA

S3 Table: Top associated SNP markers on each chromosome in case control GWAS on Nordic Seed hybrid rye (*Secale cereale* L.) breeding germplasm (n=180) using 261406 informative SNP markers from a 600K array

Table S4: Restoration of male-fertility associated SNP markers in rye (*Secale cereale* L.) F_2 mapping population (n = 181) originating from a Gülzow type hybrid rye cv. Stannos. Population was genotyped using 30K SNP array.

Marker ID	Chromosome	Position (Mbp)	Chromosome (Wheat 90K)	Position (cM, Wheat - 90K)	-log10(p)									
					Case control	Pollen production	Seed Number	Spike Number	Seed per spike	TKW	Total Grain Weight	Cumulative	Case Controle (n = 365, 20K)	Case Controle (n = 180, 600K)
RAC875_c14529_815	3R	626,81	NA	NA	7,78	6,22	6,66	1,60	6,76	6,16	6,12	39,70	NA	NA
wsnp_Ex_c2502_4675968	3R	745,45	NA	NA	3,29	3,17	3,23	0,20	2,99	2,06	3,21	17,94	0,56	1,35
c6046_935	3R	745,95	NA	NA	3,50	3,59	3,09	0,09	3,14	2,05	3,05	18,41	1,63	0,10
c5149_346	3R	747,04	NA	NA	3,35	4,09	3,52	0,19	3,13	1,98	3,39	19,46	0,70	NA
Excalibur_c5298_171	3R	769,35	NA	NA	4,75	N/A	3,33	0,00	5,20	4,25	3,49	21,03	NA	NA
c92_747	3R	798,24	NA	NA	2,62	4,04	4,28	0,38	3,71	2,06	4,24	20,96	NA	NA
AX_99346444	3R	807,11	NA	NA	3,32	4,68	4,37	0,35	4,19	2,51	4,28	23,35	0,58	0,39
AX_99840974	3R	808,70	NA	NA	3,32	4,68	4,37	0,35	4,19	2,51	4,28	23,35	0,76	3,13
c7821_382	3R	808,70	NA	NA	3,32	4,68	4,37	0,35	4,19	2,51	4,28	23,35	4,79	5,58
c15932_1173	3R	808,70	NA	NA	2,75	4,44	4,22	0,26	3,90	2,05	4,16	21,52	4,79	5,07
AX_158569168	Unknown	NA	1A	66,1	5,55	3,06	3,78	0,51	4,20	4,16	3,37	24,13	NA	NA
AX_158555391	Unknown	NA	1A	87,80	6,16	3,87	6,36	0,73	6,13	4,79	5,79	33,09	NA	NA
AX_158538397	Unknown	NA	3B	47,10	4,89	3,89	3,94	0,51	3,82	4,26	3,67	24,46	NA	NA
AX_158558079	Unknown	NA	3B	49,20	10,97	7,12	9,12	0,69	10,95	9,91	10,05	58,12	NA	NA
RFL_Contig1291_309	Unknown	NA	3B	60,52	4,87	4,18	7,00	0,68	6,49	3,87	6,43	32,85	NA	NA
Kukri_c47935_137	Unknown	NA	NA	NA	9,33	5,67	7,62	0,33	8,91	8,26	7,66	47,45	NA	NA

S5 Table: Genome-wide pairwise linkage disequilibrium in a rye (*Secale cereale* L.) F2 mapping population, comprised of 181 plants genotyped on 30K SNP array, between unmapped restoration of male fertility associated markers derived from 90K wheat SNP array and 2448 informative SNP markers with accurate mapping position on 'Lo7' rye reference genome

S6 Table: Genotype of SNP markers in 181 rye (*Secale cereale* L.) F_2 plants at region on chromosome arm 3RL associated with restoration of male fertility, **A**) all markers in the region, **B**) significant associated wheat markers

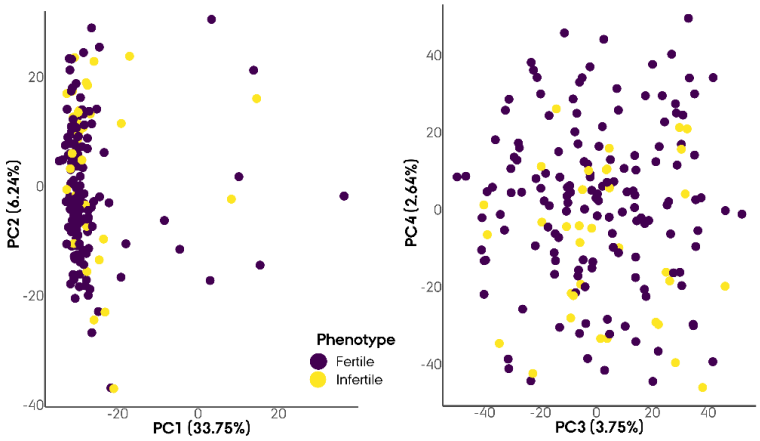
S7 Table: Gene annotation in 'Lo7' rye (*Secale cereale* L.) reference genome and gene expression in Gülzow type hybrids during flowering at region associated with restoration of male fertility on chromosome arm **A**) 3RL, and **B**) 1RS

S8 Table: Annotated **A)** pentatricopeptide repeat (PPR) proteins and restoration of male fertility-like PPR (RFL-PPR), and **B**) mitochondrial transcription termination factor (mTERF) proteins in 'Lo7' rye (*Secale cereale* L.) reference genome

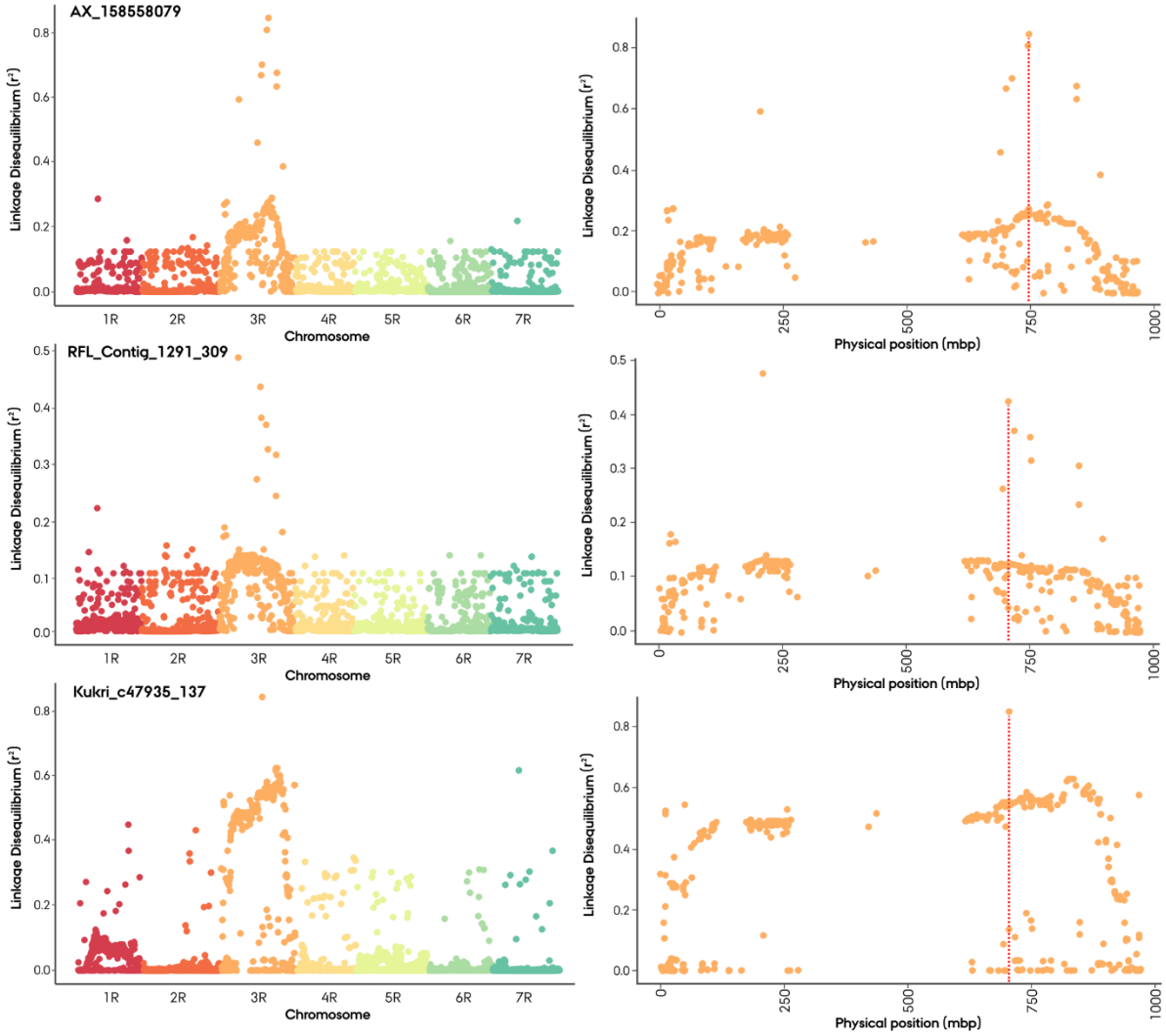
Table S9: Pairwise linkage disequilibrium of SNP markers associated with restoration of male-fertility on chromosome arm 3RL in Gülzow type hybrid rye (*Secale cereale* L.) breeding germplasm.

Marker ID	Origin	Lo7		Marker ID									
		Chromosome	Position (Mbp)	AX_99346444	AX_99840974	c7821_382	c15932_1173	AX_158569168	AX_158555391	AX_158538397	AX_158558079	RFL_Contig1291_309	Kukri_c47935_137
				3R				Unknown					
				807,11	808,70	808,70	808,70	NA	NA	NA	NA	NA	NA
c92_747	5K	3R	798,24	0,98	0,98	0,98	0,99	0,24	0,21	0,14	0,24	0,10	0,52
AX_99346444	600K	3R	807,11		1,00	1,00	0,99	0,25	0,23	0,16	0,24	0,11	0,54
AX_99840974	600K	3R	808,70			1,00	0,99	0,25	0,23	0,16	0,24	0,11	0,54
c7821_382	5K	3R	808,70				0,99	0,25	0,23	0,16	0,24	0,11	0,54
c15932_1173	5K	3R	808,70					0,25	0,22	0,15	0,24	0,10	0,54

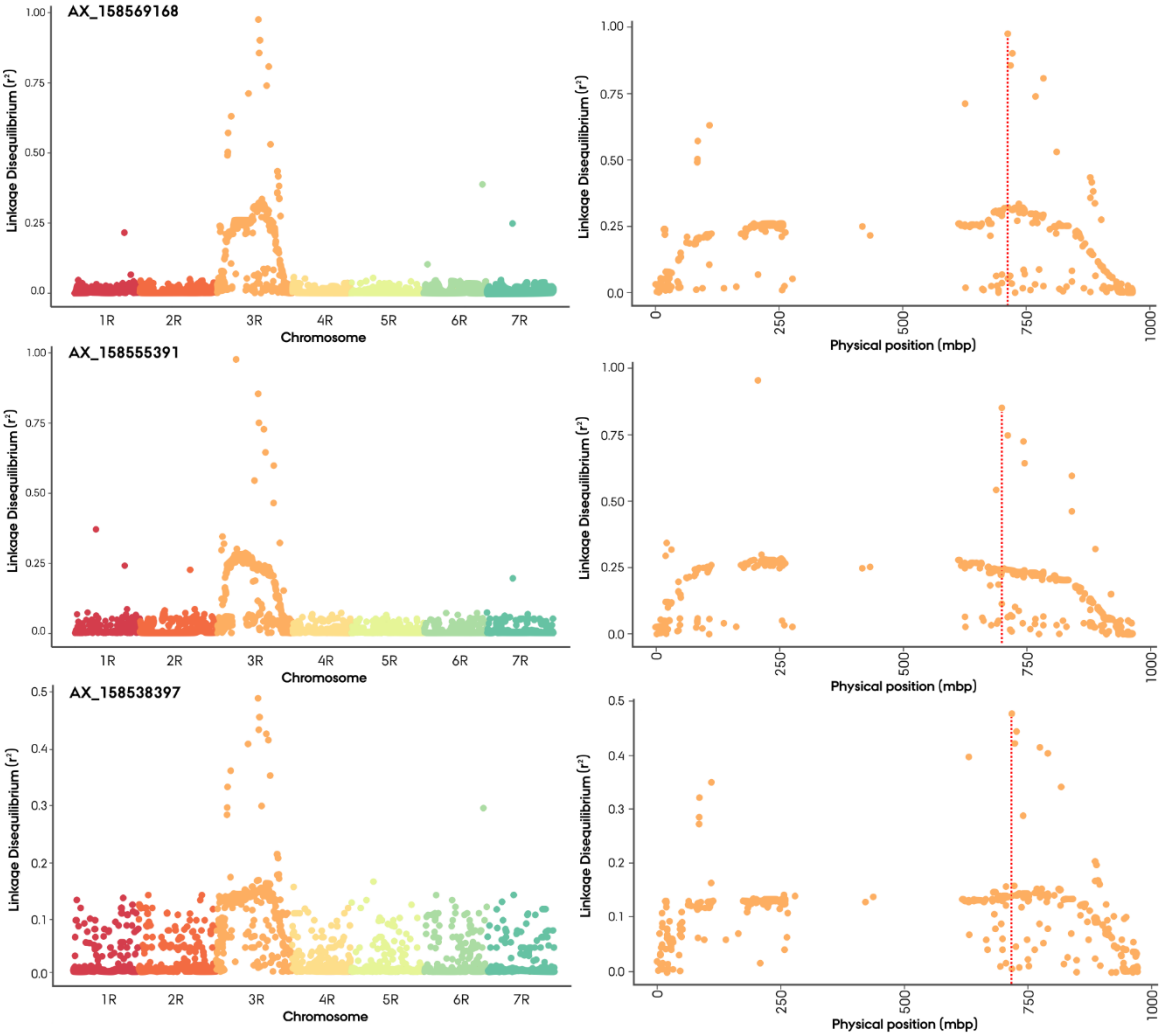
SFig. 1 Principal component analysis (PCA) on F_2 biparental mapping population (n = 181) using 3493 informative SNP markers



SFig. 2 Mapping of restoration of male-fertility associated markers derived from the 90K wheat SNP array by pairwise linkage disequilibrium towards 2448 informative SNP markers with accurate mapping position on the 'Lo7' rye (*Secale cereale* L.) reference genome.



SFig. 2: Continued



Manuscript III

Discovery of a novel leaf rust (*Puccinia recondita*) resistance gene in rye (*Secale cereale* L.) using association genomics

By

Nikolaj, M. Vendelbo[1,2*], Khalid Mahmood[1], Pernille M. Sarup[1], Mogens Hovmøller[2], Annemarie F. Justesen[2], Peter S. Kristensen[1], Jihad Orabi[1], Ahmed Jahoor[1,3]

[1] Nordic Seed A/S, Odder, Denmark

[2] Department of Agroecology, Aarhus University, Slagelse, Denmark

[3] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

Publication details

Published in MDPI Cell

cells

MDPI

Article

Discovery of a Novel Leaf Rust (*Puccinia recondita*) Resistance Gene in Rye (*Secale cereale* L.) Using Association Genomics

Nikolaj Meisner Vendelbo [1,2,*], Khalid Mahmood [1], Pernille Sarup [1], Mogens S. Hovmøller [2], Annemarie Fejer Justesen [2], Peter Skov Kristensen [1], Jihad Orabi [1] and Ahmed Jahoor [1,3]

[1] Nordic Seed A/S, Grindsnabevej 25, 8300 Odder, Denmark; khma@nordicsee.com (K.M.); pesa@nordicseed.com (P.S.); pskr@nordicseed.com (P.S.K.); jior@nordicseed.com (J.O.); ahja@nordicseed.com (A.J.)
[2] Department of Agroecology, Faculty of Technical Sciences, Aarhus University, 4200 Slagelse, Denmark; mogens.hovmoller@agro.au.dk (M.S.H.); annemariefejer.justesen@agro.au.dk (A.F.J.)
[3] Department of Plant Breeding, The Swedish University of Agricultural Sciences, 23053 Alnarp, Sweden
* Correspondence: nive@nordicseed.com

Abstract: Leaf rust constitutes one of the most important foliar diseases in rye (*Secale cereale* L.). To discover new sources of resistance, we phenotyped 180 lines belonging to a less well-characterized Gülzow germplasm at three field trial locations in Denmark and Northern Germany in 2018 and 2019. We observed lines with high leaf rust resistance efficacy at all locations in both years. A genome-wide association study using 261,406 informative single-nucleotide polymorphisms revealed two genomic regions associated with resistance on chromosome arms 1RS and 7RS, respectively. The most resistance-associated marker on chromosome arm 1RS physically co-localized with molecular markers delimiting *Pr3*. In the reference genomes Lo7 and Weining, the genomic region associated with resistance on chromosome arm 7RS contained a large number of nucleotide-binding leucine-rich repeat (NLR) genes. Residing in close proximity to the most resistance-associated marker, we identified a cluster of NLRs exhibiting close protein sequence similarity with the wheat leaf rust *Lr1* gene situated on chromosome arm 5DL in wheat, which is syntenic to chromosome arm 7RS in rye. Due to the close proximity to the most resistance-associated marker, our findings suggest that the considered leaf rust *R* gene, provisionally denoted *Pr6*, could be a *Lr1* ortholog in rye.

Keywords: hybrid breeding; Gülzow germplasm; field trial; disease progression; 600K high-density SNP array; nucleotide-binding leucine-rich repeat (NLR); NB-ARC domain; phylogenetic analysis; *Lr1*-like disease resistance protein

Citation: Vendelbo, N.M.; Mahmood, K.; Sarup, P.; Hovmøller, M.S.; Justesen, A.F.; Kristensen, P.S.; Orabi, J.; Jahoor, A. Discovery of a Novel Leaf Rust (*Puccinia recondita*) Resistance Gene in Rye (*Secale cereale* L.) Using Association Genomics. *Cells* **2022**, *11*, 64. https://doi.org/10.3390/cells11010064

Academic Editors: Francesco Mercati and Francesco Carimi

Received: 23 November 2021
Accepted: 18 December 2021
Published: 27 December 2021

1. Introduction

Winter rye (*Secale cereale* L.) is an important cereal crop in Northern and Eastern Europe, with an acreage of 903,800 ha in Poland, 636,300 ha in Germany, 146,600 ha in Denmark and 113,200 ha in the Baltic countries in 2017 [1]. Since the introduction of hybrids in the mid-1980s, the primary focus of resistance breeding in rye shifted towards ergot caused by *Claviceps purpurea* (Fr.) Tul [2,3]. Recent decades of breeding efforts and the introduction of the exotic restoration of male fertility genes have, however, reduced the importance of ergot and the resulting toxic alkaloids in rye grains [4,5]. This provides an opportunity to focus on leaf rust, also termed brown rust, which has received little scientific attention and commercial interest in recent years. In 2002, Miedaner et al. [6] reported a low level of inherent resistance to leaf rust in the predominant hybrid rye breeding gene pools Petkus and Carsten. This is still the case based on observations of top-yielding hybrid cultivars and population varieties in the Danish official trials, which showed 9.5% mean leaf rust severity in 2019 (Table S2 [7]). Similarly, the German list of recommended varieties of rye had 11.5% mean leaf rust severity in 2021 (Table S2 [8]). Under moderate to severe

infection levels, leaf rust has been reported to cause a 11–27% reduction in grain yield, in addition to considerable quality losses [9].

The cereal rusts are heteroecious, requiring two taxonomically unrelated host plants to complete their life cycle, and macrocyclic, entailing five distinct spore types corresponding to different life stages [10–12]. During the growing season, rusts can cause severe epidemics through repetitive infection and clonal reproduction of asexual urediniospores [13,14]. As obligate biotrophs, rusts require a living host for reproduction, but it may survive parts of the year as telia on rye plant debris, or as aecia on the alternate host [15]. In years with conditions conducive to leaf rust, the successive autumn-sown winter rye crop can likewise be infected, leading to an early establishment [16].

Cereal rusts are capable of migrating long distances by wind dissemination of urediniospores [17–20], which may result in exotic incursions of new races from distant areas [21]. In wheat, two new races of yellow rust, 'Warrior' (Pst7) and 'Kranich' (Pst8), first detected in Europe in 2011, are believed to have migrated from a sexual population in the near-Himalayan region of Asia [22]. Sexual recombination in rust fungi facilitates the development of new virulence combinations, driving adaptation to deployed host resistance and the emergence of novel aggressive pathotypes [22,23]. Population studies on wheat leaf and yellow rust populations, however, reveal clonal population structures and the absence of sexual recombination in many areas [24–27]. In such populations, novel genetic variation may be driven by means of mutations, somatic hybridization and internuclear exchange [23,28].

In rye, leaf rust is caused by a fungal basidiomycete *Puccinia recondita* f. sp. *secalis* Roberge. ex Desmaz. (*Prs*) [29]. Unlike the alternate hosts for the rusts of wheat, the alternate host of *Prs*, small bugloss (*Anchusa arvensis* L.), is widespread in the North European flora, being a common weed in agricultural fields [12,30–32]. Observation of high pathotype diversity and virulence complexity in the German *Prs* population may suggest the occurrence of sexual recombination in rye leaf rust [33], which may facilitate *Prs* to overcome deployed resistance (*R*) genes [23,33].

Currently, five major leaf rust *R* genes have been identified in rye, *Pr3* (1RS), *Pr4* (1RL), *Pr5* (1RL), *Pr1* (6RL) and *Pr2* (7RL) [34,35]. An additional three major leaf rust *R* genes have been identified in wheat–rye substitution and translocation lines, in the wheat gene nomenclature denoted *Lr26* (1BL-1RS), *Lr25* (4BS.4BL-2RL) and *Lr45* (2AS-2RS.2RL) [36,37].

Most major *R* genes belong to a large family of nucleotide-binding leucine-rich repeat proteins (NLR) [38]. In rye, 1,167 NLR genes have been identified in the Lo7 reference genome and 1,447 NLR genes in the Weining reference genome [39,40]. In grasses, the canonical NLR gene consists of three domains: a C-terminus leucine-rich repeat (LRR) domain, involved in pathogen effector recognition [41], a central nucleotide-binding (NB) domain functioning as a regulatory domain determining protein activation state [42] and an N-terminus coiled-coil (CC) domain believed to be involved in signaling and the induction of cell death [43].

Genomic-based breeding techniques have accelerated the introgression and pyramiding of *R* genes for enhancing resistance durability [44–46]. Recent advances in genomic resources available in rye, including the 600K high-density SNP array and chromosomal-scale reference genomes of a German inbred winter rye line Lo7, and a Chinese population rye variety Weining, respectively, constitute significant milestones in rye genomic breeding [39,40,47]. In order to expand the 'toolbox' available for resistance breeding in rye, continuous mining for the discovery of novel genetic variability in *R* genes is essential.

In this paper, we investigate leaf rust resistance in a Gülzow-based elite hybrid rye breeding germplasm. The Gülzow germplasm is differentiated from the predominant hybrid rye breeding gene pools, Petkus and Carsten, by a distinct cytoplasmic male sterility system [48,49]. Our objectives were to (I) characterize leaf rust resistance and disease progression in the assayed germplasm, (II) identify genomic regions and molecular markers associated with leaf rust resistance to facilitate marker-assisted selection for leaf rust resistance and (III) investigate in silico whether NLR genes residing in leaf rust resistance-

associated regions on the Lo7 and Weining reference genomes resemble known leaf rust *R* genes.

2. Materials and Methods

2.1. Plant Material and DNA Extraction

A panel of 180 inbred rye (*Secale cereale* L.) lines, 92 restorer and 88 non-restorer germplasms, belonging to the Gülzow-based elite hybrid rye breeding germplasm at Nordic Seed A/S (Dyngby, Denmark), were investigated in this study. Population structure and information on the genetic characteristics of the accessions were investigated in a recent study by Vendelbo et al. [50]. DNA extraction was done using an adapted SDS-based method according to USDA [51], after Pallotta et al. [52], on an equivalent of 75 mg plant material collected from the primary leaves of two seven-day old seedlings per line. DNA concentration and 260/280 nm absorption ratio of samples were measured using an Epoch™ microplate spectrophotometer (Biotek®, Santa Clara, CA, USA) and evidence of fragmentation by size visualization on 1.2% agarose gel.

2.2. Molecular Marker Resource and SNP Genotyping

Samples of each line containing 200 ng high-molecular-weight gDNA with ≥1.8 260/280nm ratio were sent for single-nucleotide polymorphism (SNP) genotyping at Eurofins Genomics Europe Genotyping (Aarhus, Denmark). Genotyping was done using a 600K SNP array with 600,843 SNP markers on a Affymetrix GeneTitan™ Scanner platform (Thermo Fisher Scientific Inc., Waltham, MA, USA) [47]. Marker map for the 600K SNP array on the Lo7 reference genome and evaluation of its performance in the assayed germplasm was recently presented in a study by Vendelbo et al. [53]. The marker map was acquired from https://doi.org/10.5281/ZENODO.5086235 (Access date: 9 July 2021). Prior to analysis, markers were filtered for marker allele frequency ≥0.05, missing individual score ≤ 0.2 and missing marker score ≤ 0.1 to identify informative markers.

2.3. Collection of Puccinia Recondite f. sp. secalis Populations

To establish a representative bulk inoculum of *Puccinia recondita* f. sp. *secalis* (*Prs*) for inoculation of field trials, leaf samples of rye in the field were collected from four locations in the period May-July 2018 in Denmark and Northern Germany. In Denmark, the locations were Dyngby in Jutland (55.9479° N, 10.2572° E), Flakkebjerg on Zealand (55.3255° N, 11.3826° E) and Holeby on Lolland (54.6998° N, 11.4511° E). In Northern Germany, the location was Nienstädt (52.3451° N, 9.1664° E). Multiplication of *Prs* was done at the Global Rust Reference Centre (GRRC), Institute of Agroecology, Aarhus University (Flakkebjerg, Denmark). The detailed protocol has been provided at https://doi.org/10.5281/zenodo.5478060 (Access date: 9 July 2021). Spores from each leaf sample were multiplied individually, resulting in 10 'unique' *Prs* field populations from each of the four sampled locations. For multiplication of *Prs*, a 50:50 mixture of hybrid rye cvs. KWS Bono and KWS Florano was selected on the basis of high leaf rust susceptibility reported in the Danish official trials from 2015 to 2018 [7].

2.4. Field Trial

Field trials were conducted at three locations, two situated in Denmark at Gylling in Central Jutland (55.8946° N, 10.1705° E) and Flakkebjerg on Zealand (55.3216° N, 11.3901° E), and one located in Northern Germany at Nienstädt in Niedersachsen (52.3556° N, 9.2270° E). Trials were sown in a Seedmatic® layout with a single parcel (1.0 m × 1.25 m) consisting of six rows of approximately 25 plants per breeding line, with a between-row distance of 25 cm and between-parcel distance of 40 cm. Each block comprised 18 parcels, with three blocks per replicate, of which the first was laid down in numerical order and the second in an incomplete randomized block design [54]. Hybrid cv. KWS Serafino was included as a 'resistant' control, and cv. KWS Binntto as 'susceptible', selected on the basis of disease severity recordings in the Danish official trials from 2015 to 2018 [7]. At Nienstädt

and Flakkebjerg, a leaf-rust-susceptible spreader row consisting of 50:50 hybrid cvs. KWS Binntto and KWS Serafino was sown as a spreader row at 2nd and 4th row position in each parcel to facilitate artificial inoculation. The trial site at Gylling was not artificially inoculated and no spreader rows were sown here. The Gylling trial site was sown on the 18th of September 2018, Flakkebjerg the 25th of September and Nienstädt the 4th of October in 2018. The trial at Nienstädt was repeated and sown on the 13th of October 2019.

2.5. Field Inoculation

Flakkebjerg and Nienstädt were artificially inoculated using seedlings with sporulating rust that had been inoculated with *Prs* field populations collected in 2018. The Gylling trial site in 2019 was kept for natural infection to study the progression of leaf rust disease under non-inoculated conditions. For each of the two sites, four sowing trays were prepared, each containing 35 multiplication pots. Multiplication of field trial inoculum was done according to the protocol provided at https://doi.org/10.5281/zenodo.5478060 (Access date: 9 July 2021). Inoculation of trays was done using an airbrush compression system in accordance with Thach et al. [27]. A spore inoculum solution was prepared for each of the trials in Denmark by mixing one tube, containing ≈ 6–7 mg spores, for each of the 30 *Prs* field populations collected at the three Danish locations in 2018 in a single tube by shaking for 60 s. The spore inoculum was split into four equal portions, one per tray, of approximately 50 mg, transferred to a 50 mL airbrush glass container and suspended in 4 mL 3MTM NovecTM 7100 engineering fluid. The inoculum solution for the trial in Northern Germany was prepared following the same procedure, using the *Prs* populations collected at Nienstädt. At 14 DAI, rusts were sporulating on the seedlings and trials were inoculated by brushing a single multiplication pot over three consecutive spreader rows. Three days later, multiplication pots were brushed in a similar manner across the same spreader rows in the opposite direction, plastic pots removed and seedlings planted. Trials were artificially inoculated in April in both years, following the same procedure to ensure uniform disease pressure.

2.6. Disease Scoring

Trials were scored for leaf rust severity a minimum of two times using a 1–9 scoring scale (Supplementary Table S1, [55]). The lines were scored by evaluating all plants of the individual breeding line within the plot. The experiment at Gylling, Denmark was scored four times in 2019 from the first detection of leaf rust in May until crop senescence at the end of June to study the disease progression under untreated conditions.

2.7. Analysis of Disease Scoring Data

Area under disease progress curve (AUDPC) was calculated at the non-treated site, Gylling, Denmark, using the agrocolae (v. 1.3–5) package in R [56]. For the purpose of interpretation, disease severity scale was adjusted to calculate a corrected AUDPC starting from zero (completely resistant with no evidence of chlorosis). Based on AUDPC and the disease progression curve at Gylling, lines were placed into four groups in order to characterize the assayed germplasms' qualitative and quantitative resistance to leaf rust. The groups comprised (i) resistant ('R'), with an AUDPC less than 20, (ii) partially resistant ('P–R'), with an AUDPC of 20–80, (iii) partially susceptible ('P–S') and (iv) susceptible ('S'), with an AUDPC between 70 and 170. While the AUDPC intervals of groups overlapped, each was distinguished by their disease progression; breeding lines assigned to (i) 'R' had a near-linear curve with disease severity less than 2, (ii) 'PS' had a stable inclining curve with a terminal severity score less than 6, (iii) 'P–S' had a late occurrence with a terminal severity score between 6 and 8, and (iv) 'S' had an early occurrence with a terminal severity score between 6 and 8. A pairwise t-test using a standard Bonferroni correction for multiple comparison was done on the four groups' AUDPC scores using inherent functions in R.

In order to correct the resistance phenotype for effects of replicate, block position, population, location-year and $G \times E$ interaction effect, four linear mixed models were

constructed using the lme4 (v. 1.1.26) package in R. The models were used for (1) individual locations per parental population, (2) individual locations using the entire assayed germplasm, (3) all locations per parental population and (4) all locations using the entire assayed germplasm

$$y = u + b + r + G + \varepsilon$$

$$y = u + b + r + P + G + \varepsilon$$

$$y = u + b + r + G + E + G \times E + \varepsilon$$

$$y = u + b + r + P + G + E + G \times E + \varepsilon$$

where μ is the general mean, b is the block, r is the replicate, P is the population, G is the line id, E is the location-year, $G \times E$ is the genotype, i.e., line and environment interaction effect, and ε is the residuals. The parameters b, r and $G \times E$ were set as random effects and P, E and G were set as fixed effects. The random effects and residuals were assumed to be independent, normally distributed variables described as follows: b~N (0, I σ_b^2), r~N (0, I σ_r^2), $G \times E$~N (0, I $\sigma_{G\times E}^2$) and ε~N (0, I σ_ε^2). The best linear unbiased estimator (BLUE) for the line effect, referred to as the resistance value, was used as phenotypic input for GWAS.

The broad sense heritability per plot (H^2) was extracted using a modified model no. 4 with line set as random effect, distributed G~N (0, I σ_G^2) and estimated as

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_b^2 + \sigma_r^2 + \sigma_{G\times E}^2 + \sigma_\varepsilon^2}$$

where σ_l^2 is line variance, σ_b^2 is block variance, σ_r^2 is replicate variance, $\sigma_{G\times E}^2$ is genotype–environment interaction variance, and σ_ε^2 is the residual variance.

2.8. Genome-Wide Association Study

Discovery of leaf rust resistance-associated SNP markers was done by genome-wide association study (GWAS) using the genomic association and prediction integration tool (GAPIT) (v.3) package in R [57]. The Manhattan plot was colorized using the RColorBrewer (v.1.1-2) R package colour palette [58]. GWAS was done using both the mixed linear model (MLM) and the Bayesian information and Linkage Disequilibrium Iteratively Nested Keyway (BLINK) method [59]. BLINK uses a multiple loci test for MLM by combining a fixed effects model, Bayesian information content and linkage disequilibrium information, collectively improving the statistical power in the GWAS. Markers that are in linkage disequilibrium with the most significant marker at a site are excluded in BLINK. A standard Bonferroni-corrected threshold of $\alpha = 0.05$ was used as the significance threshold. To investigate whether identified resistance-associated SNP markers on chromosome arm 1RS resided in proximity to known *Pr* genes (*Pr3*, *Pr4*, *Pr5*), flanking co-segregating markers were extracted and anchored to the Lo7 and Weining reference genomes using the NCBI blast function [35,60,61].

2.9. Phylogenetic Analysis of Lines

In order to investigate the phylogenetic distribution of resistance, a neighbor-joining clustering analysis of breeding lines was done using available SNP marker data with the Euclidean genetic distance measure in the ape (v. 5.3) R package [62]. The tree was constructed after 10.000 bootstrapping iterations, with weak nodes ($\leq$80% recurrence) collapsed into multifurcations. Circular neighbor-joining tree was generated using the iTOL (v. 5) online tool (http://itol.embl.de/, access date: 5 October 2021), enabling a color visualization of the resistance response of each line at the three field trial locations and the distribution of lines carrying the resistance concentric circles [63].

2.10. Phylogenetic Analysis and In Silico Characterization of Nucleotide-Binding Leucine-Rich Repeat Genes in Leaf Rust Resistance-Associated Regions

The leaf rust resistance-associated sites in the Weining reference genome were identified by the mapping of resistance-associated markers using the same procedure as described previously for the Lo7 reference genome. Nucleotide-binding leucine-rich repeat (NLR) genes residing in leaf rust resistance-associated regions in the Lo7 and Weining reference genomes were identified using the NLR annotation provided in a recent study by Vendelbo et al. [53], available at https://doi.org/10.5281/zenodo.5085854 (Access date: 9 July 2021). Coding sequences of potential candidate NLR genes were extracted using an online data repository [39,40]. Gene structures of NLR genes were predicted using the AUGUSTUS (3.4.0) program [64].

To investigate whether NLR genes residing in leaf rust resistance-associated sites on the Lo7 and Weining reference genomes resembled known leaf rust *R* genes, the NCBI blastp function was used for a protein–protein search in the online database and a phylogenetic analysis was conducted using a panel of cloned cereal rust *R* genes as reference [61]. The phylogenetic analysis was performed using NLR genes' conserved NB-ARC (nucleotide-binding adaptor shared by APAF-1, R proteins and CED-4) domain sequences. The NB-ARC sequences of NLR genes residing within the sites on the Lo7 and Weining reference genomes were extracted from the data repository file referred to above by Vendelbo et al. [53]. The panel of known NLR-type *R* genes comprised leaf rust (*Lr1*, *Lr10*, *Lr21*, *Lr22a*), stem rust (*Sr13*, *Sr22*) and yellow rust (*Yr5*, *Yr10*, *Yr28*). NB-ARC sequences of reference NLR genes were obtained from the UniProt online database [65]. Phylogenetic analysis was conducted using a pipeline developed by Toparslan et al. [66] in R. Multiple sequence alignment of NB-ARC domain sequences was done using the msa (v. 1.20.1) and pairwise genetic distance based on identity calculated using the seqinr (4.2-8) package in R [67,68]. A tree was constructed for the respective NLR repertoire of Lo7 and Weining reference genome and visualized using the ggtree (v. 2.2.4) R package [69].

2.11. Graphical Editing

Graphs and figures were outputted from R in svg format and manually curated using the Inkscape (v. 1.1) program (https://inkscape.org/, accessed on 7 September 2021).

3. Results

3.1. 600 K SNP Genotyping of Panel

Quality filtration of markers for low minor allele frequency, missing markers and missing individual scores across the panel led to the identification of 261,406 informative markers. Markers were homogeneously distributed across the rye genome, with an average of 32,676 markers per chromosome and mean marker-to-marker distance of 25.54 kb.

3.2. Phenotyping of Leaf Rust Resistance

In both 2019 and 2020, field trials demonstrated a high level of leaf rust disease, with a clear segregation of resistance within the assayed germplasm. Based on the AUDPC and disease progression at the natural infection site, Gylling, lines were divided into four groups (Supplementary Tables S3 and S5). The resistant ('R') group consisted of 48 restorer and 23 non-restorer germplasm lines, displaying a mean final disease severity (1–9) of 1.38 ± 0.84 standard deviation (SD) across trials (Supplementary Table S5). The partially resistant ('P-R') group consisted of nine restorer and five non-restorer germplasm lines, displaying a mean final disease severity of 3.85 ± 2.29 SD across trials. The partially susceptible ('P-S') group consisted of 12 restorer and 52 non-restorer germplasm lines, displaying a mean final disease severity of 7.21 ± 1.47 SD across trials. Finally, the susceptible ('S') group consisted of 23 restorer and 8 non-restorer germplasm lines, displaying a mean final disease severity of 7.21 ± 1.34 SD across year-location. AUDPCs for each group were 0.97 ± 3.6 SD for the 'R' group, 41.4 ± 22.9 SD for the 'P-R' group, 84.6 ± 25.8 SD of the 'P-S' group and 116.4 ± 26.6 SD for the 'S' group (Figure 1A). Calculation of a pairwise t-test showed that

the AUDPC distribution of all groups differed significantly ($p < 0.05$). Susceptible control hybrid cv. KWS Binntto demonstrated a mean final disease severity of 7.29 ± 0.86 SD across trials, with an AUDPC of 144.1 (Supplementary Table S4, Figure 1B). Resistant control hybrid cv. KWS Serafino demonstrated a mean final disease severity of 6.75 ± 1.03 SD across trials, with an AUDPC of 91.3.

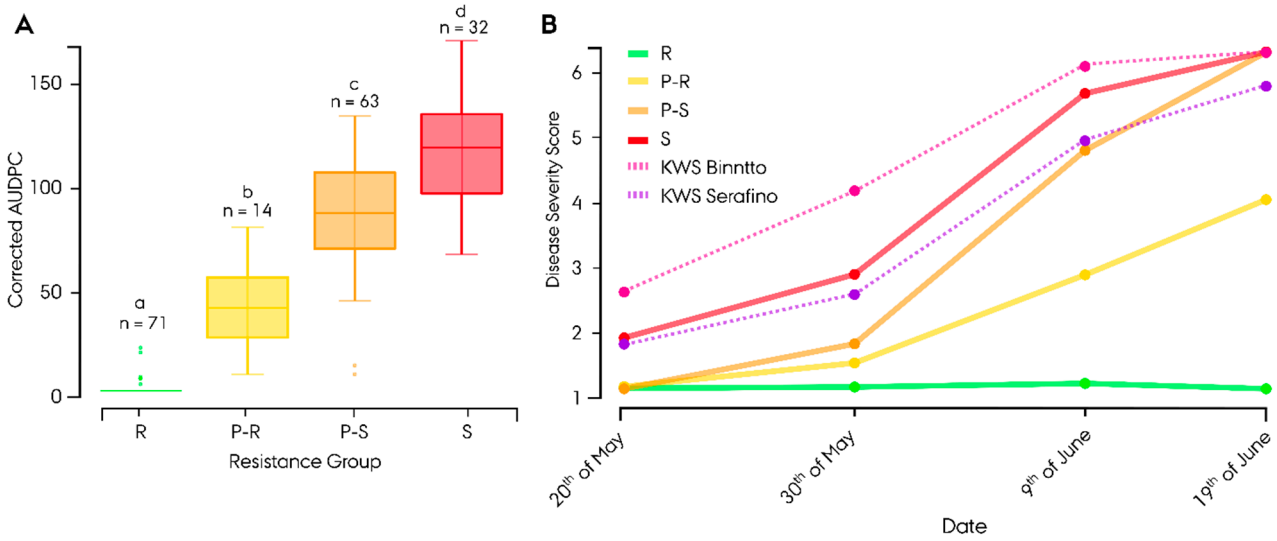


Figure 1. Leaf rust disease severity in 180 hybrid rye (*Secale cereale* L.) breeding lines grouped according to resistance response at Gylling (Denmark) field trial in 2019. (**A**) Area under disease progress curve (AUDPC) boxplot with standard error of each assigned resistance response group—resistant ('R'), partially resistant ('P-R'), partially susceptible ('P-S') and susceptible ('S') groups—with different letters indicating significant differences ($p < 0.05$). (**B**) Mean disease progression curve of each group during the growing season, including resistant control hybrid cv. KWS Binntto and susceptible control hybrid cv. KWS Serafino.

The resistance spectrum of breeding lines was visualized in a circular neighbor-joining dendrogram in iTOL, with concentric circles added to integrate the phenotypic score at each trial site (Figure 2). The phylogenetic analysis revealed a uniform distribution of resistant lines across the tree in both parental populations, while the majority of susceptible lines in the restorer population were largely found to form a secluded clade.

3.3. Genome-Wide Association Study

For the identification of SNP markers associated with leaf rust resistance, a genome-wide association study (GWAS) was done on each of the individual field trial locations, across all trials and on AUDPC using both the entire germplasm and individual parental populations (Supplementary Figure S1). GWAS conducted on the entire germplasm and non-restorer germplasm population led to a consistent finding of a highly significant peak ($-\log_{10}(p)$ = 9.7–48.5) on chromosome arm 7RS from 1.56 to 4.85 Mb, with the most associated marker at 4.7 Mb (Figure 3A, Supplementary Table S6). The most leaf rust resistance-associated SNP markers on chromosome arm 7RS explained between 15.9 and 27.1% of the phenotypic variance (Supplementary Table S6). In the GWAS conducted on the entire population across all trials, a significant peak ($-\log_{10}(p)$ = 7.9) was identified at 115.1 Mb on chromosome arm 1RS, explaining 11.8% of the phenotypic variance (Figure 3A). The *SCM9* and *Xscm1* markers co-segregating with *Pr3*, and flanking either side of the gene, mapped to 96.7 and 137.6 Mb on chromosome arm 1RS of Lo7 and 118.1 to 184.1 Mb, respectively, on the Weining reference genome. None of the GWAS analyses conducted on the restorer population alone led to the observation of significant associated SNP markers (Supplementary Figure S1). Non-significant peaks were, however, observed on chromosome arms 1RS, 2RS, 2RL 3RS, 6RL, 7RS and 7RL in the restorer population,

explaining between 6.9 and 10.2% of the phenotypic variance. While additional markers with a moderately significant association were identified in GWAS on the entire germplasm and non-restorer germplasm, these were disregarded due to inadequate consistency in the discovery relationship. Using MLM-GWAS, the leaf rust resistance-associated region on chromosome arm 7RS was found to span 11 Mb from the distal tip (Figure 3B).

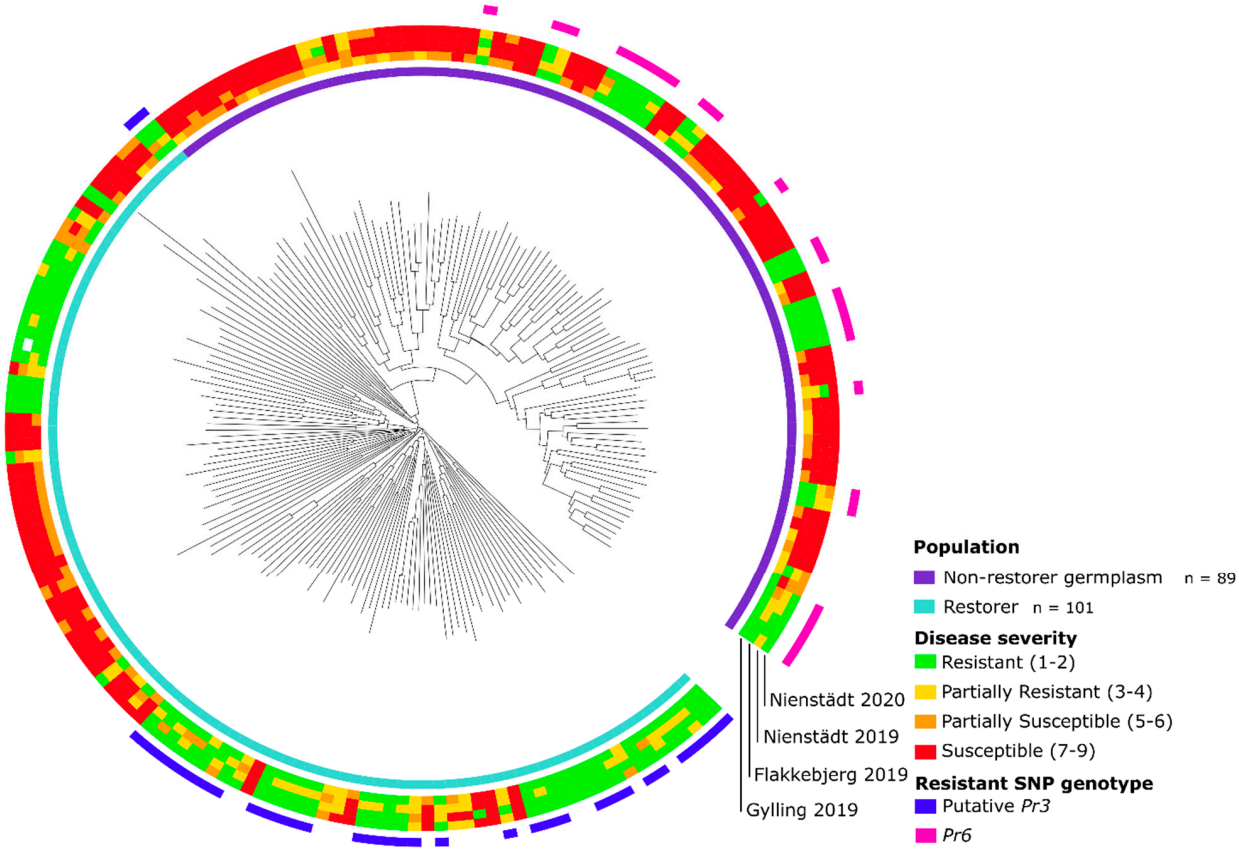


Figure 2. Circular neighbor-joining dendrogram of 180 hybrid rye (*Secale cereale* L.) breeding lines using 261,406 informative SNP markers. Leaf rust resistance (1–9) at four field trials in Denmark and Northern Germany in 2019 and 2020 displayed by concentric circles around the dendrogram. Lines carrying the resistant genotype of single-nucleotide polymorphism (SNP) markers associated with leaf rust resistance gene on chromosome arm 1RS (putative *Pr3*) and 7RS (*Pr6*) are likewise displayed in the concentric circles.

The resistant allele of the most resistance-associated marker, AX-99370891, on chromosome arm 7RS was exclusive to the non-restorer germplasm population and found in all 23 resistant lines, two out of nine partially resistant lines and one susceptible line (Figure 2, Supplementary Table S6). The resistant allele of the most resistance-associated marker, AX-99442596, on chromosome arm 1RS was exclusive to the restorer population and found in 33 out of 48 resistant lines and two out of nine partially resistant lines (Figure 2, Supplementary Table S6).

The plot-based broad sense heritability (H^2) was high for all of the linear mixed models conducted both on individual and across parental populations and locations, ranging from 0.79 to 0.92 (Supplementary Figure S1, Supplementary Table S7). In the non-restorer population, the H^2 ranged from 0.82 to 0.92, and in the restorer population, the H^2 ranged from 0.79 to 0.87.

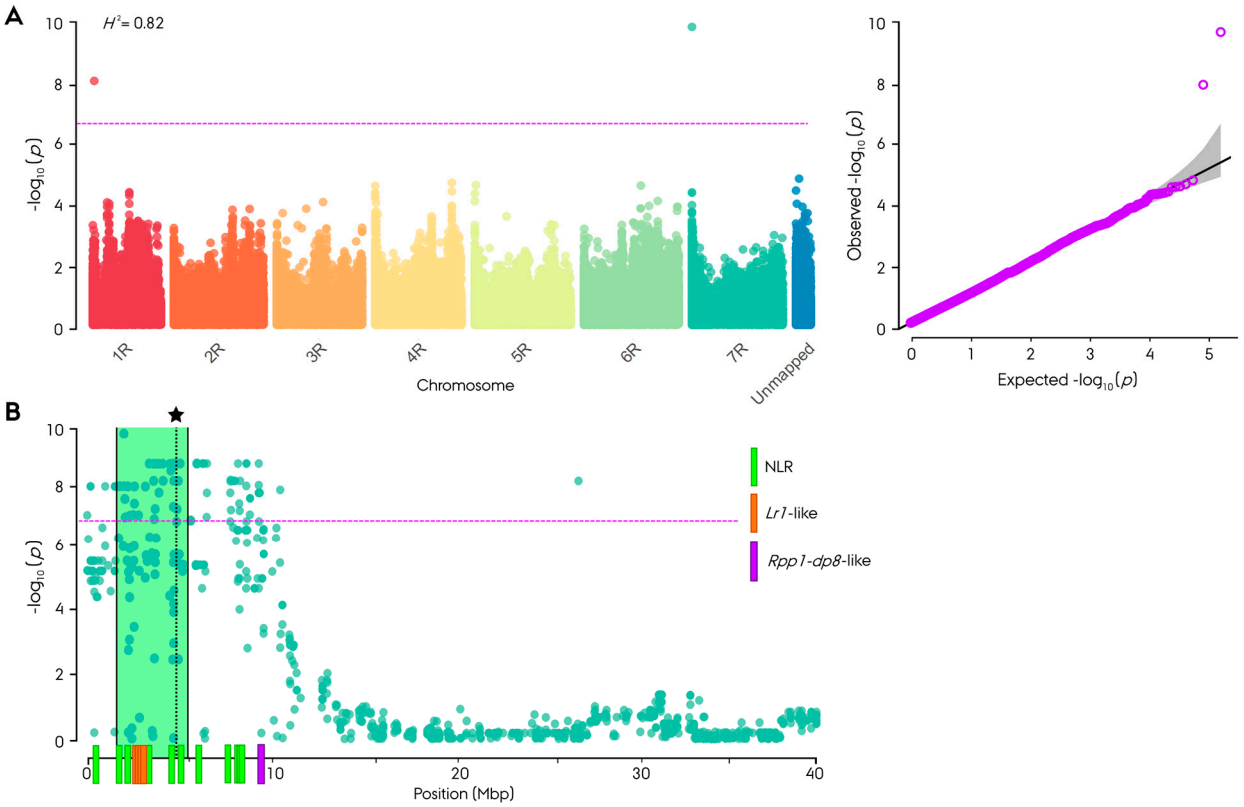


Figure 3. Manhattan plot for genome-wide association study (GWAS) using BLINK method on leaf rust disease resistance in 180 hybrid rye (*Secale cereale* L.) breeding lines using 261,406 informative SNP markers anchored to the Lo7 reference genome. (**A**) Using best linear unbiased estimator (BLUE) resistance value across four field trials in 2019 and 2020 for the entire germplasm as phenotypic input, including Q-Q plot. (**B**) Excerpt of the chromosome arm 7RS from GWAS using the same phenotypic input as in (**A**) with MLM method instead. The span of most associated marker position is highlighted in green, mean position of most associated marker by an asterisk and NLR genes in the Lo7 reference genome by vertical bars. The purple line represents the Bonferroni-adjusted significance threshold based on informative markers.

3.4. Phylogenetic Analysis and In Silico Characterization of Nucleotide-Binding Leucine-Rich Repeat Genes in Leaf Rust Resistance-Associated Block on Chromosome Arm 7RS

Mapping of 442 out of 685 resistance-associated markers on chromosome arm 7RS in the Lo7 reference genome to the Weining reference genome led to the positioning of the resistance-associated region from 5.8 to 23.6 Mb on chromosome arm 7RS (Supplementary Table S8). Gene mining in the resistance-associated region led to the identification of 33 nucleotide-binding leucine-rich repeat (NLR) genes in the Lo7 reference genome and 38 NLR genes in the Weining reference genome (Supplementary Table S9).

Phylogenetic analysis and in silico characterization led to the identification of a large cluster of full-length ('complete') NLR genes in both reference genomes, showing $\leq$80.7% sequence similarity at a complete alignment with wheat leaf rust *R* gene Lr1 (Figure 4, Table 1 and Table S10). In the Lo7 reference genome, the cluster consisted of five NLR genes situated at 2.7 to 2.9 Mb, with the most leaf rust resistance-associated markers from 1.59 to 4.9 Mb (Supplementary Table S6). In the Weining reference genome, the cluster consisted of six NLR genes situated at 12.2 to 14.5 Mb, with the most leaf rust resistance-associated markers positioned from 10.0 to 12.8 Mb (Supplementary Table S6). The resistance-associated region on chromosome arm 7RS also housed NLR genes in both reference genomes, showing $\leq$ 59.2% sequence similarity, aligning from 350 to 534 of the entire 1020 aa protein sequence, with a putative rust resistance protein *Rp1-dp8* in *Brachypodium distachyon* (Figure 4, Tables 1 and S10). The *Rp1-dp8*-like NLR gene was

situated at 9.6 Mb in the Lo7 reference genome, and, in the Weining reference genome, the two respective NLR genes were situated at 18.9 and 19.1 Mb.

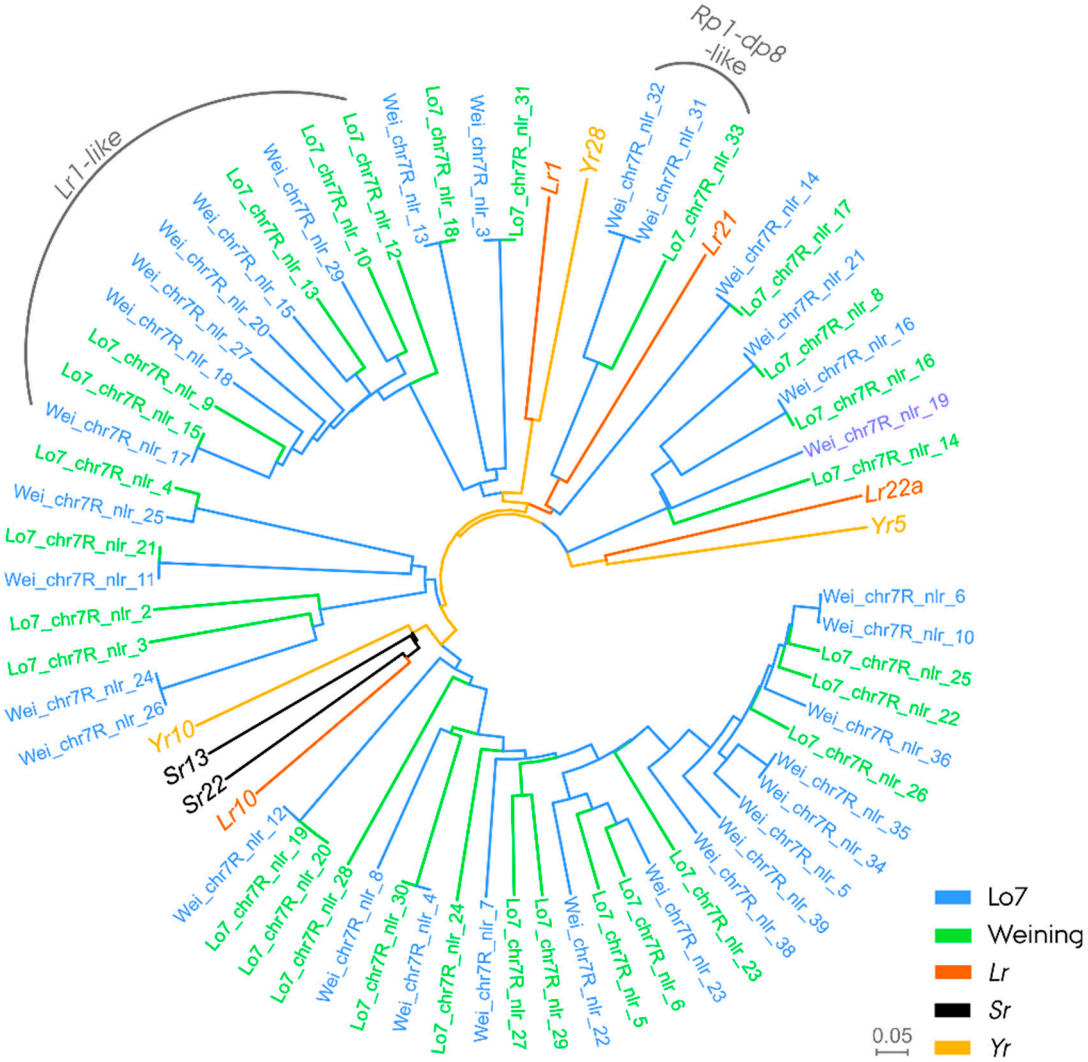


Figure 4. Phylogenetic relationship of nucleotide-binding leucine-rich repeat (NLR) genes in a leaf rust resistance association region on rye (*Secale cereale* L.) chromosome arm 7RS in the reference genomes Lo7 and Weining. The tree was constructed using the central NB-ARC domain sequence. Panels of known wheat leaf rust (*Lr*), stripe rust (*Yr*) and stem rust (*Sr*) genes are included as reference.

Table 1. Characteristics of nucleotide-binding leucine-rich repeat (NLR) genes residing in leaf resistance-associated region on rye (*Secale cereale* L.) chromosome arm 7RS in the reference genomes Lo7 and Weining showing similarity with known leaf rust resistance genes.

Reference Genome	NLR ID	Position (Mbp)	Gene Length (bp)	Predicted Protein Sequence Length (aa)	Coding Sequence (aa)	BlastP				
						Hit	Species	Alignment Length (aa)	Identity (%)	Gaps (%)
Lo7	Lo7_chr7R_nlr_9	2.37	3294	1098	1408	*Lr1 disease protein*	*Triticum aestivum*	1183	84.60	4.18
	Lo7_chr7R_nlr_10	2.41	3141	1047	1326	*Lr1 disease protein*	*Triticum aestivum*	1270	80.48	6.31
	Lo7_chr7R_nlr_12	2.80	3258	1086	1387	*Lr1 disease protein*	*Triticum aestivum*	1341	80.67	5.44
	Lo7_chr7R_nlr_13	2.81	3246	1082	1429	*Lr1 disease protein*	*Triticum aestivum*	1195	81.99	4.00
	Lo7_chr7R_nlr_15	2.87	3207	1068	1438	*Lr1 disease protein*	*Triticum aestivum*	1195	84.69	3.75
	Lo7_chr7R_nlr_33	9.57	2294	526	-	*Putative rust resistance protein Rp1-dp8*	*Brachypodium distachyon*	350	62.90	17.74
Weining	Wei_chr7R_nlr_15	12.21	3687	1075	-	*Lr1 disease protein*	*Triticum aestivum*	1064	83.72	4.29
	Wei_chr7R_nlr_17	12.35	3207	1069	-	*Lr1 disease protein*	*Triticum aestivum*	1054	85.41	3.88
	Wei_chr7R_nlr_18	12.46	6438	1730	-	*Lr1 disease protein*	*Triticum aestivum*	1113	86.03	3.87
	Wei_chr7R_nlr_20	12.53	3294	1098	-	*Lr1 disease protein*	*Triticum aestivum*	1078	84.32	3.87
	Wei_chr7R_nlr_27	14.41	1392	464	-	*Lr1 disease protein*	*Triticum aestivum*	464	81.66	3.41
	Wei_chr7R_nlr_29	14.52	11,906	1044	-	*Lr1 disease protein*	*Triticum aestivum*	915	82.43	5.75
	Wei_chr7R_nlr_31	18.91	2054	500	-	*Putative rust resistance protein Rp1-dp8*	*Brachypodium distachyon*	534	54.68	25.72
	Wei_chr7R_nlr_32	19.09	2053	440	-	*Putative rust resistance protein Rp1-dp8*	*Brachypodium distachyon*	366	59.95	18.11

4. Discussion

4.1. Evidence of Quantitative Resistance in the Germplasm

Most race-specific resistance (*R*) genes in plants are subject to classical gene-for-gene interactions with avirulence gene(s) in the pathogen, giving rise to effector-triggered immunity responses [38,70]. In turn, selection drives the emergence of virulent pathogen genotypes, permitting the evasion of R protein recognition, often termed the 'arms-race' [71,72]. As a result, monogenic inherited *R* genes exerting race-specific resistance under the gene-for-gene principle have been associated with short durability and loss of effect for disease control. In Canada, the wheat leaf rust pathogen *P. triticina* has been closely surveyed since the early 1900s, documenting, amongst others, the rapid loss of effectiveness by race-specific *R* genes such as *Lr1*, *Lr13*, *Lr14a* and *Lr26* [73–75]. Contrary to *R* genes conferring race-specific resistance, quantitative resistance conferring broad-spectrum resistance has been associated with enhanced durability [76,77]. Quantitative resistance, also referred to as partial resistance, is expressed as a susceptible infection type with reduced infection frequency and severity [78]. While major *R* genes exerting quantitative resistance have been identified, such as *Lr34* in wheat [79], quantitative resistance is often governed by several low-to-intermediate-effect quantitative trait loci (QTL) [80]. The potential additive effect, however, may permit a high level of resistance by the pyramiding of multiple QTLs, conferring quantitative resistance to leaf rust [81–83]. The adoption of marker-assisted selection (MAS) in modern plant breeding systems has aided in breeding for oligogenic traits, such as quantitative resistance [44,84]. In a large hybrid wheat breeding system, Beukert et al. [85] found that MAS constituted an efficient strategy for the introgression of leaf rust resistance QTLs. While quantitative resistance has been observed in inbred lines

and population varieties of rye, no genes or QTLs have, to our knowledge, been deployed in commercial hybrid rye cultivars [6,86,87].

In the assayed hybrid rye breeding germplasm, we observed a subset of breeding lines in both parental populations (belonging to the assigned 'P-R' group), which demonstrated a moderate to high level of quantitative resistance. These breeding lines were characterized by the later occurrence of leaf rust and reduced disease progression compared to susceptible lines. In the non-restorer germplasm population, none of the quantitative resistant lines were found to harbor the resistant haplotype of the most leaf rust resistance-associated markers on chromosome arm 7RS. This finding could suggest that the quantitatively resistant lines do not carry the major *Pr* gene, instead either harboring rare variants of a large effect, or several common QTLs of low to intermediate effect [88]. In both cases, this would dramatically reduce the phenotypic variance explained and hence the statistical power of GWAS to infer marker linkage.

In the restorer population, however, several of the quantitatively resistant lines carried the resistance allele of the putative *Pr3*-associated marker on chromosome arm 1RS. If the causative gene is *Pr3*, this could be explained by the observation of *Prs* pathotypes virulent to *Pr3* in Germany by Roux et al. [35] in 2004. However, the strict conservation of the putative *Pr3*-associated marker amongst resistant or partially resistant lines indicates that the *R* gene remains largely effective towards a large fraction of current pathotypes in Northern Germany and Denmark. However, variation in copy number [89] or structural variation [90] of the putative *Pr3* gene may also give rise to the observed quantitative resistance in these lines. This could be the case for the quantitatively resistant non-restorer germplasm lines, potentially carrying a unique variant of the resistant locus on chromosome arm 7RS.

4.2. Discovery of a Novel Major Pr Gene on Rye Chromosome Arm 7RS in the Non-Restorer Germplasm Population

The use of genome-wide association studies (GWAS) has become a routine strategy for the mining of *R* genes in crop species [91,92]. Exploiting recent advances in rye genomic resources, we successfully identified a genomic region on the chromosome arm 7RS that was significantly associated with leaf rust resistance [39,40]. With no previous *Pr* gene identified on chromosome arm 7RS, we provisionally denote the new *Pr* gene discovered in this study *Pr6*, in accordance with the nomenclature presented by Wehling et al. [34].

While *Pr6* is a new finding in rye, several leaf rust *R* genes have been identified in wheat and barley chromosomal segments syntenic to the rye chromosome arm 7RS. During *Triticeae* speciation, a series of recurrent translocation events gave rise to major patterns of chromosomal rearrangements, disturbing the collinearity of orthologous chromosomes [93,94]. In barley, the chromosome arm 5HL, which is syntenic to the distal tip of chromosome arm 7RS, harbors three major leaf rust *R* genes: *Rph9*, *Rph9.z* and *Rph12* [95–97]. In wheat, the 4A, 5B and 5D chromosomes are syntenic to rye chromosome arm 7RS [40]. In total, ten major leaf rust *R* genes have been identified on wheat chromosomes syntenic to rye chromosome arm 7RS, comprising two on 4AL (*Lr28*, *Lr30*), three on 5BS (*LrK1*, *Lr52*), two on 5BL (*Lr18*, *Lr*), three on 5DS (*Lr57*, *Lr70*, *Lr76*) and two on the 5DL (*Lr1*, *LrSyn137*) chromosome [98].

4.3. Discovery of a Putative Pr3 Gene in the Restorer Population

While the majority of resistant lines in the assayed germplasm belonged to the restorer population, no genomic region significantly associated with resistance could be identified by GWAS on the parental population alone, including the *Pr6* locus. However, several non-significant peaks were identified that could potentially correspond to both known, such as *Pr6*, and novel *Pr* genes [34,35]. GWAS used BLUE estimated resistance values across all trial locations on the entire population, which resulted in a resistance-associated SNP marker physically co-localizing with molecular markers delimiting *Pr3* on chromosome arm 1RS [35]. Intriguingly, the resistant allele of the putative *Pr3*-associated marker was

strictly confined to the restorer population. While this indicates that the putative *Pr3* locus is prevalent amongst resistant and partially resistant restorer lines, several did not carry the resistant allele. In conjunction with the finding of several non-significant, low-to-intermediate-effect QTLs in GWAS on the restorer population, this suggests the presence of additional, less prevalent *R* genes or low-to-intermediate-effect QTLs.

The ability of GWAS to establish an association between a marker and trait of interest is influenced by several biological factors [88]. In a recent population study by Vendelbo et al. [50] on the assayed germplasm, the population genetic characteristics diverged considerably, with the restorer population showing high genetic diversity, effective population size and low linkage disequilibrium The higher frequency of effective recombination events will in turn cause a more rapid decay in linkage; hence, influencing the extent of linkage between non-functional markers and the *R* gene of interest needed in GWAS to establish a significant association [88,99]. In heterogeneous cross-fertilizing crop species, the rate of decay is often rapid compared to self-pollinated species [100]. Additionally, the existence of several less prevalent *Pr* genes and low-to-intermediate-effect QTLs in various combinations would, in addition to the low sample size, further reduce the phenotypic variance explained by genes needed for GWAS to establish a significant association [92]. The potential existence of a more complex genetic architecture underlying leaf rust resistance in the restorer population is supported by the high level and diverse spectra of leaf rust resistance, discovery of several low-to-intermediate-effect QTLs and the absence of the putative *Pr3* locus in some resistant lines.

4.4. Pr6, a Potential Ortholog to Wheat Leaf Rust Resistance Gene Lr1 on Rye Chromosome Arm 7RS

Enabled by recent advances in rye genomic resources, we investigated whether the identified distal region of chromosome arm 7RS associated with *Pr6* harbored NLRs resembling known leaf rust *R* genes. In plants, NLR genes have been observed to accumulate in large clusters at recombination hotspots in the subtelomeric region, contributing to the rapid generation of novel genetic variation in NLR genes [99,101]. To detangle the NLR diversity, we conducted a phylogenetic analysis using the conserved NB-ARC domain, including cloned cereal rust genes as a reference [102,103]. The analysis led to the identification of a large clade of paralogous NLR genes showing a close evolutionary relationship and protein sequence similarity with wheat leaf rust *R* gene *Lr1*. In wheat, *Lr1* has been mapped to the 5D chromosome syntenic to rye chromosome arm 7RS [40,104]. Initial fine mapping of *Lr1* revealed a close linkage of the gene to a molecular marker, *Xpsr567* [105,106]. Intriguingly, a successive study mapped the *Xpsr567* marker to the distal region of rye chromosome arm 7RS [107]. These findings suggest that the novel leaf rust *R* gene *Pr6* could potentially be a rye ortholog to *Lr1*. A similar observation has been made on powdery mildew by Hurni et al. [108], who demonstrated that the powdery mildew resistance genes *Pm8* on rye 1RS and *Pm3* on the syntenic chromosome arm 1AS in wheat are orthologous genes. Despite an independent evolution since the species diverged 7 million years ago, the orthologous powdery mildew *R* genes retained a similar resistance function. *Pr6* in rye and *Lr1* in wheat could be another such example. However, the in silico study presented here only constitutes a preliminary finding of a potential co-evolution between a wheat and rye leaf rust *R* gene. In order to investigate whether *Pr6* constitutes an *Lr1* ortholog, identification of the causative gene, e.g., by resistance gene enrichment sequencing (RenSeq) analysis, followed by the transformation of a susceptible non-restorer germplasm line, would be required [109].

5. Conclusions

In contrast to previously investigated hybrid rye breeding germplasms, the assayed Gülzow germplasm displayed a high level of qualitative as well as quantitative resistance to leaf rust, providing opportunities for the development of leaf-rust-resistant rye hybrids. By performing GWAS on 261,406 informative SNP markers, we identified a putative *Pr3*

gene confined to the restorer population and a novel *Pr* gene on chromosome arm 7RS, provisionally denoted *Pr6*, confined to the non-restorer germplasm population. Using recent advances in rye genomic resources, we identified a large cluster of *Lr1*-like NLR genes residing in close proximity to *Pr6*. With wheat leaf rust *R* gene *Lr1* situated on 5D homologous to rye 7RS, our finding suggests that *Pr6* could potentially be a rye ortholog to *Lr1*. Despite the discovery of two *Pr* genes, our findings show that GWAS was impeded by several co-existing biological factors. To unveil the genetics underlying leaf rust resistance in the genetically diverse restorer population, we recommend developing bi- and multiparent mapping populations for the restorer, in addition to increasing the sample size [110,111]. This approach would equally permit the concurrent investigation of the genetics underlying leaf rust resistance in the quantitatively resistant lines, constituting a valuable genetic resource for enhancing the resistance durability [83]. As *R* genes in rye can be introgressed into wheat by chromosomal translocation and substitution lines, gene mining in rye serves a dual purpose, accentuating the relevance of studies in rye [60,112].

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cells11010064/s1, Figure S1: Manhattan plot for genome-wide association study (GWAS) using BLINK method of leaf rust disease resistance in an entire hybrid rye (*Secale cereale* L.) breeding germplasm ('All'), or parental populations, restorer ('R', n = 92) and non-restorer germplasm ('NRG', n = 88) using 261,406 informative SNP markers, Table S1: 1–9 scoring scale for leaf rust disease severity in rye (*Secale cereale* L.), Table S2: Leaf rust disease severity of top-yielding hybrid rye (*Secale cereale* L.), hybrid rye population mixtures and population varieties tested in the Danish official trials in 2019 and in the German list of recommended varieties in 2021, Table S3: Leaf rust disease severity score (1–9) of 180 Nordic Seed hybrid rye (*Secale cereale* L.) elite breeding lines across four field trials in Denmark and Northern Germany in 2019 and 2020, Table S4: Leaf rust disease severity score (1–9) of two hybrid rye (*Secale cereale* L.) cultivars tested in six replicates at four field trials in Denmark and Northern Germany in 2019 and 2020, Table S5: Distribution of 180 Nordic Seed hybrid rye (*Secalae cereale* L.) elite breeding lines in four leaf rust resistance groups based on area under disease progression curve and disease progression pattern, Table S6: Metrics and genotype of top-most leaf rust resistance associated markers in Nordic Seed hybrid rye (*Secale cereale* L.) elite breeding lines, Table S7: Variance components and plot broad sense heritability of leaf rust resistance in 180 Nordic Seed hybrid rye (*Secale cereale* L.) inbred lines evaluated at four field trials in Denmark and Northern Germany. The material comprised of 88 non-restorer germplasm lines (N), 92 restorer lines (R), Table S8: Mapping position of leaf rust resistance-associated markers on chromosome 7RS on Lo7 reference genome to Weining reference genome, Table S9: Annotated nucleotide-binding leucine-rich repeat (NLR) genes residing in a leaf rust resistance-associated region on chromosome arm 7RS in Nordic Seed hybrid rye (*Secale cereale* L.) elite breeding germplasm in Lo7 and Weining reference genomes, Table S10: In silico characterization of nucleotide-binding leucine-rich repeat genes residing in leaf rust resistance-associated region on rye (*Secale cereale* L.) chromosome arm 7RS in Lo7 and Weining reference genomes, Supplementary Material S1: NB-ARC sequence of nucleotide-binding leucine-rich repeat (NLR) genes residing in leaf rust resistance-associated region on rye (*Secale cereale* L.) chromosome arm 7RS in Lo7 and Weining reference genomes, Supplementary Material S2: Genotype of 261,406 Informative single-nucleotide polymorphism (SNP) markers in 180 Nordic Seed hybrid rye (*Secale cereale* L.) elite breeding lines.

Author Contributions: Conceptualization N.M.V., P.S., M.S.H., A.F.J., P.S.K., J.O. and A.J.; Data curation, N.M.V. and P.S.; Funding acquisition, J.O. and A.J.; Investigation, N.M.V., M.S.H., A.F.J., P.S.K., J.O. and A.J.; Methodology, N.M.V., K.M., P.S., M.S.H., A.F.J., P.S.K., J.O. and A.J.; Project administration, N.M.V., M.S.H., A.F.J., J.O. and A.J.; Resources, K.M., P.S., M.S.H., A.F.J., P.S.K., J.O. and A.J.; Software, N.M.V., K.M. and P.S.; Supervision, P.S., M.S.H., A.F.J., J.O. and A.J.; Validation, N.M.V. and M.S.H.; Visualization, N.M.V.; Writing—original draft, N.M.V.; Writing—review and editing, K.M., P.S., M.S.H., A.F.J., P.S.K., 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) and Pajbjerg Foundation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data availability statement: The datasets presented in this study can be found in the Supplementary Materials.

Acknowledgments: We would like to thank laboratory technician Hanne Svenstrup at Nordic Seed A/S for her contribution to the genotypic data collection. Special thanks to laboratory technician Ellen Jørgensen at the Global Rust Reference Center, Institute of Agroecology, Aarhus University (Slagelse, Denmark) for her kind mentorship in the recovery, inoculation and multiplication procedures of cereal rusts. Thanks to trial technician Jacob Sørensen for assistance in the multiplication of collected rust samples, and lead field technician Anders Almskou-Dahlgaard, both at the Institute of Agroecology, Aarhus University (Slagelse, Denmark), for the establishment of the field trials at Flakkebjerg. Many thanks to rye breeding assistants Johannes Hiller, Anette Deterding and Marlene Walbrodt at Nordic Seed Germany GmbH (Nienstädt, Germany) for their efforts in seed multiplication and activities related to the Nienstädt field trial. Thanks to field technician Henrik Christensen, assistant barley breeder Lotte Olesen and assistant barley breeding technician Helma Lübbe at Nordic Seed A/S (Dyngby, Denmark) for their efforts in the establishment of the field trials at Dyngby. The 600K SNP genotyping was provided by Eurofins Genomics (Skejby, Denmark).

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. FAOSTAT. Food and Agriculture Data. Available online: https://fao.org (accessed on 20 August 2021).
2. Miedaner, T.; Mirdita, V.; Rodemann, B.; Drobeck, T.; Rentel, D. Genetic variation of winter rye cultivars for their ergot (*Claviceps purpurea*) reaction tested in a field design with minimized interplot interference. *Plant Breed.* **2010**, *129*, 58–62. [CrossRef]
3. Mielke, H. Investigations on the control of ergot. *Nachr. Dtsch. Pflanzenschutzd.* **1993**, *45*, 97–102.
4. Miedaner, T.; Kodisch, A.; Raditschnig, A.; Eifler, J. Ergot alkaloid contents in hybrid rye are reduced by breeding. *Agriculture* **2021**, *11*, 526. [CrossRef]
5. Miedaner, T.; Geiger, H.H. Biology, genetics, and management of ergot (*Claviceps* spp.) in rye, sorghum, and pearl millet. *Toxins* **2015**, *7*, 659–678. [CrossRef] [PubMed]
6. 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]
7. Sortsinfo. The Danish Official Trial Records. Available online: https://sortinfo.dk/ (accessed on 10 February 2021).
8. Bundessortenamt. Beschreibende Sortenlisten—Getreide, Mais Öl- und Faserpflanzen, Leguminosen, Rüben, Zwischenfrüchte. Available online: https://www.bundessortenamt.de/bsa/sorten/beschreibende-sortenlisten/download-bsl-im-pdf-format (accessed on 1 October 2021).
9. 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]
10. Rodriguez-Algaba, J.; Walter, S.; Sørensen, C.K.; Hovmøller, M.S.; Justesen, A.F. Sexual structures and recombination of the wheat rust fungus Puccinia striiformis on Berberis vulgaris. *Fungal Genet. Biol.* **2014**, *70*, 77–85. [CrossRef]
11. Jin, Y.; Szabo, L.J.; Carson, M. Century-old mystery of Puccinia striiformis life history solved with the identification of Berberis as an alternate host. *Phytopathology* **2010**, *100*, 432–435. [CrossRef] [PubMed]
12. 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]
13. Roelfs, A. Epidemiology of the cereal rusts in North America. *Can. J. Plant Pathol.* **1989**, *11*, 86–90. [CrossRef]
14. Zadoks, J. Plant disease epidemiology in the twentieth century: A picture by means of selected controversies. *Plant Dis.* **2001**, *85*, 808–816. [CrossRef] [PubMed]
15. Zhao, J.; Wang, M.; Chen, X.; Kang, Z. Role of alternate hosts in epidemiology and pathogen variation of cereal rusts. *Annu. Rev. Phytopathol.* **2016**, *54*, 207–228. [CrossRef] [PubMed]
16. Roelf, A.P. Development and impact of regional cereal rust epidemics. *Plant Dis. Epidemiol.* **1986**, *1*, 129–150.
17. 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]
18. 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] [PubMed]
19. Singh, R.P.; Hodson, D.P.; Jin, Y.; Lagudah, E.S.; Ayliffe, M.A.; Bhavani, S.; Rouse, M.N.; Pretorius, Z.A.; Szabo, L.J.; Huerta-Espino, J.; et al. Emergence and spread of new races of wheat stem rust fungus: Continued threat to food security and prospects of genetic control. *Phytopathology* **2015**, *105*, 872–884. [CrossRef] [PubMed]

20. Hermansen, J.E.; Torp, U.; Prahm, L.P. Studies of transport of live spores of cereal mildew and rust fungi across the North Sea. *Grana* **1978**, *17*, 41–46. [CrossRef]
21. Hovmøller, M.; Justesen, A.; Brown, J. Clonality and long-distance migration of *Puccinia striiformis* f. sp. tritici in north-west Europe. *Plant Pathol.* **2002**, *51*, 24–32. [CrossRef]
22. Hovmøller, M.S.; Walter, S.; Bayles, R.A.; Hubbard, A.; Flath, K.; Sommerfeldt, N.; Leconte, M.; Czembor, P.; Rodriguez-Algaba, J.; Thach, T.; et al. Replacement of the European wheat yellow rust population by new races from the centre of diversity in the near-Himalayan region. *Plant Pathol.* **2016**, *65*, 402–411. [CrossRef]
23. 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] [PubMed]
24. Goyeau, H.; Berder, J.; Czerepak, C.; Gautier, A.; Lanen, C.; Lannou, C. Low diversity and fast evolution in the population of Puccinia triticina causing durum wheat leaf rust in France from 1999 to 2009, as revealed by an adapted differential set. *Plant Pathol.* **2012**, *61*, 761–772. [CrossRef]
25. Kolmer, J. Collections of *Puccinia triticina* in different provinces of China are highly related for virulence and molecular genotype. *Phytopathology* **2015**, *105*, 700–706. [CrossRef]
26. Ali, S.; Rodriguez-Algaba, J.; Thach, T.; Sorensen, C.K.; Hansen, J.G.; Lassen, P.; Nazari, K.; Hodson, D.P.; Justesen, A.F.; Hovmoller, M.S. Yellow rust epidemics worldwide were caused by pathogen races from divergent genetic lineages. *Front. Plant Sci.* **2017**, *8*, 1057. [CrossRef] [PubMed]
27. 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]
28. Li, F.; Upadhyaya, N.M.; Sperschneider, J.; Matny, O.; Nguyen-Phuc, H.; Mago, R.; Raley, C.; Miller, M.E.; Silverstein, K.A.T.; Henningsen, E.; et al. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nat. Commun.* **2019**, *10*, 5068. [CrossRef]
29. Kim, W.; Shang, H.; Samborski, D. Electrophoretic analysis of detergent-soluble polypeptides of *Puccinia recondita* f. sp. tritici, *P. recondita* f. sp. secalis, *P. hordei*, and *P. coronata*. *Can. J. Plant Pathol.* **1985**, *7*, 287–293. [CrossRef]
30. 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]
31. 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]
32. Andreasen, C.; Stryhn, H. Increasing weed flora in Danish beet, pea and winter barley fields. *Crop. Prot.* **2012**, *36*, 11–17. [CrossRef]
33. 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]
34. 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]
35. 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]
36. 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]
37. 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]
38. 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] [PubMed]
39. 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]
40. 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]
41. Krasileva, K.V.; Dahlbeck, D.; Staskawicz, B.J. Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell* **2010**, *22*, 2444–2458. [CrossRef]
42. Takken, F.L.; Goverse, A. How to build a pathogen detector: Structural basis of NB-LRR function. *Curr. Opin. Plant Biol.* **2012**, *15*, 375–384. [CrossRef] [PubMed]
43. Wang, J.; Han, M.; Liu, Y. Diversity, structure and function of the coiled-coil domains of plant NLR immune receptors. *J. Integr. Plant Biol.* **2021**, *63*, 283–296. [CrossRef]
44. Miedaner, T.; Korzun, V. Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* **2012**, *102*, 560–566. [CrossRef] [PubMed]
45. Mundt, C.C. Pyramiding for resistance durability: Theory and practice. *Phytopathology* **2018**, *108*, 792–802. [CrossRef]
46. Koller, T.; Brunner, S.; Herren, G.; Hurni, S.; Keller, B. Pyramiding of transgenic Pm3 alleles in wheat results in improved powdery mildew resistance in the field. *Theor. Appl. Genet.* **2018**, *131*, 861–871. [CrossRef]

47. 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] [PubMed]
48. Vendelbo, N.M.; Mahmood, K.; Sarup, P.; Kristensen, P.S.; Orabi, J.; Jahoor, A. Genomic Scan of Male Fertility Restoration Genes in a 'Gülzow'Type Hybrid Breeding System of Rye (*Secale cereale* L.). *Int. J. Mol. Sci.* **2021**, *22*, 9277. [CrossRef] [PubMed]
49. Miedaner, T.; Glass, C.; Dreyer, F.; Wilde, P.; Wortmann, H.; Geiger, H.H. Mapping of genes for male-fertility restoration in 'Pampa' CMS winter rye (*Secale cereale* L.). *Theor. Appl. Genet.* **2000**, *101*, 1226–1233. [CrossRef]
50. 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]
51. 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 (accessed on 7 July 2021).
52. 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. *Proc. Tenth Int. Wheat Genet. Symp.* **2003**, 789–791.
53. Vendelbo, N.; Mahmood, K.; Sarup, P.; Kristensen, P.; Orabi, J.; Jahoor, A. Discovery of a novel powdery mildew (Blumeria graminis) resistance locus in rye (*Secale cereale* L.). *Sci. Rep.* **2021**, *11*, 23057. [CrossRef] [PubMed]
54. Patterson, H.D.; Hunter, E.A. The efficiency of incomplete block designs in national list and recommend list cereal variety trials. *J. Agric. Sci.* **1983**, *101*, 427–433. [CrossRef]
55. Arias, N.; Lafarga, A.; Virto, C. RustWatch: Recommendations for Coordinated and Harmonized Wheat Rust Surveillance in Europe Based on Input from Stakeholders. Available online: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c8e0c491&appId=PPGMS (accessed on 6 June 2021).
56. De Mendiburu, F.; Reinhard, S. Agrocolae-Ten years of open source statistical tool for experiments in breeding, agriculture and biology. *PeerJ PrePrints* **2015**, *3*, e1748.
57. 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] [PubMed]
58. Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **2011**, *3*, 180–185. [CrossRef]
59. Huang, M.; Liu, X.; Zhou, Y.; Summers, R.M.; Zhang, Z. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* **2019**, *8*, 1–12. [CrossRef] [PubMed]
60. Mago, R.; Miah, H.; Lawrence, G.J.; Wellings, C.R.; Spielmeyer, 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]
61. NCBI. National Center for Biotechnology Information. Available online: https://www.ncbi.nlm.nih.gov (accessed on 13 June 2021).
62. Paradis, E.; Schliep, K. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **2019**, *35*, 526–528. [CrossRef]
63. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, 256–259. [CrossRef] [PubMed]
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]
65. The UniProt Consortium. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [CrossRef]
66. Toparslan, 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] [PubMed]
67. Bodenhofer, U.; Bonatesta, E.; Horejs-Kainrath, C.; Hochreiter, S. msa: An R package for multiple sequence alignment. *Bioinformatics* **2015**, *31*, 3997–3999. [CrossRef]
68. Charif, D.; Lobry, J.R. SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural Approaches to Sequence Evolution*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 207–232.
69. Yu, G. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinform.* **2020**, *69*, e96. [CrossRef]
70. Lolle, S.; Stevens, D.; Coaker, G. Plant NLR-triggered immunity: From receptor activation to downstream signaling. *Curr. Opin. Immunol.* **2020**, *62*, 99–105. [CrossRef]
71. Zhang, S.; Wang, L.; Wu, W.; He, L.; Yang, X.; Pan, Q. Function and evolution of Magnaporthe oryzae avirulence gene AvrPib responding to the rice blast resistance gene Pib. *Sci. Rep.* **2015**, *5*, 11642. [CrossRef]
72. Chen, J.; Upadhyaya, N.M.; Ortiz, D.; Sperschneider, J.; Li, F.; Bouton, C.; Breen, S.; Dong, C.; Xu, B.; Zhang, X. Loss of AvrSr50 by somatic exchange in stem rust leads to virulence for Sr50 resistance in wheat. *Science* **2017**, *358*, 1607–1610. [CrossRef] [PubMed]
73. Wang, X.; Bakkeren, G.; McCallum, B. Virulence and molecular polymorphisms of the wheat leaf rust fungus Puccinia triticina in Canada from 1997 to 2007. *Botany* **2010**, *88*, 575–589. [CrossRef]
74. McCallum, B.D.; Hiebert, C.W.; Cloutier, S.; Bakkeren, G.; Rosa, S.B.; Humphreys, D.G.; Marais, G.F.; McCartney, C.A.; Panwar, V.; Rampitsch, C.; et al. A review of wheat leaf rust research and the development of resistant cultivars in Canada. *Can. J. Plant Pathol.* **2016**, *38*, 1–18. [CrossRef]

75. Cowger, C.; Mehra, L.; Arellano, C.; Meyers, E.; Murphy, J.P. Virulence differences in *Blumeria graminis* f. sp. *tritici* from the Central and Eastern United States. *Phytopathology* **2018**, *108*, 402–411. [CrossRef] [PubMed]
76. Singh, R.P.; Huerta-Espino, J.; William, H.M. Genetics and breeding for durable resistance to leaf and stripe rusts in wheat. *Turk. J. Agric. For.* **2005**, *29*, 121–127.
77. Fukuoka, S.; Saka, N.; Mizukami, Y.; Koga, H.; Yamanouchi, U.; Yoshioka, Y.; Hayashi, N.; Ebana, K.; Mizobuchi, R.; Yano, M. Gene pyramiding enhances durable blast disease resistance in rice. *Sci. Rep.* **2015**, *5*, 7773. [CrossRef] [PubMed]
78. 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] [PubMed]
79. Rinaldo, A.; Gilbert, B.; Boni, R.; Krattinger, S.G.; Singh, D.; Park, R.F.; Lagudah, E.; Ayliffe, M. The Lr34 adult plant rust resistance gene provides seedling resistance in durum wheat without senescence. *Plant Biotechnol. J.* **2017**, *15*, 894–905. [CrossRef]
80. Pinto da Silva, G.B.; Zanella, C.M.; Martinelli, J.A.; Chaves, M.S.; Hiebert, C.W.; McCallum, B.D.; Boyd, L.A. Quantitative trait loci conferring leaf rust resistance in hexaploid wheat. *Phytopathology* **2018**, *108*, 1344–1354. [CrossRef]
81. Das, M.K.; Raaram, S.; Mundt, C.C.; Kronstad, W.E. Inheritance of slow-rusting resistance to leaf rust in wheat. *Crop. Sci.* **1992**, *32*, 1452–1456. [CrossRef]
82. Pilet-Nayel, M.L.; Moury, B.; Caffier, V.; Montarry, J.; Kerlan, M.C.; Fournet, S.; Durel, C.E.; Delourme, R. Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Front. Plant Sci.* **2017**, *8*, 1838. [CrossRef]
83. Huerta-Espino, J.; Singh, R.; Crespo-Herrera, L.A.; Villasenor-Mir, H.E.; Rodriguez-Garcia, M.F.; Dreisigacker, S.; Barcenas-Santana, D.; Lagudah, E. Adult plant slow rusting genes confer high levels of resistance to rusts in bread wheat cultivars from Mexico. *Front. Plant Sci.* **2020**, *11*, 1–15. [CrossRef]
84. Cobb, J.N.; Biswas, P.S.; Platten, J.D. Back to the future: Revisiting MAS as a tool for modern plant breeding. *Theor. Appl. Genet.* **2019**, *132*, 647–667. [CrossRef] [PubMed]
85. 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*, 594113. [CrossRef] [PubMed]
86. Reichel, A. Untersuchungen zur Horizontalen Resistenz von Roggen und Weizen Gegenüber Braunrost. Ph.D. Thesis, University of Halle, Halle, Germany, 1981.
87. Solodukhina, O.V. Genetic characterization of rye accessions with regard to leaf rust resistance. *Russ. J. Genet.* **2002**, *38*, 497–506. [CrossRef]
88. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **2013**, *9*, 29. [CrossRef]
89. Coriton, O.; Jahier, J.; Leconte, M.; Huteau, V.; Trotoux, G.; Dedryver, F.; de-Vallavieille-Pope, C. Double dose efficiency of the yellow rust resistance gene Yr17 in bread wheat lines. *Plant Breed.* **2020**, *139*, 263–271. [CrossRef]
90. Ellis, J.G.; Lawrence, G.J.; Luck, J.E.; Dodds, P.N. Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. *Plant Cell* **1999**, *11*, 495–506. [CrossRef]
91. Liu, H.J.; Yan, J. Crop genome-wide association study: A harvest of biological relevance. *Plant J.* **2019**, *97*, 8–18. [CrossRef] [PubMed]
92. Alqudah, A.M.; Sallam, A.; Stephen Baenziger, P.; Borner, 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]
93. Li, W.; Challa, G.S.; Zhu, H.; Wei, W. Recurrence of chromosome rearrangements and reuse of dna breakpoints in the evolution of the triticeae genomes. *G3 Genes Genomes Genet.* **2016**, *6*, 3837–3847. [CrossRef]
94. Devos, K.M. Updating the 'crop circle'. *Curr. Opin. Plant Biol.* **2005**, *8*, 155–162. [CrossRef] [PubMed]
95. Dracatos, P.M.; Khatkar, M.S.; Singh, D.; Park, R.F. Genetic mapping of a new race specific resistance allele effective to Puccinia hordei at the Rph9/Rph12 locus on chromosome 5HL in barley. *BMC Plant Biol.* **2014**, *14*, 1598. [CrossRef] [PubMed]
96. Borovkova, I.; Jin, Y.; Steffenson, B. Chromosomal location and genetic relationship of leaf rust resistance genes Rph9 and Rph12 in barley. *Phytopathology* **1998**, *88*, 76–80. [CrossRef]
97. Martis, M.M.; Zhou, R.; Haseneyer, G.; Schmutzer, T.; Vrana, J.; Kubalakova, M.; Konig, S.; Kugler, K.G.; Scholz, U.; Hackauf, B.; et al. Reticulate evolution of the rye genome. *Plant Cell* **2013**, *25*, 3685–3698. [CrossRef]
98. McIntosh, R.A.; Ducovsky, W.J.; Appels, R.; Xia, X.C. Catalogue of Gene Symbols for Wheat: 2017 Supplement. KOMUGI Integrated Wheat Science Database. Available online: https://shigen.nig.ac.jp/wheat/komugi/about/about.jsp (accessed on 6 July 2021).
99. Chen, N.W.G.; Thareau, V.; Ribeiro, T.; Magdelenat, G.; Ashfield, T.; Innes, R.W.; Pedrosa-Harand, A.; Geffroy, V. Common bean subtelomeres are hot spots of recombination and favor resistance gene evolution. *Front. Plant Sci.* **2018**, *9*, 1185. [CrossRef] [PubMed]
100. Flint-Garcia, S.A.; Thornsberry, J.M.; Buckler, E.S., IV. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* **2003**, *54*, 357–374. [CrossRef] [PubMed]
101. Nieri, D.; Di Donato, A.; Ercolano, M.R. Analysis of tomato meiotic recombination profile reveals preferential chromosome positions for NB-LRR genes. *Euphytica* **2017**, *213*, 206. [CrossRef]
102. Mondragon-Palomino, M.; Meyers, B.C.; Michelmore, R.W.; Gaut, B.S. Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. *Genome Res.* **2002**, *12*, 1305–1315. [CrossRef] [PubMed]

103. Florence, J.; Vernaldi, S.; Maekawa, T. Evolution and conservation of plant NLR functions. *Front. Immunol.* **2013**, *4*, 297. [CrossRef]
104. Feuillet, C.; Messmer, M.; Schachermayr, G.; Keller, B. Genetic and physical characterization of the Lr1 leaf rust resistance locus in wheat (*Triticum aestivum* L.). *Mol. Gen. Genet. MGG* **1995**, *248*, 553–562. [CrossRef] [PubMed]
105. Ling, H.Q.; Zhu, Y.; Keller, B. High-resolution mapping of the leaf rust disease resistance gene Lr1 in wheat and characterization of BAC clones from the Lr1 locus. *Theor. Appl. Genet.* **2003**, *106*, 875–882. [CrossRef] [PubMed]
106. Cloutier, S.; McCallum, B.D.; Loutre, C.; Banks, T.W.; Wicker, T.; Feuillet, C.; Keller, B.; Jordan, M.C. Leaf rust resistance gene Lr1, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large psr567 gene family. *Plant Mol. Biol.* **2007**, *65*, 93–106. [CrossRef]
107. Khlestkina, E.K.; Than, M.H.; Pestsova, E.G.; Roder, M.S.; Malyshev, S.V.; Korzun, V.; Borner, A. Mapping of 99 new microsatellite-derived loci in rye (*Secale cereale* L.) including 39 expressed sequence tags. *Theor. Appl. Genet.* **2004**, *109*, 725–732. [CrossRef]
108. Hurni, S.; Brunner, S.; Buchmann, G.; Herren, G.; Jordan, T.; Krukowski, P.; Wicker, T.; Yahiaoui, N.; Mago, R.; Keller, B. Rye Pm8 and wheat Pm3 are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J.* **2013**, *76*, 957–969. [CrossRef]
109. 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]
110. 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] [PubMed]
111. Ren, Y.; Hou, W.; Lan, C.; Basnet, B.R.; Singh, R.P.; Zhu, W.; Cheng, X.; Cui, D.; Chen, F. QTL analysis and nested association mapping for adult plant resistance to powdery mildew in two bread wheat populations. *Front. Plant Sci.* **2017**, *8*, 1212. [CrossRef] [PubMed]
112. Crespo-Herrera, L.A.; Garkava-Gustavsson, L.; Åhman, I. A systematic review of rye (*Secale cereale* L.) as a source of resistance to pathogens and pests in wheat (*Triticum aestivum* L.). *Hereditas* **2017**, *154*, 14. [CrossRef] [PubMed]

Discovery of a novel leaf rust (*Puccinia recondita*) resistance gene in rye (*Secale cereale* L.) using association genomics

Nikolaj, M. Vendelbo, Khalid Mahmood, Pernille M. Sarup, Mogens Hovmøller, Annemarie F. Justesen, Peter S. Kristensen, Jihad Orabi, Ahmed Jahoor

Supplementary material

Caption of all supplementary files has been given below and a subset of supplementary figures and tables included in the dissertation. The extended supplementary materials has been provided at an open-access Zenodo data repository (https://zenodo.org/record/5510292).

Supplementary Table S1: 1-9 scoring scale for leaf rust disease in rye (*Secale cereale* L.).

Group	Score	Phenotype	Leaf Coverage (%)	Incidence (%)
Resistant	1	No disease	0-0.01	0
	2	Few colonies/pustules/spots per plant	0.1	1-49
Partial Resistant	3	Few colonies/pustules/spots per tiller	0.5	51-100
	4	Several colonies/pustules/spots on lower leaves: < 10 % coverage	1	100
Partial Susceptible	5	Lower leaves: 10-25 % coverage	5	100
	6	Upper leaves: <10% coverage, lower leaves: 25-50 % coverage	10	100
Susceptible	7	Upper leaves: <25% coverage, lower leaves: 50-75 % coverage	25	100
	8	Half of leaves dead with lower leaves: 75-95 % coverage	50	100
	9	No green	75-100	100

Supplementary Table S2: Disease severity across top-yielding hybrid rye cultivars evaluated in the Danish official trial in 2019 at nine locations.

	% Leaf affected at location										
Cultivar	Horsens	Odder	Skive	Gørding	Tinglev	Årslev	Søllested	Skælskør	Gørløse	Mean	Max
KWS Berado	3	1	0,1	0,5	10	10	3	18	10	6	18
KWS Binntto	3	5	1	0	10	18	10	18	25	10	25
KWS Jethro	0,1	5	1	0,1	10	10	0,5	10	10	5	10
KWS Livado	5	0,5	0,5	3	3	3	3	5	10	3,7	10
KWS Loretto	1	18	3	5	8	18	10	10	25	11	25
KWS Receptor	8	5	3	0,5	18	25	5	33	33	15	33
KWS Serafino	8	1	0,5	0,5	10	8	5	33	10	8	33
KWS Skylor	10	10	25	10	18	18	10	33	33	19	33
KWS Tayo	1	0,1	0,1	0,1	1	8	3	25	18	6	25
KWS Vinetto	0,1	3	0,5	3	10	10	8	18	18	8	18
SU Arvid 90+10% population	5	1	5	0	18	18	1	25	18	10	25
SU Performer 90+10% population	1	3	8	5	5	3	3	25	18	8	25
SU Pluralis 90+10% population	5	5	1	0,1	10	8	0,1	10	25	7	25
Stannos	10	3	18	18	33	33	18	50	8	21	50
DL 12	0,5	0,5	1	0,5	1	18	0,5	10	10	4,7	18
Piano	3	5	5	0,5	10	8	5	25	25	10	25
Mean	4	4	5	3	11	14	5	22	19	10	25

Supplementary Table S3: Leaf rust disease severity score (1-9) of 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines across four locations in Denmark and Northern Germany in 2019 and 2020. Area under disease progression curve (AUDPC) was determined at the non-treated trial site Gylling in 2019.

Supplementary Table S4: Leaf rust disease severity score (1-9) of two hybrid rye (*Secale cereale* L.) cultivars across four field trial locations in Denmark and Northern Germany in 2019 and 2020. The cv. KWS Serafino was selected as 'resistant' control, and cv. KWS Binntto as 'susceptible' control based on available information prior to the trial.

	Dyngby 2019								Flakkebjerg 2019						Nienstädt 2019				Nienstädt 2020			
	15/5		28/5		11/6		22/6		20/5		3/6		24/6		22/5		19/6		16/6		30/6	
Cultivar	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean	Score	Mean
KWS Serafino	2 1 1 2 1 1	1	3 4 2 2 1 1	2	6 5 4 4 5 5	5	5 6 4 6 6 6	6	2 2 4 3 2 2	3	4 4 5 4 4 3	4	7 6 7 6 6 7	7	5 5 4 5 6 5	5	8 7 7 7 7 7	7	6 7 5 6 6 6	6	8 8 7 8 8 8	8
KWS Binntto	1 4 2 3 2 3	3	3 5 4 4 4 4	4	5 7 6 6 5 6	6	6 6 7 9 6 6	7	3 1 2 2 3 3	2	5 4 1 4 5 4	4	7 7 7 8 6 7	7	6 6 6 5 6 5	6	8 8 7 7 8 8	8	6 6 6 6 6 6	6	8 8 7 8 8 8	8

Supplementary Table S5: Disease severity distribution of leaf rust in 180 Nordic Seed hybrid rye (*Secale cereal* L.) breeding lines phenotyped at found trial locations in Denmark and Northern Germany in 2019 and 2020. **A**) Mean disease severity (1-9) and standard deviation (SD) for each assigned resistance group; R ('resistant'), P-R ('partial resistant'), P-S ('partial susceptible'), and S ('susceptible) at each trial site, and recorded date. **B**) Distribution of restorer (R) and non-restorer germplasm (NRG) lines in each of the assigned resistance groups.

A

		Dates							
		15/5		28/5		11/6		22/6	
Gylling 2019	Resistance Group	mean	SD	mean	SD	mean	SD	mean	SD
	R	1,01	0,08	1,03	0,20	1,08	0,46	1,00	0,00
	P-R	1,00	0,00	1,39	0,79	2,68	1,66	3,71	2,12
	P-S	1,00	0,00	1,70	0,80	4,70	1,43	6,22	1,56
	S	1,79	0,96	2,77	1,27	5,58	1,49	6,23	1,09
		20/5		3/6		24/6			
Flakkebjerg 2019		mean	SD	mean	SD	mean	SD		
	R	1,32	0,37	1,15	0,55	1,89	1,28		
	P-R	1,32	0,55	2,71	1,43	3,86	1,74		
	P-S	1,86	0,83	3,82	1,54	7,44	1,90		
	S	2,16	0,87	4,35	1,22	7,37	1,35		
		22/5		19/6					
Nienstädt 2019		mean	SD	mean	SD				
	R	1,31	0,86	1,35	1,08				
	P-R	2,68	1,63	3,43	2,32				
	P-S	4,35	1,49	7,49	1,35				
	S	4,98	1,18	7,35	1,46				
		16/6		30/6					
Nienstädt 2020		mean	SD	mean	SD				
	R	1,13	0,79	1,27	1,05				
	P-R	2,25	2,49	3,00	2,60				
	P-S	6,72	1,49	7,68	1,06				
	S	6,87	1,48	7,47	1,47				

B

Resistance Spectrum	n R	n NRG	Total
R	48	23	71
P-R	9	5	14
P-S	12	52	63
S	23	8	32

Supplementary Table S6: Top 1000 leaf rust resistance associated 600K SNP markers from BLINK Genome-wide association study in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines using BLUP estimated resistance value on field trial phenotyping data.

Supplementary Table S7:

a: Informative 600K SNP array markers in the region on chromosome arm 7RS associated with leaf rust resistance in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines.

b: Mapping of 600K SNP array markers in the region on chromosome arm 7RS associated with leaf rust resistance in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines on the 'Lo7' reference genome to the 'Weining' rye reference genome using NCBI blastn function.

Supplementary Table S8: Annotated nucleotide-binding leucine rich repeat (NLR) genes on rye (*Secale cereale*) reference genome **A**) 'Lo7', and **B**) 'Weining', residing in the leaf rust resistance associated block on chromosome arm 7RS in Nordic Seed hybrid rye breeding germplasm (n = 180) in 'Lo7' reference genome.

Supplementary Table S9: NB-ARC domain of nucleotide-binding leucine rich repeat genes residing in leaf rust resistance associated block on rye (*Secale cereale* L.) 7RS in a hybrid rye breeding germplasm (n = 180) in 'Lo7' and 'Weining' reference genome.

Supplementary Table S10:

a: Characteristics of candidate leaf rust resistance nucleotide binding leucine rich repeat genes residing in leaf resistance associated block on rye (*Secale cereale* L.) chromosome arm 7RS.

b: Pairwise Multiple sequence alignment identity of candidate leaf rust resistance nucleotide binding leucine rich repeat genes residing in leaf resistance associated block on rye (*Secale cereale* L.) chromosome arm 7RS.

Supplementary Figure S1: GWAS using BLINK method on 261,406 informative SNP markers in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines using BLUP estimated LR resistance value for all individual trial sites and across.

Supplementary Figure S2: Phylogenetic tree of 1,027 nucleotide-binding leucine-rich repeat (NLR) proteins NB-ARC domain in 'Weining' rye (*Secale cereale* L.) reference genome. NB-ARC domains of known NLR genes have been included as reference. NLR genes residing in a brown rust resistance block in the sub-teleomeric region of chromosome arm 7RS (LR 7RS) have been colored in light green.

Supplementary material S1: NB-ARC sequence of nucleotide-binding leucine-rich repeat (NLR) proteins in leaf rust resistance associated region on chromosome arm 7RS in the rye reference genome 'Lo7' and 'Weining'.

Manuscript IV

Discovery of resistance genes in rye by long-read target sequencing and association genetics

By

Nikolaj M. Vendelbo[1,4]*, Khalid Mahmood[1], Burkhard Steuernagel[2], Brande B.H. Wulff[2,3], Pernille Sarup[1], Mogens Hovmøller[4], Annemarie F. Justesen[4], Peter S. Kristensen[1,] Jihad Orabi[1], and Ahmed Jahoor[1,5]

[1] Nordic Seed A/S, Odder, Denmark

[2] John Innes Centre, Norwich Research Park, Norwich, UK

[3] Center for Desert Agriculture, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

[4] Department of Agroecology, Faculty of Technology, Aarhus University, Slagelse, Denmark

[5] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

Publication details

To be submitted to New Phytologist

Discovery of resistance genes in rye by long-read target sequencing and association genetics

Nikolaj M. Vendelbo[1,4], Khalid Mahmood[1], Burkhard Steuernagel[2], Brande B.H. Wulff[2,3], Pernille Sarup[1], Mogens Hovmøller[4], Anne-Marie Fejer Justesen[4], Peter S. Kristensen[1,] Jihad Orabi[1], Ahmed Jahoor[1,5]

[1] Nordic Seed A/S, Odder, Denmark
[2] John Innes Centre, Norwich Research Park, Norwich, UK
[3] Center for Desert Agriculture, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
[4] Department of Agroecology, Faculty of Technology, Aarhus University, Slagelse, Denmark
[5] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

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

Summary

● Most released rye cultivars are susceptible to leaf rust due to the low resistance exhibited by the predominant hybrid rye breeding gene pools Petkus and Carsten.

● To discover new sources of leaf rust resistance, we phenotyped a diverse panel of 92 inbred lines from the less prevalent Gülzow germplasm using six distinct isolates of *Puccinia recondita* f. sp. *secalis*. We identified 51 lines with resistance against all isolates. A genome-wide association study using 261,406 informative single-nucleotide polymorphisms (SNPs) revealed five genomic regions associated with resistance on chromosome arms 1RS, 1RL, 2RL, 5RL and 7RS.

● In order to identify the underlying resistance genes against *Puccinia recondita* (*Pr*), we sequenced the complement of rye nucleotide binding leucine-rich repeat (NLR) intracellular immune receptors 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 highlighted four candidate NLR-encoding *Pr* genes, one of which mapping to the same genomic region as molecular markers delimiting the *R* gene *Pr3* on chromosome arm 1RS and the top SNP associated with resistance in the diversity panel.

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 (Miedaner and Sperling, 1995). Because the predominant heterotic gene pools 'Petkus' and 'Carsten' used for the breeding of rye hybrids suffer from low inherent resistance, most newly released cultivars are highly susceptible to leaf rust (Supplementary Table S1, Miedaner et al., 2002; Sortsinfo, 2021). 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 (Miedaner et al., 2011). Unlike wheat leaf rust, stripe rust and stem rust the alternate plant host of *Prs*, small bugloss (*Anchusa arvensis* L.), is widespread in the Northern European flora and a common weed in agricultural fields (Anikster et al., 1997; Andreasen and Stryhn, 2012; Hanzlik and Gerowitt, 2012; De Mol et al., 2015). 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 (Hovmoller et al., 2008; Miedaner et al., 2011; Kolmer et al., 2019; Figueroa et al., 2020).

Host resistance represents a sustainable and more environmentally conscious alternative to chemical control strategies (Nelson et al., 2018). 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) (Wehling et al., 2003; Roux et al., 2004; Vendelbo et al., 2021a). 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) (McIntosh et al., 1995; Friebe et al., 1996). The majority of *R* genes encode intracellular nucleotide-binding leucine-rich repeat (NLR) immune receptor proteins (Kourelis and Van Der Hoorn, 2018). 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 (Qi et al., 2012; Monino-Lopez et al., 2021); (ii) a conserved nucleotide-binding (NB) domain in the middle of the protein that is involved in regulation of NLR activation (Williams et al., 2011); and (iii), an N-terminal coiled-coil (CC) domain mediating the downstream immune signalling cascade (Maekawa et al., 2011). In the

reference genome of the German inbred grain-type rye line Lo7, 1,167 NLR genes have been reported with this number reaching 1,447 in the Chinese forage-type Weining (Li et al., 2021; Rabanus-Wallace et al., 2021).

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 (Miedaner and Korzun, 2012; Beukert et al., 2020; Liu et al., 2020). 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 (Hickey et al., 2017; Ghosh et al., 2018). 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 (Bajgain et al., 2016; Rollar et al., 2021). Historical recombination events naturally present within a diverse panel for a crop specie can also be exploited for mining *R* genes and QTLs through genome-wide association studies (GWAS) and SNP genotype data (Alqudah et al., 2020). GWAS identifies a genomic region ing a trait of interest and, predominantly, a non-functional marker associated with the trait that can be readily implemented in MAS (Korte and Farlow, 2013). 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 exhibit considerable divergence with the reference sequence (Florence et al., 2013).

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 (Arora et al., 2019). 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 (Dong and Ronald, 2019). However, the *de*

novo assembly of NLR genes from short-reads data is a complex task due to their high copy number and sequence similarity between paralogues (Jupe et al., 2013). This limitation was previously alleviated by coupling RenSeq with long-read PacBio® single-molecule real-time (SMRT) sequencing (Witek et al., 2016; Xing et al., 2018). Since then, improvements to the PacBio® HiFi sequel II system have demonstrated a 99.9% accuracy in recall rate of single-nucleotide variants (SNVs) and 96% for insertions and deletions (InDels) (Wenger et al., 2019). 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*) (Steuernagel et al., 2016; Witek et al., 2016; Xing et al., 2018; Arora et al., 2019; Barbey et al., 2019; Van de Weyer et al., 2019; Seong et al., 2020).

The objective of this study was to determine the efficacy of SMRT AgRenSeq for the mining of leaf rust *R* genes in rye.

Methods

Collection of Danish *Puccinia recondita* f. sp. *secalis* samples, their multiplication and single pustule isolation

With no molecular markers nor 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 (Table S2, Fig. S1). Multiplication of field samples and isolation of single pustule isolates (SPI) were performed as described by Vendelbo et al. (2021a).

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-grain 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 section of two coleoptiles and primary leaves was excised after 7 d (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 d prior to freeze drying for an additional 2 d. DNA extraction was performed using an adapted SDS-based

method according to USDA (2006) according to Pallotta et al. (2003). The DNA concentration and 260/280–nm absorption ratio were measured using an Epoch™ microplate spectrophotometer (Biotek®). Fragmentation of genomic DNA was assessed by size separation on a 1.2% (w/v) agarose gel.

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. (1971) and Hovmoller et al. (2017). Seedlings with a mean IT between 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. (2021a). In the trial, eight seeds per line were sown in a 28-hole tray containing a coarse-grain sphagnum substrate. The trial followed a partial randomized design with two repetitions for each of the two trial replicates, hence producing four scorings per line per SPI evaluated.

At 14 d after sowing, each tray was inoculated with 30 mg of SPI spores solubilized in 4 mL 3M™ Novec™ 7100 (Sigma-Aldrich®, Missouri, USA) engineering fluid according to the method described by Thach et al. (2015). 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.

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 ≥1.8 were sent for SNP genotyping at Eurofins Genomics Europe Genotyping (Aarhus, Denmark). Genotyping was accomplished using a 600K SNP array with 600,843 SNP markers on an Affymetrix GeneTitan™ Scanner platform (Thermo Fisher Scientific, Massachusetts, US) platform (Bauer et al., 2017). Mapping of the SNP markers to the Lo7 rye reference genome and characterisation of its performance in the assayed germplasm was recently investigated by (Vendelbo et al., 2021b). The marker map is available at https://doi.org/10.5281/ZENODO.5086235.

Data analysis

Bioinformatics analysis was done in the R Studio (v. 1.3.959) interface with R statistical software (v. 4.0.1) with various predesigned packages (RStudio Team, 2015; R Core Team,

2021). Prior to analysis of the SNP genotyping data, markers were filtered for marker allele frequency of at least 5%, missing individual score below 20%, and missing marker score below 10% to identify informative markers. Computationally demanding tasks were run on 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.

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 (Lipka et al., 2012). The phenotypic input for GWAS included all individual recordings for the six SPI *Prs* scorings averaged across the two replicates.

***In silico* test of the bait library**

Before using the 60 K Triticeae nucleotide-binding leucine-rich repeat (NLR) bait library (Tv_1) designed by Steuernagel et al. (2016), an *in silico* test was run to evaluate its capture success in 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) (NCBI, 2021). 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. (2013), who reported that ~80% sequence identify between bait and NLR sequence is sufficient for enrichment. The Tv_1 bait library is available at https://github.com/steuernb/MutantHunter/. 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 %).

Phylogenetic analysis and pair-wise selection of restorer lines

To reduce the false association from non-causative background, 10 pairs of lines were selected for Single-Molecule Real Time *R* gene enrichment and Sequencing (SMRT RenSeq), each comprising of two phylogenetically closely related lines but exhibiting a divergent resistance profile. Phylogenetic analysis was performed with a neighbour-joining clustering of lines using measures of their Euclidean genetic distance using the ape (v. 5.3) R package (Paradis and Schliep, 2019). 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/), allowing a colour-based visualization the resistance response presented by each line against the six *Prs* SPIs as concentric circles (Letunic and Bork, 2019).

Single-molecule real-time high-fidelity resistance gene enrichment sequencing

The construction of PacBio® RenSeq libraries was outsourced to Arbor Biosciences (Michigan, USA). A minimum of 5 µg genomic DNA per sample 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 60K Triticeae NLR bait library (Steuernagel et al., 2016). 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.

***De novo* assembly and NLR annotation**

The CCSs were assembled with HiCanu (v. 2.0; -pacbio-hifi, trimReadsCoverage = 2, errorRate = 0.01, genomeSize = 8.8m, minOverlapLength = 500, minReadLength = 1000) (Nurk et al., 2020). The expected genome size of the rye NLR repertoire was estimated using NLR data from the recent rye reference genome Lo7 (Rabanus-Wallace et al., 2021) with an addition of 40% flanking region in accordance with the study by Van de Weyer et al. (2019). Contigs associated with NLR genes were annotated using NLR-Parser (v.3) and NLR-Annotator (https://github.com/steuernb/NLR-Annotator) (Steuernagel et al., 2015; Steuernagel et al., 2020). NLR contigs were mapped to the rye reference genome Lo7 using NCBI BLASTN (v. 2.9.0+) with a significance threshold set to 1×10^{-5}, selecting the physical position of the top hit (NCBI, 2021). Contig positions were kept if they met the criteria of ≥90% alignment length and ≥85% alignment identity.

***In vitro* test of the bait library**

To investigate the Tv_1 bait library performance *in vitro*, the individual restorer line SMRT RenSeq data were subjected to a similar analysis as 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 five lines. The distribution of bait alignment, bait alignment identity, and bait category distribution were visualized using the ggplot2 (v. 3.3.3) R package (Wickham, 2011). The distribution of baits per NLR contig and NLR contig per bait was visualized using the

YaRrr R (v. 0.1.5) package. The R scripts for the *in silico* and *in vitro* assessment of the 60K bait library on the panel are provided as supporting information.

_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 ≤10 per line (Marcais and Kingsford, 2011). 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 ≤3 lines. *k*-mer processing of the CCS reads was done following the Java source code published at https://github.com/steuernb/AgRenSeq.

Association genetics RenSeq analysis

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

$$\text{Stackman}^{'}\text{s IT score} = 2 - \left(\frac{4}{9} * Severity\ Score_{\,0-9}\right)$$

Identification of candidate leaf rust resistance genes amongst the NLR-annotated contigs was done by AgRenSeq analysis following the Java pipeline provided by Arora et al. (2019) at https://github.com/steuernb/AgRenSeq and general linear regression model (AgRenSeq-GLM) using the Python module provided at https://github.com/kgaurav1208/AgRenSeq_GLM. This modulated analysis utilises a linear regression model for each of the *k*-mers to account for population structure amongst lines and likelihood test for nested models to output a *P*-value for each of the *k*-mers' association to resistance. A standard Bonferroni-corrected threshold of α = 0.025 was used as the significance threshold.

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 2,000 bp and with 95% sequence identity (NCBI, 2021). 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) (GeneiousPrime, 2020). The gene structure of the manually assembled consensus NLR contigs was predicted using the AUGUSTUS (3.4.0) program (Stanke et al., 2004). The prediction of protein domain structures was done using InterPro-Scan (Jones et al., 2014). Prediction of LRR motifs was done using LRRpredictor (v.

1.0) (Martin et al., 2020). Coding sequences were mapped to the Lo7 and Weining reference genomes using NCBI BLASTN (v. 2.9.0+) (Li et al., 2021; NCBI, 2021; Rabanus-Wallace et al., 2021).

To distinguish unique NLR genes from homologues 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. (2020) in R. Multiple sequence alignment of protein sequences was done using 'Clustal Omega' in the msa (v. 1.20) R package. A neighbour-joining tree was constructed using Nei's standard genetic distance and visualized using the ggtree (v. 2.2.4) R package (Yu, 2020). 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 (NCBI, 2021; UniProt, 2021).

Identification of single nucleotide variants (SNVs) and insertions/deletions (InDels) between candidate gene variants was done by multiple sequence alignment using the Multiple Sequence Comparison by Log- Expectation (MUSCLE) method in Geneious Prime (v. 2020.2.3) with ≤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 (Vendelbo et al., 2021b). Annotation files are available at https://doi.org/10.5281/zenodo.5085854. If present in the reference genomes (≥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 genome by BLASTN (Roux et al., 2004; Mago et al., 2005).

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/).

Results

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 (Table S2, Fig. 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) (Table S3). Four lines showed a SPI-specific resistance response. Fig. 1 illustrates the 10 infection types as a resource for future studies of leaf rust resistance in rye.

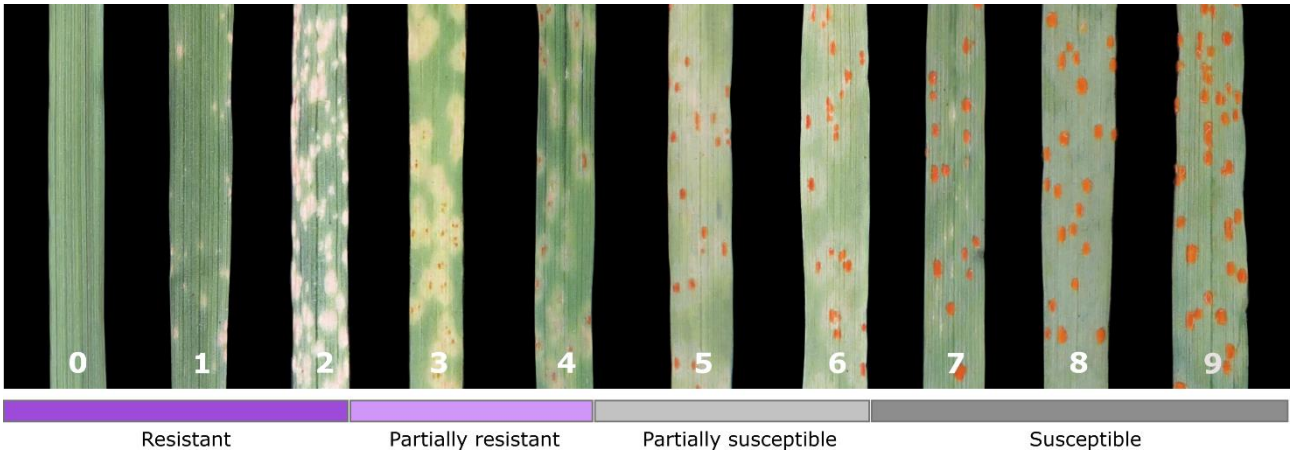


Fig. 1. Infection type (IT) response scale for leaf rust in rye (*Secale cereale* L.) caused by the fungal pathogen *Puccinia recondita* f. sp. *secalis* according to Hovmøller., *et al.* (2017) and McNeal., *et al*. (1971). In terms of virulence/avirulence, IT 0–6 are considered 'avirulent' and 7–9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

Genome-Wide Association Study

Quality filtering of markers from the 600 K SNP array allowed the identification of 261,406 informative markers (Datafile 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 (Fig. S2, 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, Fig. 2). However, none of the markers exhibited a *p*-value above the Bonferroni significance threshold of $-\text{Log}_{10}(p) = 6.72$. 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	$-\text{Log}_{10}(p)$	Phentypic 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

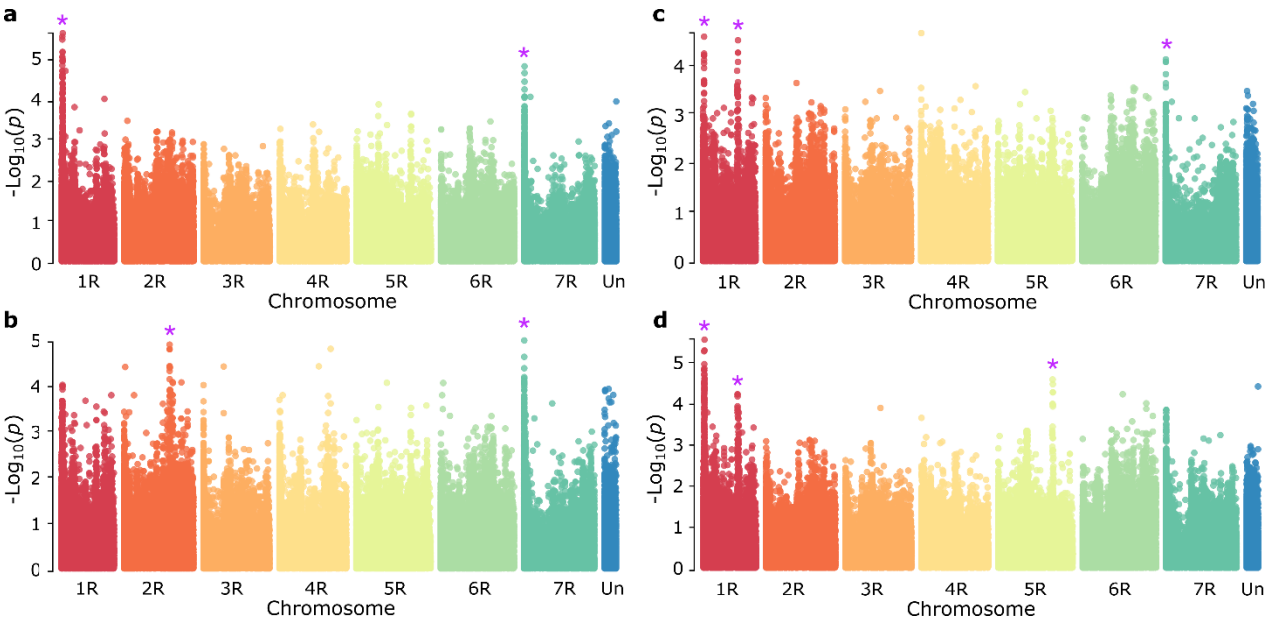


Fig. 2. Manhattan plots for genome wide association study (GWAS) for leaf rust resistance in 92 inbred rye lines. GWAS was performed using 261,406 informative SNP markers mapped to the Lo7 reference genome. The entire germplasm collection was phenotyped for resistance against six distinct *Puccinia recondita* f. sp. *secalis* single pustule isolates (SPIs) in a greenhouse, as 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.

In silico test of 60K Triticeae NLR bait library for rye

Prior to RenSeq analysis, we assessed the performance of the 60 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 1,125 out of the 1,167 annotated NLR sequences in the Lo7 reference (Table S7). Each bait mapped on average to 5 ± 6 (standard deviation, SD) unique NLR sequences, with a mean of 140 ± 164 (SD) unique baits mapped per NLR sequence (Table S5).

Phylogenetic analysis and pair-wise 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 (Fig. 3a). The corresponding shortlisted resistant lines exhibited five different infection types (ITs) spectra: resistant (ITs_1), partially resistant (ITs_2), and three different SPI-specific spectra (ITs_{3-5}) (Fig. 3b).

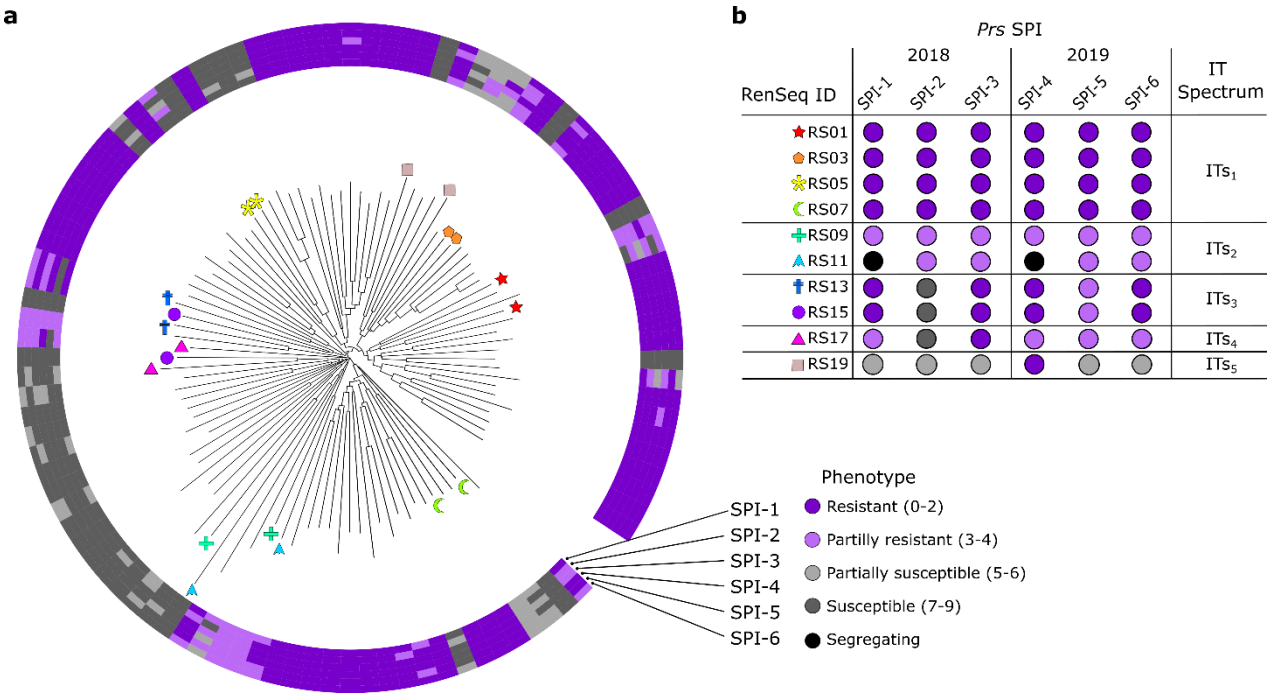


Fig. 3. Phylogenetic analysis and type of leaf rust infection response in 92 inbred rye lines. (**a**) Circular neighbour-joining dendrogram indicating the infection types (IT) against six geographically distinct Danish *Puccinia recondita* f. sp. *secalis* single pustule isolates, shown as concentric circles. Ten pairs, comprising two closely related lines with divergent IT, selected for SMRT RenSeq are marked by stars. (**b**) IT spectrum of the 10 resistant lines.

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® 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 (Table S6). *De novo* assembly using HiCanu produced 482 to 2,122 contigs per line, with a mean of 1,330 contigs. NLR annotation of the assemblies led to the identification of 288 to 1,102 NLR contigs per line, with a mean of 646 contigs (Table S6), consisting of 151 to 573 partial (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 60 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 ± 2,910 (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 five lines (Fig. S3e). Across lines, 75.9% of the aligned baits were shared with Lo7. Baits were found to align with an average of 113 bp ± 14.9 (SD) on average, with a mean alignment identity of 91.9% ± 4.3 (SD) (Fig. S3a-b). Approximately 0.5% of the baits aligned with less than 80% alignment identity. Each bait mapped to an average average of 4 ± 7 SD unique NLR contigs, with a mean of 63 ± 82 (SD) unique baits per NLR contig (Fig. S3c-d, Table S5).

_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 (Fig. 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 potential involved in resistance against leaf rust (Fig. S6). In total, we identified 25 NLR contigs harbouring resistant-specific *k*-mers were identified across the shortlisted resistant lines (Fig. 4a-b, Table S8a).

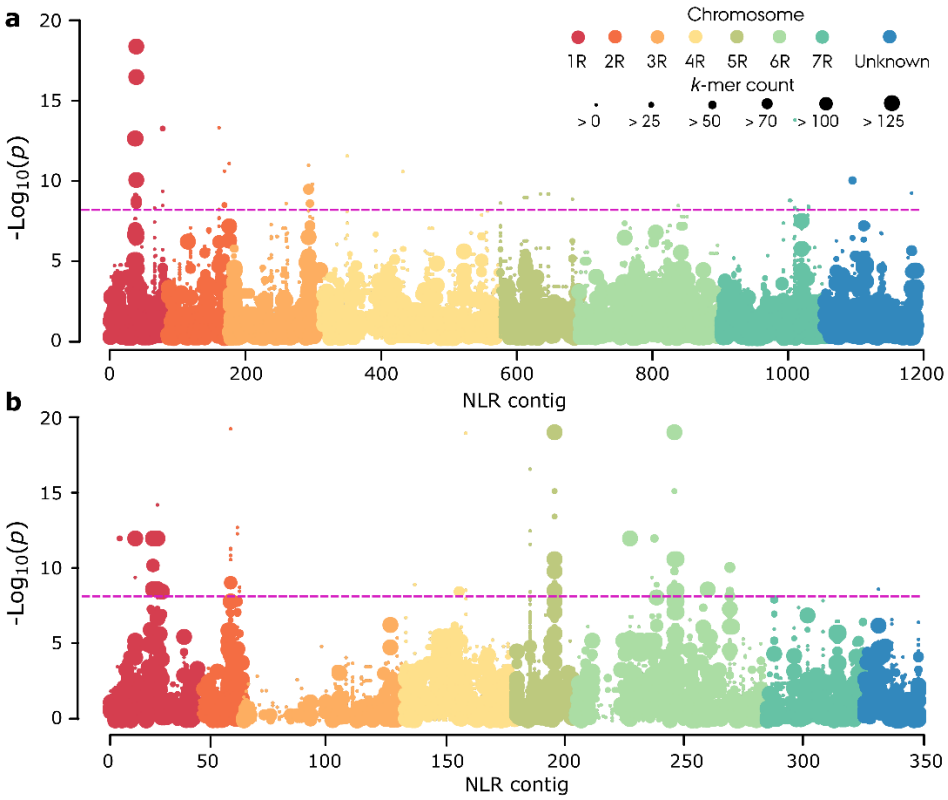


Fig. 4. Identification of candidate leaf rust resistance NLR genes by AgRenSeq in 20 inbred rye lines. Each circle column on the x-axis represents an NLR contig from the SMRT RenSeq assembly of a resistant line anchored to the Lo7 rye reference genome, and each dot on the y-axis represents a *k*-mer 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.

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 harbouring the candidate NLR contig, and trimmed and assembled them. This manual assembly step increased the contig size by 2 kb ± 1.5 kb (SD), resulting in a mean contig size of 7.6 kb ± 1.8 kb (SD) (Table S8). Out of the 25 contigs, 16 contigs were considered to be complete (full length) NLR genes, with the remaining nine being partial (incomplete). On average, the full-length NLRs encoded an NLR protein of 1,147 amino acids (aa) ± 297 (SD). To distinguish unique NLR genes from homologous contigs within the panel, we performed a phylogenetic analysis based on sequence alignment of the

NLR proteins (Fig. 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.

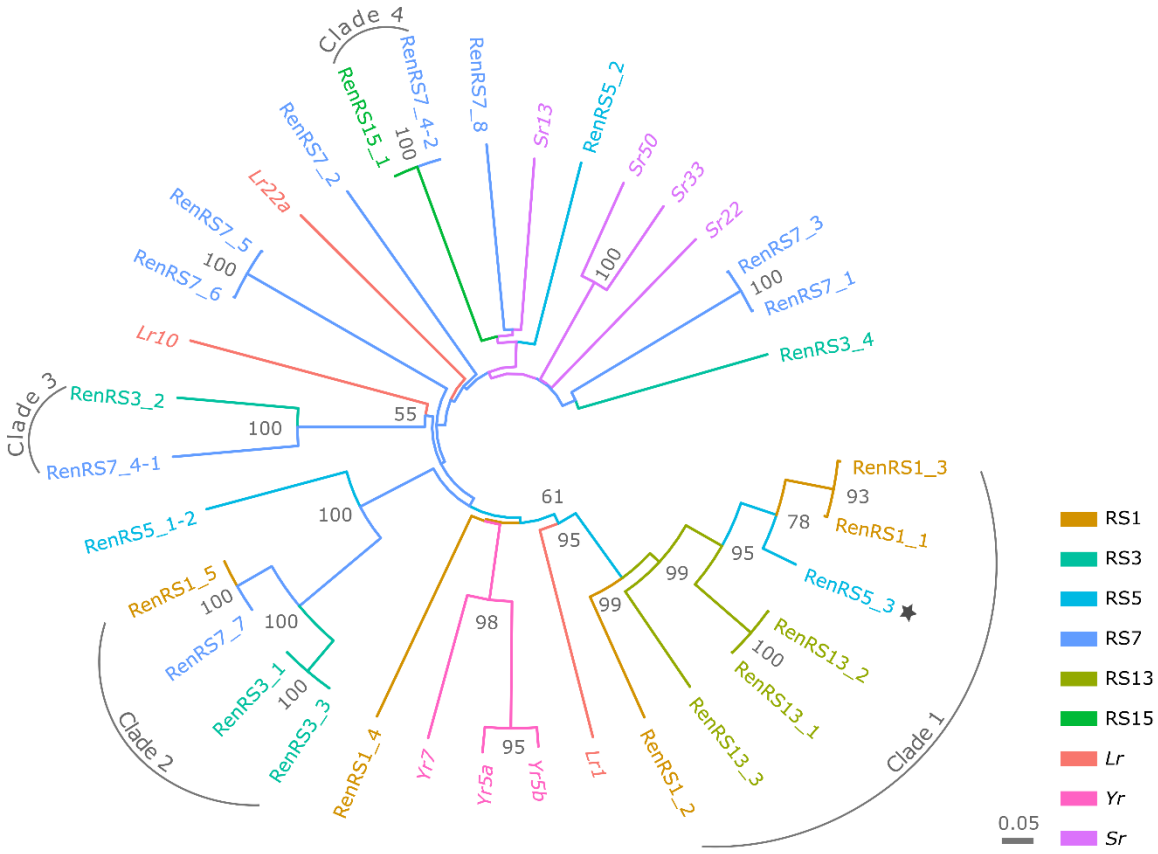


Fig. 5. Phylogenetic relationship of candidate leaf rust resistance (*R*) NLR genes in 20 inbred rye lines. The tree was constructed using protein sequences, with bootstrapping values set as percent recurrence of from 10,000 iterations shown for nodes with ≥50% recurrence. Candidate leaf rust *R* gene RenRS5_3 is marked with 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 1,414 to 1,700 aa (Fig. 5, 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 identified from the GWAS analysis with the 600K SNPs, spanning from 101 to 117 Mb (Table 1, Fig. 6a, Table S4). This genomic region contained five NLR genes in In the Lo7 reference genome. Of the 2,558 SNP markers positioned within the region on the Lo7 chromosome arm 1RS, 1,817 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 of five NLR gene each within the region spanning 130–145 Mb (Fig. 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 *Xmwg68* marker co-segregating with *Lr26* mapped to 16.6 Mb in Lo7 and 19.5 Mb in Weining, with *Lr26* positioned distal to *Xmwg68* towards the telomeric tip and 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 (Fig. 6a-b).

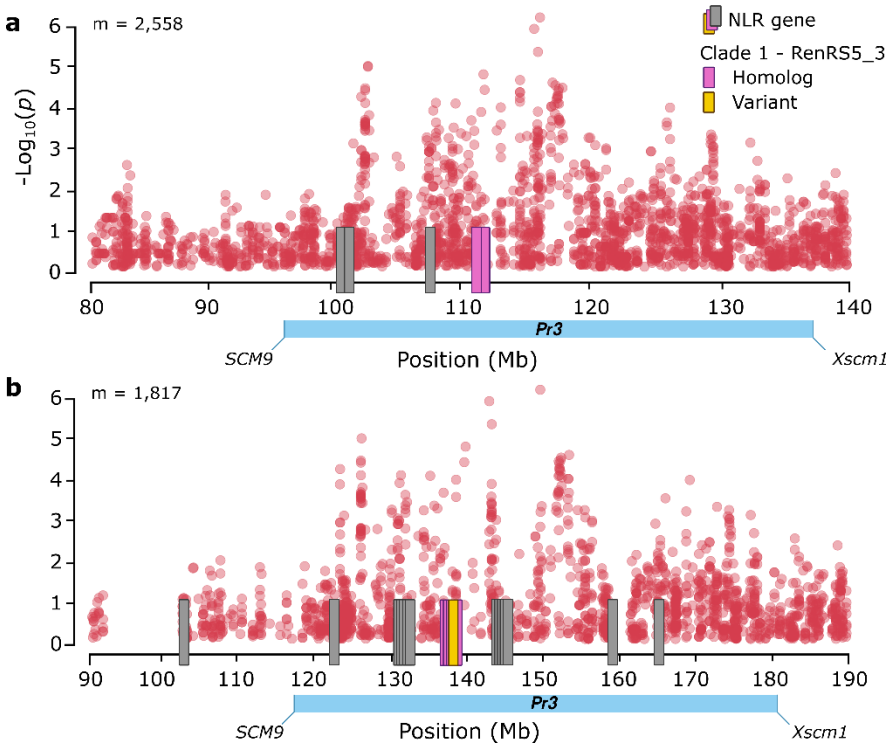


Fig. 6. Visualization of SNP markers associated by GWAS with leaf rust resistance on chromosome arm 1RS in 92 inbred rye lines. (**a**) Lo7 reference genome with 2,558 associated markers. (**b**) Weining reference genome with 1,817 associated markers. The region harbouring the known rye leaf rust resistance gene *Pr3* is marked as a blue bar based on flanking genetic markers *SCM9* and *Xscm1* (Roux et al., 2004). The position of annotated NLR genes are marked as vertical bars. NLR genes showing ≥80% sequence identity to the clade 1 member RenRS5_3 candidate *Pr* gene are in pink ('Homologue') and genes with ≥95% sequence identity are in orange ('Variant').

Clade 2 consisted of four NLR contigs encoding near-identical canonical NLR proteins of 1,383 to 1,417 aa, distinguished by intraspecific polymorphisms in the distal end of the C-terminal LRR domain (Table 2, Table S8). Clade 3 contained two NLR contigs encoding canonical NLR

proteins of 950 and 1,050 aa that exhibited considerable intraspecific polymorphisms. Clade 4 was defined by two identical NLR contigs encoding a non-canonical 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 relationship identified in resistant inbred rye lines. Contigs were mapped to the rye Lo7 reference genome.

		Assemblies		**Anchoring position in Lo7**	
Clade	**NLR contigs**	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).

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 for candidate *Pr* genes for further characterization. The NLRs in this clade mapped to the region on chromosome arm 1RS dsplaying the greatest association with leaf rust resistance (Fig. 2), which also overlapped with the genomic location of the known leaf rust resistance gene *Pr3* (Fig. 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 homologues (Fig. S7). We then selected RenRS5_3 for further study, as its complete annotation further facilitated its characterization. The RenRS5_3 contig was present (with ≥ 95% sequence identity) in 7 out of the 20 rye lines analyzed (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 199–3,313 bp and 3,404–4,015 bp (Fig. 7a). The corresponding NLR protein was predicted to have 1,306 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 1,306 containing 25 LRR motifs (Fig. 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 IT_1 group with 99.9% sequence identity; the second variant, denoted S, was conserved amongst susceptible or

partially resistant lines belonging to the IT_2 category with 97.8% sequence identity (Fig. 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 1,105, and a 32-aa C-terminal extension of the LRR domain in the R variant (Fig. 7c).

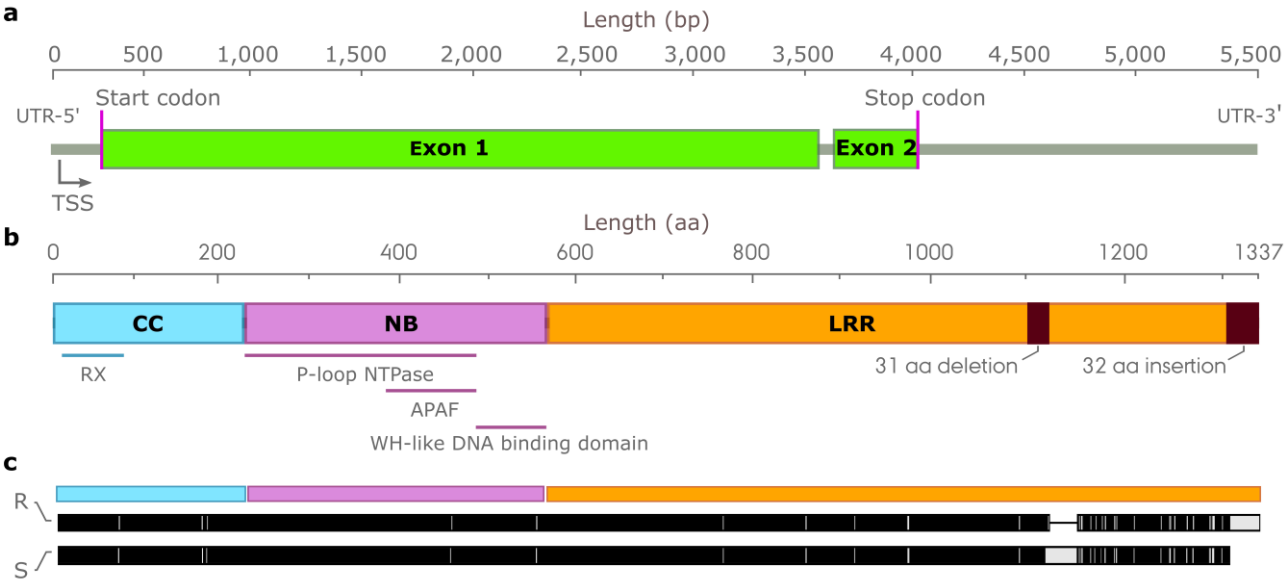


Fig. 7. Characterization of the candidate leaf rust resistance nucleotide-binding leucine-rich repeat (NLR) gene RenRS5_3. (**a)** Gene structure, (**b**) NLR protein motif and domain structure, coiled-coil (CC), nucleotide-binding (NB), leucine-rich repeat (LRR), and (**c**) Alignment of RenRS3_5 resistant (R) and susceptible (S) variants with single amino acid substitutions marked as white lines.

We identified NLR contigs with ≥97.8% similarity with the coding sequence of RenRS5_3 in three lines belonging to the IT categories IT_1 and IT_3. These contigs exhibited a more complex gene structure and encoded a non-canonical NLR protein of 1,314 aa lacking the N-terminal CC domain and showing ≥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 (Fig. 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 ≥80.2% sequence similarity (Fig. 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 identityads with the ‘S’ variant.

Discussion

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 analysed in this study boasts a high level of resistance to leaf rust (Supplementary Table S1, Miedaner et al., 2002). We discovered here five intermediate-effect genomic regions on chromosome arms 1RS, 1RL, 2RL, and 7RS which are known to harbour *Pr* genes in rye (Wehling et al., 2003; Roux et al., 2004; Vendelbo et al., 2021a). 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 between true associations characterized by, a distinct marker peak from false-positive outliers (Gaikpa et al., 2021). 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 a 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) (Korte and Farlow, 2013; Alqudah et al., 2020). 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 to the seed-mother population (Vendelbo et al., 2020). 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 (Wu et al., 2016). 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 (Fig. 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 (Gibson, 2012; Korte and Farlow, 2013; Alqudah et al., 2020). Furthermore, several rye lines also displayed partial resistance against all six SPIs 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 (Sucher et al., 2017). 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.

Test of NLR capture by the bait library

The 60 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. speltoides*) (Steuernagel et al., 2016). 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 (Rabanus-Wallace et al., 2021). During the course of our *in vitro* assessment, we observed a similar requirement for a minimum alignment identity between bait and target for efficient capture to that reported by Jupe et al. (2013), hence 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 observations was further supported by the considerably lower number of NLR annotated contigs per line (with mean of 646 NLRs) compared to the 1,167 NLRs present in the Lo7 reference genome. This discrepancy clearly falls outside the ≤40% expected range accessional size variation of the NLR repertoire reported in previous RenSeq studies in *Ae. tauschii* and Arabidopsis (Arora et al., 2019; Van de Weyer et al., 2019). Our assessment of bait library performance offered a glimpse into NLR capture across the Gülzow rye germplasm. However, additional data and analysis will be 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 (Arora et al., 2019). This technique relies on the capacity of baits to align to a broad span of orthologous NLR genes with divergence sequence from the original NLR bait design library. In the pioneering application of RenSeq by Jupe et al. (2013), 68.5% of the enriched reads corresponded to genes that did not feature in the bait library design, hence demonstrating the plasticity of bait NLR capture.

***k-mer* association genetics with SMRT RenSeq data**

Discovery of *R* genes using association genetics on subsequences depends on the detection of *k*-mer sequences specific to resistant lines, often differentiated by a SNV or InDel in the underlying NLR sequence. In contrast to the pioneering AgRenSeq study using Illumina short-read sequencing (Arora et al., 2019), here we used the updated PacBio® SMRT sequel II HiFi system to achieve a higher precision in SNV and InDel calling (Wenger et al., 2019). When

assessing the effect of sample size on *R* gene detection using AgRenSeq, Arora et al. (2019) concluded that a panel of 80 diverse *Ae. tauschii* accessions is 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 an 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 non-redundant candidate *Pr* genes. Comparative analysis of mapping positions was consistent with two genomic regions identified by GWAS on chromosome arms 1RS and 5RL.

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 centre of a genomic region associated with leaf rust resistance on chromosome arm 1RS (Roux et al., 2004). 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 (Halterman and Wise, 2004; Xie et al., 2020). In flax (*Linum usitatissimum* L.) loss of a repeated unit within the LRR domain of *M* has been associated with inactivation of rust resistance (Anderson et al., 1997). In the Weining reference genome, the paralogous NLR cluster has undergone considerable gene expansion, although one of the paralogues 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 cross-over, intra-cluster rearrangements, and gene conversion events (Barragan and Weigel, 2021). In plants, several NLR clusters form complex loci of multiple paralogous genes encoding NLR proteins with different resistance specificities against distinct races of a certain pathogen (Jones et al., 1993; Wei et al., 1999). In maize (*Zea mays* L.), the rust resistance locus *Rp1* multiple copies of paralogous genes in various haplotype combinations carrying the *Rp1-A* and *Rp1-H* specificities (Smith et al., 2004). The evolutionary capacity of complex NLR loci was highlighted in a study by Richter et al. (1995), 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 likewise explain the presence of the *Pr* gene paralogue identified in the panel, accentuating the rapid generation of novel genetic variation in NLR clusters (Barragan and Weigel, 2021).

Chromosome arm 1RS from 'Petkus' rye has been widely deployed for wheat improvement of rust resistance through chromosomal translocation, carrying amongst others the two leaf rust *R* genes *Pr3* and *Lr26* (Lukaszewski, 1990; Mago et al., 2005). We used the RFLP markers *SCM9* and *Xscm1,* which genetically co-segregate with *Pr3* (Roux et al., 2004), 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 a Russian population of rye variety 'Jaroslavna', segregating as a single dominant locus (Roux et al., 2004). Whether the candidate *Pr* gene is *Pr3,* a paralogue, 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 will provide a valuable genetic resource for the development of functional markers and for future genetic engineering (Dong and Ronald, 2019). With most breeding companies having invested in the infrastructure for MAS, molecular markers can now be easily integrated directly into commercial breeding programs (Miedaner and Korzun, 2012).

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 60 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 nearly leaf rust resistance gene *Pr3* may be the underlying candidate *Pr3* gene or a paralogue or variant thereof.

Acknowledgements

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 solving.

Author contributions:

All authors were involved in the study design. NMV. performed the leaf rust collection, recovery, multiplication and trial inoculation, phenotyping, bioinformatic analysis, visual output and wrote the manuscript. KM assisted in DNA extraction of the germplasm. MSH oversaw the initial collection, recovery and multiplication of *Prs* populations. PSK oversaw seed multiplication. JO was responsible for all communications with Trait Genetics and Eurofins Genomics, who conducted the SNP genotyping. BS mentored NMV in the bioinformatics processing of SMRT RenSeq data and interpretation of results. KM, BS, BBHW, MSH, AFJ, PS, JO and AJ were involved in the intellectual input for the study, including interpretation of results. All authors were involved in the conceptualization of the study and revision of the manuscript.

Competing Interests

The author(s) declare no competing interests.

Financial statement

The research was funded by Innovation Fund Denmark (grant no. 8053-00085B) and Pajbjerg Foundation.

Data availability

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.

References

Alqudah, A.M., Sallam, A., Baenziger, P.S., and Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley–A review. *Journal of Advanced Research* 22**,** 119-135.

Anderson, P.A., Lawrence, G.J., Morrish, B.C., Ayliffe, M.A., Finnegan, E.J., and Ellis, J.G. (1997). Inactivation of the flax rust resistance gene M associated with loss of a repeated unit within the leucine-rich repeat coding region. *The Plant Cell* 9**,** 641-651.

Andreasen, C., and Stryhn, H. (2012). Increasing weed flora in Danish beet, pea and winter barley fields. *Crop Protection* 36**,** 11-17.

Anikster, Y., Bushnell, W., Roelfs, A., Eilam, T., and Manisterski, J. (1997). Puccinia recondita causing leaf rust on cultivated wheats, wild wheats, and rye. *Canadian Journal of Botany* 75**,** 2082-2096.

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., Johnson, R., Enk, J., Periyannan, S., Singh, N., Asyraf Md Hatta, M., Athiyannan, N., Cheema, J., Yu, G., Kangara, N., Ghosh, S., Szabo, L.J., Poland, J., Bariana, H., Jones, J.D.G., Bentley, A.R., Ayliffe, M., Olson, E., Xu, S.S., Steffenson, B.J., Lagudah, E., and Wulff, B.B.H. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature Biotechnology* 37**,** 139-143.

Bajgain, P., Rouse, M.N., Tsilo, T.J., Macharia, G.K., Bhavani, S., Jin, Y., and Anderson, J.A. (2016). Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. *PloS One* 11**,** e0155760.

Barbey, C.R., Lee, S., Verma, S., Bird, K.A., Yocca, A.E., Edger, P.P., Knapp, S.J., Whitaker, V.M., and Folta, K.M. (2019). Disease resistance genetics and genomics in octoploid strawberry. *G3: Genes, Genomes, Genetics* 9**,** 3315-3332.

Barragan, A.C., and Weigel, D. (2021). Plant NLR diversity: the known unknowns of Pan-NLRomes. *The Plant Cell* 33**,** 814-831.

Bauer, E., Schmutzer, T., Barilar, I., Mascher, M., Gundlach, H., Martis, M.M., Twardziok, S.O., Hackauf, B., Gordillo, A., Wilde, P., Schmidt, M., Korzun, V., Mayer, K.F., Schmid, K., Schon, C.C., and Scholz, U. (2017). Towards a whole-genome sequence for rye (Secale cereale L.). *The Plant Journal* 89**,** 853-869.

Beukert, U., Thorwarth, P., Zhao, Y., Longin, C.F.H., Serfling, A., Ordon, F., and Reif, J.C. (2020). Comparing the potential of marker-assisted selection and genomic prediction for improving rust resistance in hybrid wheat. *Frontiers in Plant Science* 11.

De Mol, F., Von Redwitz, C., and Gerowitt, B. (2015). Weed species composition of maize fields in G ermany is influenced by site and crop sequence. *Weed Research* 55**,** 574-585.

Dong, O.X., and Ronald, P.C. (2019). Genetic engineering for disease resistance in plants: recent progress and future perspectives. *Plant Physiology* 180**,** 26-38.

Figueroa, M., Dodds, P.N., and Henningsen, E.C. (2020). Evolution of virulence in rust fungi - multiple solutions to one problem. *Current Opinion in Plant Biology* 56**,** 20-27.

Florence, J., Vernaldi, S., and Maekawa, T. (2013). Evolution and conservation of plant NLR functions. *Frontiers in Immunology* 4**,** 297.

Friebe, B., Jiang, J., Raupp, W.J., Mcintosh, R.A., and Gill, B.S. (1996). Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91**,** 59-87.

Gaikpa, D.S., Kessel, B., Presterl, T., Ouzunova, M., Galiano-Carneiro, A.L., Mayer, M., Melchinger, A.E., Schön, C.-C., and Miedaner, T. (2021). Exploiting genetic diversity in two European maize landraces for improving Gibberella ear rot resistance using genomic tools. *Theoretical and Applied Genetics* 134**,** 793-805.

Geneiousprime (2020). Geneious Prime (v. 2020.2.3). https://www.geneious.com.

Ghosh, S., Watson, A., Gonzalez-Navarro, O.E., Ramirez-Gonzalez, R.H., Yanes, L., Mendoza-Suárez, M., Simmonds, J., Wells, R., Rayner, T., and Green, P. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nature Protocols* 13**,** 2944-2963.

Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics* 13**,** 135-145.

Halterman, D.A., and Wise, R.P. (2004). A single-amino acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. *The Plant Journal* 38**,** 215-226.

Hanzlik, K., and Gerowitt, B. (2012). Occurrence and distribution of important weed species in German winter oilseed rape fields. *Journal of Plant Diseases and Protection* 119**,** 107-120.

Hickey, L.T., Germán, S.E., Pereyra, S.A., Diaz, J.E., Ziems, L.A., Fowler, R.A., Platz, G.J., Franckowiak, J.D., and Dieters, M.J. (2017). Speed breeding for multiple disease resistance in barley. *Euphytica* 213**,** 64.

Hovmoller, M.S., Rodriguez-Algaba, J., Thach, T., and Sorensen, C.K. (2017). "Race typing of Puccinia striiformis on wheat," in *Wheat Rust Diseases: Methods and Protocols*. (New York, NY: Humana Press), 29-40.

Hovmoller, M.S., Yahyaoui, A.H., Milus, E.A., and Justesen, A.F. (2008). Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* 17**,** 3818-3826.

Jones, D., Dickinson, M., Balint-Kurti, P., Dixon, M., and Jones, J. (1993). 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. *Molecular Plant Microbe Interactions* 6**,** 348-348.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., Mcanulla, C., and Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30**,** 1236-1240.

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G.J., Maclean, D., Cock, P.J., Leggett, R.M., Bryan, G.J., Cardle, L., Hein, I., and Jones, J.D. (2013). 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. *The Plant Journal* 76**,** 530-544.

Kolmer, J.A., Ordonez, M.E., German, S., Morgounov, A., Pretorius, Z., Visser, B., Goyeau, H., Anikster, Y., and Acevedo, M. (2019). Multilocus genotypes of the wheat leaf rust fungus puccinia triticina in worldwide regions indicate past and current long-distance migration. *Phytopathology* 109**,** 1453-1463.

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9**,** 1-9.

Kourelis, J., and Van Der Hoorn, R.A. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *The Plant Cell* 30**,** 285-299.

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* 47**,** 256-259.

Li, G., Wang, L., Yang, J., He, H., Jin, H., Li, X., Ren, T., Ren, Z., Li, F., Han, X., Zhao, X., Dong, L., Li, Y., Song, Z., Yan, Z., Zheng, N., Shi, C., Wang, Z., Yang, S., Xiong, Z., Zhang, M., Sun, G., Zheng, X., Gou, M., Ji, C., Du, J., Zheng, H., Dolezel, J., Deng, X.W., Stein, N., Yang, Q., Zhang, K., and Wang, D. (2021). A high-quality genome assembly highlights rye genomic characteristics and agronomically important genes. *Nature Genetics* 53**,** 574-584.

Lipka, A.E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P.J., Gore, M.A., Buckler, E.S., and Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28**,** 2397-2399.

Liu, R., Lu, J., Zhou, M., Zheng, S., Liu, Z., Zhang, C., Du, M., Wang, M., Li, Y., and Wu, Y. (2020). Developing stripe rust resistant wheat (Triticum aestivum L.) lines with gene pyramiding strategy and marker-assisted selection. *Genetic Resources and Crop Evolution* 67**,** 381-391.

Lukaszewski, A.J. (1990). Frequency of 1RS. 1AL and 1RS. 1BL translocations in United States wheats. *Crop Science* 30**,** 1151-1153.

Maekawa, T., Cheng, W., Spiridon, L.N., Töller, A., Lukasik, E., Saijo, Y., Liu, P., Shen, Q.-H., Micluta, M.A., and Somssich, I.E. (2011). Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host & Microbe* 9**,** 187-199.

Mago, R., Miah, H., Lawrence, G.J., Wellings, C.R., Spielmeyer, W., Bariana, H.S., Mcintosh, R.A., Pryor, A.J., and Ellis, J.G. (2005). High-resolution mapping and mutation analysis separate the rust resistance genes Sr31, Lr26 and Yr9 on the short arm of rye chromosome 1. *Theoretical and Applied Genetics* 112**,** 41-50.

Marcais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27**,** 764-770.

Martin, E.C., Sukarta, O.C., Spiridon, L., Grigore, L.G., Constantinescu, V., Tacutu, R., Goverse, A., and Petrescu, A.-J. (2020). LRRpredictor—a new LRR motif detection method for irregular motifs of plant NLR proteins using an ensemble of classifiers. *Genes* 11**,** 286.

Mcintosh, R.A., Frieb, B., Jiang, J., and Gill, B.S. (1995). Cytogenetical studies in wheat XVI. Chromosomal location of a new gene for resistance to leaf rust in a Japanese wheat-rye translocation line. *Euphytica* 82**,** 141-147.

Mcneal, F.H., Konzak, C.F., Smith, E.P., Tate, W.S., and Russell, T.S. (1971). A uniform system for recording and processing cereal research data. *No. REP-10904. CIMMYT*.

Miedaner, T., Gey, A.-K.M., Sperling, U., and Geiger, H.H. (2002). Quantitative-genetic analysis of leaf-rust resistance in seedling and adult-plant stages of inbred lines and their testcrosses in winter rye. *Plant Breeding* 121**,** 475-479.

Miedaner, T., Klocke, B., Flath, K., Geiger, H.H., and Weber, W.E. (2011). Diversity, spatial variation, and temporal dynamics of virulences in the German leaf rust (Puccinia recondita f. sp. secalis) population in winter rye. *European Journal of Plant Pathology* 132**,** 23-35.

Miedaner, T., and Korzun, V. (2012). Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* 102**,** 560-566.

Miedaner, T., and Sperling, U. (1995). Effect of leaf rust and yield components of winter rye hybrids and assessment of quantitative resistance. *Journal of Phytopathology* 143**,** 725-730.

Monino-Lopez, D., Nijenhuis, M., Kodde, L., Kamoun, S., Salehian, H., Schentsnyi, K., Stam, R., Lokossou, A., Abd-El-Haliem, A., and Visser, R.G. (2021). 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. *The Plant Journal*.

Ncbi (2021). *National Center for Biotechnology Information* [Online]. Bethesda (MD): National Library of Medicine, USA. Available: https://www.ncbi.nlm.nih.gov [Accessed 2021 13/6].

Nelson, R., Wiesner-Hanks, T., Wisser, R., and Balint-Kurti, P. (2018). Navigating complexity to breed disease-resistant crops. *Nature Reviews Genetics* 19**,** 21-33.

Nurk, S., Walenz, B.P., Rhie, A., Vollger, M.R., Logsdon, G.A., Grothe, R., Miga, K.H., Eichler, E.E., Phillippy, A.M., and Koren, S. (2020). HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *bioRxiv*.

Pallotta, M.A., Warner, P., Fox, R.L., Kuchel, H., Jefferies, S.J., and Langridge, P. (2003). Marker assisted wheat breeding in the southern region of Australia. *Proceedings of the Tenth International Wheat Genetics Symposium***,** 789-791.

Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35**,** 526-528.

Qi, D., Deyoung, B.J., and Innes, R.W. (2012). Structure-function analysis of the coiled-coil and leucine-rich repeat domains of the RPS5 disease resistance protein. *Plant Physiology* 158**,** 1819-1832.

R Core Team (2021). *R a language and environment for statistical computing. R foundation for statistical Computing, Vienna, Austria* [Online]. Available: https://www.R-project.org/ [Accessed 2021 8/2].

Rabanus-Wallace, M.T., Hackauf, B., Mascher, M., Lux, T., Wicker, T., Gundlach, H., Báez, M., Houben, A., Mayer, K.F.X., Guo, L., Poland, J., Pozniak, C.J., Walkowiak, S., Melonek, J., Praz, C., Schreiber, M., Budak, H., Heuberger, M., Steuernagel, B., Wulff, B., Börner, A., Byrns, B., Čížková, J., Fowler, D.B., Fritz, A., Himmelbach, A., Kaithakottil, G., Keilwagen, J., Keller, B., Konkin, D., Larsen, J., Li, Q., Myśków, B., Padmarasu, S., Rawat, N., Sesiz, U., Sezgi, B., Sharpe, A., Šimková, H., Small, I., Swarbreck, D., Toegelová, H., Tsvetkova, N., Voylokov, A.V., Vrána, J., Bauer, E., Bolibok-Bragoszewska, H., Doležel, J., Hall, A., Jia, J., Korzun, V., Laroche, A., Ma, X.-F., Ordon, F., Özkan, H., Rakoczy-Trojanowska, M., Scholz, U., Schulman, A.H., Siekmann, D., Stojałowski, S., Tiwari, V., Spannagl, M., and Stein, N. (2021). Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. *Nature Genetics* 53**,** 564-573.

Richter, T.E., Pryor, T.J., Bennetzen, J.L., and Hulbert, S.H. (1995). New rust resistance specificities associated with recombination in the Rp1 complex in maize. *Genetics* 141**,** 373-381.

Rollar, S., Serfling, A., Geyer, M., Hartl, L., Mohler, V., and Ordon, F. (2021). QTL mapping of adult plant and seedling resistance to leaf rust (Puccinia triticina Eriks.) in a multiparent advanced generation intercross (MAGIC) wheat population. *Theoretical and Applied Genetics* 134**,** 37-51.

Roux, S.R., Hackauf, B., Linz, A., Ruge, B., Klocke, B., and Wehling, P. (2004). Leaf-rust resistance in rye (Secale cereale L.). 2. Genetic analysis and mapping of resistance genes Pr3, Pr4, and Pr5. *Theoretical and Applied Genetics* 110**,** 192-201.

Rstudio Team (2015). *Rstudio: integrated development for R. RStudio, Inc., Boston* [Online]. Available: http://www.rstudio.com [Accessed 2021 2/8].

Seong, K., Seo, E., Witek, K., Li, M., and Staskawicz, B. (2020). Evolution of NLR resistance genes with noncanonical N-terminal domains in wild tomato species. *New Phytologist* 227**,** 1530-1543.

Smith, S.M., Pryor, A.J., and Hulbert, S.H. (2004). Allelic and haplotypic diversity at the rp1 rust resistance locus of maize. *Genetics* 167**,** 1939-1947.

Sortsinfo (2021). *The Danish official trial records* [Online]. Available: https://sortinfo.dk/ [Accessed 10/2 2021].

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research* 32**,** W309-312.

Steuernagel, B., Jupe, F., Witek, K., Jones, J.D., and Wulff, B.B. (2015). NLR-parser: rapid annotation of plant NLR complements. *Bioinformatics* 31**,** 1665-1667.

Steuernagel, B., Periyannan, S.K., Hernandez-Pinzon, I., Witek, K., Rouse, M.N., Yu, G., Hatta, A., Ayliffe, M., Bariana, H., Jones, J.D., Lagudah, E.S., and Wulff, B.B. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology* 34**,** 652-655.

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., Jones, J.D.G., Uauy, C., Keller, B., Ridout, C.J., and Wulff, B.B.H. (2020). The NLR-annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiology* 183**,** 468-482.

Sucher, J., Boni, R., Yang, P., Rogowsky, P., Buchner, H., Kastner, C., Kumlehn, J., Krattinger, S.G., and Keller, B. (2017). The durable wheat disease resistance gene Lr34 confers common rust and northern corn leaf blight resistance in maize. *Plant Biotechnology Journal* 15**,** 489-496.

Thach, T., Ali, S., Justesen, A., Rodriguez‐Algaba, J., and Hovmøller, M. (2015). Recovery and virulence phenotyping of the historic 'Stubbs collection'of the yellow rust fungus Puccinia striiformis from wheat. *Annals of Applied Biology* 167**,** 314-326.

Toparslan, E., Karabag, K., and Bilge, U. (2020). A workflow with R: Phylogenetic analyses and visualizations using mitochondrial cytochrome b gene sequences. *PLoS One* 15**,** e0243927.

Uniprot, C. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research* 49**,** 480-489.

Usda (2006). *United States Department of Agriculture: wheat and barley DNA extraction protocol (96-well plate format)* [Online]. Available: https://www.ars.usda.gov/ARSUserFiles/60701500/SmallGrainsGenotypingLaboratory/Protocols/wheat%20and%20barleyDNA%20extraction_original.pdf [Accessed 2021 7/7].

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., and Bemm, F. (2019). A species-wide inventory of NLR genes and alleles in Arabidopsis thaliana. *Cell* 178**,** 1260-1272. e1214.

Vendelbo, N.M., Mahmood, K., Sarup, P., Orabi, J., Hovmøller, M., Justesen, A.F., Kristensen, P.S., and Jahoor, A. (2021a). Discovery of a novel leaf rust (Puccinia recondita) resistance in rye (Secale cereale L.). *MDPI Cells*.

Vendelbo, N.M., Mahmood, K., Sarup, P., Orabi, J., Kristensen, P.S., and Jahoor, A. (2021b). Discovery of a novel powdery mildew (Blumeria graminis) resistance locus in rye (Secale cereale L.). *Scientific reports***,** In press.

Vendelbo, N.M., Sarup, P., Orabi, J., Kristensen, P.S., and Jahoor, A. (2020). Genetic structure of a germplasm for hybrid breeding in rye (Secale cereale L.). *PLoS One* 15**,** e0239541.

Wehling, P., Linz, A., Hackauf, B., Roux, S.R., Ruge, B., and Klocke, B. (2003). Leaf-rust resistance in rye (Secale cereale L.). 1. Genetic analysis and mapping of resistance genes Pr1 and Pr2. *Theoretical and Applied Genetics* 107**,** 432-438.

Wei, F., Gobelman-Werner, K., Morroll, S.M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P., and Wise, R.P. (1999). 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* 153**,** 1929-1948.

Wenger, A.M., Peluso, P., Rowell, W.J., Chang, P.-C., Hall, R.J., Concepcion, G.T., Ebler, J., Fungtammasan, A., Kolesnikov, A., and Olson, N.D. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology* 37**,** 1155-1162.

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics* 3**,** 180-185.

Williams, S.J., Sornaraj, P., Decourcy-Ireland, E., Menz, R.I., Kobe, B., Ellis, J.G., Dodds, P.N., and Anderson, P.A. (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant-Microbe Interactions* 24**,** 897-906.

Witek, K., Jupe, F., Witek, A.I., Baker, D., Clark, M.D., and Jones, J.D. (2016). Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology* 34**,** 656-660.

Wu, Y., San Vicente, F., Huang, K., Dhliwayo, T., Costich, D.E., Semagn, K., Sudha, N., Olsen, M., Prasanna, B.M., Zhang, X., and Babu, R. (2016). Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theoretical and Applied Genetics* 129**,** 753-765.

Xie, J., Guo, G., Wang, Y., Hu, T., Wang, L., Li, J., Qiu, D., Li, Y., Wu, Q., and Lu, P. (2020). A rare single nucleotide variant in Pm5e confers powdery mildew resistance in common wheat. *New Phytologist* 228**,** 1011-1026.

Xing, L., Hu, P., Liu, J., Witek, K., Zhou, S., Xu, J., Zhou, W., Gao, L., Huang, Z., and Zhang, R. (2018). Pm21 from Haynaldia villosa encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Molecular Plant* 11**,** 874-878.

Yu, G. (2020). Using ggtree to visualize data on tree-like structures. *Current Protocols in Bioinformatics* 69**,** e96.

Discovery of resistance genes in rye by long-read target sequencing and association genetics

Nikolaj M. Vendelbo[1], Khalid Mahmood, Burkhard Steuernagel, Brande B.H. Wulff, Pernille Sarup, Mogens Hovmøller, Annemarie F. Justesen, Peter S. Kristensen, Jihad Orabi, and Ahmed Jahoor

Supplementary material

Caption of all supplementary files has been given below and a subset of supplementary figures and tables included in the dissertation. The extended supplementary materials has been provided at an open-access Zenodo data repository (https://zenodo.org/record/5510300).

Table S1: Leaf rust severity of top-yielding hybrid rye (*Secale cereale* L.) and hybrid rye-population mixtures tested in the Danish official trials in 2019 across nine locations.

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.

Isolate	Country	Region	Location	Survey site	Latitude	Longitude	Collection date
SPI-1	Denmark	Zealand	Flakkebjerg	Trial	55,3161	11,3686	05-07-2018
SPI-2	Denmark	Jutland	Dyngby	Trial	55,9527	10,2414	10-07-2018
SPI-3	Denmark	Lolland	Holeby	Trial	54,7001	11,4488	04-10-2018
SPI-4	Denmark	Jutland	Peterslyst	Farmers field	56,0810	10,0838	18-05-2019
SPI-5	Denmark	Jutland	Gylling	Farmers field	55,8944	10,1688	20-05-2019
SPI-6	Denmark	Fyn	Vissenbjerg	Farmers field	55,3949	10,1839	20-05-2019

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, SPI-3) or 2019 (SPI-4, SPI-5, SPI-6).

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.

Table S5. Metrics of *in silico* assessment of the 60 K Triticeae nucleotide-binding leucine-rich repeat (NLR) bait library on the rye (*Secale cereal* L.) reference genome Lo7, and *in vitro* assessment in 20 inbred rye lines analyzed using single-molecule real-time resistance gene enrichment sequencing.

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.

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.

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.

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.

Datafile S1. Genotype information for 261,406 informative 600 K SNP array markers in 92 inbred rye (*Secale cereale* L.) lines (.Rdata).

Fig. S1. Geographical location of the origin of the six *Puccinia recondita* f. sp. *secalis* single pustule isolates collected in Denmark in 2018 and 2019.

Fig. 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.

Fig. S3. Characteristics of NLR target sequencing in 20 inbred rye (*Secale cereale* L.) lines using a 60K 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.

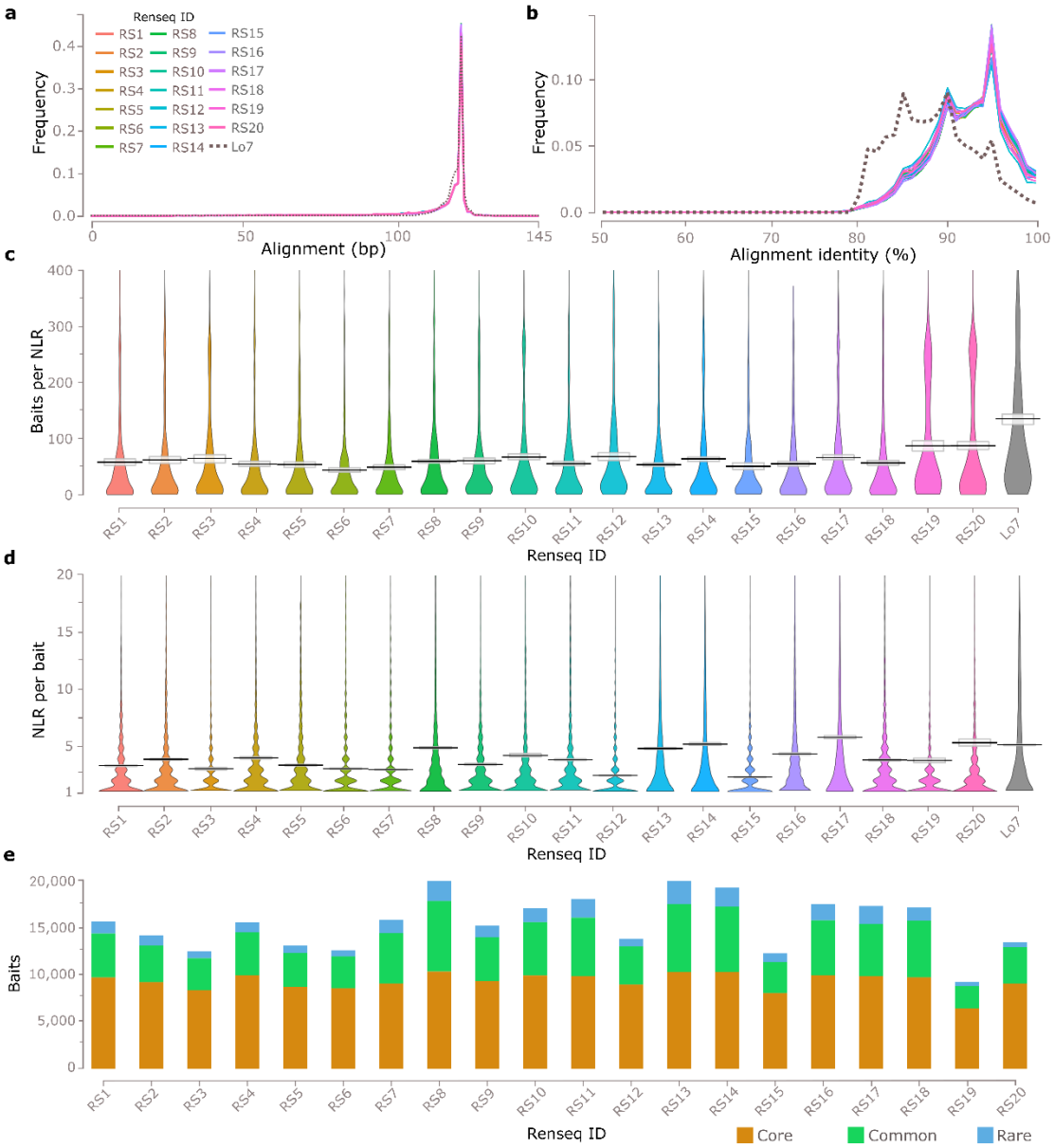


Fig. S4. Identification of resistance associated nucleotide-binding leucine-rich repeat (NLR) annotated contig across 20 inbred rye (*Secale cereale* L.) lines by association genetics resistance gene enrichment sequencing (AgRenSeq) analysis. (**a**) Standard AgRenSeq with candidate leaf rust resistance gene highlighted in red. (**b**), AgRenSeq with NLR contigs mapped to the Lo7 reference genome. (**c**), AgRenSeq-GLM.

Fig. S5. Association genetics resistance gene enrichment sequencing (AgRenSeq) analysis for the identification of NLR contigs harbouring resistant-specific *k*-mers in 20 inbred rye (*Secale cereale* L.) lines.

Fig. S6. Association genetics resistance gene enrichment sequencing (AgRenSeq) GLM analysis for the identification of NLR contigs harbouring resistant-specific *k*-mers in 20 inbred rye (*Secale cereale* L.) lines.

Fig. 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.

Data repository: All scripts and data needed to reproduce the visual analysis conducted in the manuscript

Manuscript V

Discovery of a novel powdery mildew (*Blumeria graminis*) resistance gene in rye (*Secale cereale* L.)

By

Nikolaj M. Vendelbo[1,2], Khalid Mahmood [1], Pernille Sarup [1], Peter S. Kristensen [1], Jihad Orabi [1], and Ahmed Jahoor [1,3]

[1] Nordic Seed A/S, Odder, Denmark

[2] Department of Agroecology, Aarhus University, Slagelse, Denmark

[3] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden

Publication details

Published in Nature Scientific Reports

scientific reports

Check for updates

OPEN

Discovery of a novel powdery mildew (*Blumeria graminis*) resistance locus in rye (*Secale cereale* L.)

N. M. Vendelbo[1,2]✉, K. Mahmood[1], P. Sarup[1], P. S. Kristensen[1], J. Orabi[1] & A. Jahoor[1,3]

Powdery mildew is one of the most destructive diseases in the world, causing substantial grain yield losses and quality reduction in cereal crops. At present 23 powdery mildew resistance genes have been identified in rye, of which the majority are in wheat-rye translocation lines developed for wheat improvement. Here, we investigated the genetics underlying powdery mildew resistance in the Gülzow-type elite hybrid rye (*Secale cereale* L.) breeding germplasm. In total, 180 inbred breeding lines were genotyped using the state-of-the-art 600 K SNP array and phenotyped for infection type against three distinct field populations of *B. graminis* f. sp. *secalis* from Northern Germany (2013 and 2018) and Denmark (2020). We observed a moderate level of powdery mildew resistance in the non-restorer germplasm population, and by performing a genome-wide association study using 261,406 informative SNP markers, we identified a powdery mildew resistance locus, provisionally denoted *PmNOS1*, on the distal tip of chromosome arm 7RL. Using recent advances in rye genomic resources, we investigated whether nucleotide-binding leucine-rich repeat genes residing in the identified 17 Mbp block associated with *PmNOS1* on recent reference genomes resembled known *Pm* genes.

Powdery mildew (PM) is one of the most devasting diseases globally[4]. In cereals, the causative agent of PM is the ascomycete fungus *Blumeria graminis* (DC.) speer (*Bg*), which is capable of inflicting severe grain yield loss (≥ 20%) and quality reduction in cereals[5–8]. In periods of conducive conditions, such as frequent precipitation and low to moderate temperatures, *Bg* can cause severe epidemics by repeated infections and clonal reproduction[9]. As an obligate biotroph, *Bg* is highly specialized to its host species, divided into several distinct *'formae specialis'* (f. sp.)[10], and dependent on a living host for survival and reproduction. In the later stages of the growing season, or in periods of unconducive conditions, chasmothecia structures are formed[11]. The chasmothecia field inoculum facilitates the infection of volunteer plants or the successive autumn sown winter crop, allowing *Bg* to overwinter as dormant mycelia[12]. Long-distance dispersal by wind of conidiospores drives the large spatial variability in the *Bg* population[13–15]. In barley (*Hordeum vulgare* L.), the virulence complexity of *B. graminis* f. sp. *hordei* has been observed to increase with the prevailing wind direction from west to east in Europe[16]. Rather than geographical origin, evidence suggests that local use of host resistance and fungicides are the primary factors influencing the European *B. graminis* f. sp. *hordei* population[17]. These findings underline the seriousness of long-distance dispersal of *B. graminis* spores, which allows novel and aggressive pathotypes to evolve through the recombination of distinct pathotypes and rapidly spread[18]. In rye, the causative agent of PM, *B. graminis* f. sp. *secalis*, has attracted little scientific interest in recent years[19,20].

To reduce our dependency on pesticides, host resistance constitutes a sustainable alternative for farmers to ensure crop productivity[21]. At present, 23 major PM resistance (*R*) genes have been identified in rye (Table 1). However, it is likely that a subset of colocalized *R* genes are allelic, as there was no evidence that these genes were distinct from existing *R* genes prior to denotation.

The majority of characterized *Pm* genes encode intracellular nucleotide-binding leucine-rich repeat (NLR) proteins that recognize pathogen effector molecules, leading to an effector-triggered immunity resistance response[43–45]. Canonical NLR genes are composed of three domains[46]. In grasses, the N-terminus is composed of a coiled-coil (CC) domain believed to be involved in signaling and the induction of cell death[47]. In the center, a nucleotide-binding adaptor shared by APAF-1, R proteins, and the CED-4 (NB-ARC) domain functions as a

[1]Nordic Seed A/S, Odder, Denmark. [2]Department of Agroecology, Faculty of Technical Sciences, Aarhus University, Slagelse, Denmark. [3]Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden. ✉email: nive@nordicseed.com

Chromosome	Arm	Gene name		References
		Rye	Wheat	
1R	Short	*Pm1*	*Pm8*	[22]
		Pm1	*Pm17*	[23]
		Pm1	*PmCn17*	[24]
		Pm		[2]
		Pm		[25]
	Long	*PmSESY*		[26]
	ND	*Pm7*	–	[27]
2R	Long	*Pm2*	–	[28]
		–	*PmJZHM2RL*	[29]
		Pm	–	[30,31]
	ND	–	*Pm7*	[32]
		Pm8	*Pm3*	[31]
		Pm		[33]
3R	Short	*Pm3*	–	[34]
4R	Long	*Pm*	–	[35,36]
	ND	*Pm*	–	[37]
		Pm6	–	[34]
5R	Long	*Pm4*	–	[38]
6R	Short	–	*Pm56*	[39]
	Long	–	*Pm20*	[40]
		Pm5	–	[41]
		Pm	–	[31,36]
7R	Long	*Pm*	–	[42]

Table 1. Location and origin of powdery mildew resistance genes in rye (*Secale cereale* L.) and wheat-(*Triticum aestivum* L.) rye translocation lines with rye being the donor parent. After[1–3]. ND: Interchromosomal position not determined.

regulatory domain determining the protein activation state[48]. Last, in the C-terminus, a leucin-rich repeat (LRR) domain is involved in effector recognition[49]. In rye, 770 canonical NLR genes have been identified in the 'Lo7' reference genome[50].

Rye has been an important source for the improvement of PM resistance in wheat (*Triticum aestivum* L.) by chromosomal translocation of segments housing *R* genes[3,25,42]. While PM is effectively controlled by host genetic resistance in cereals, it has historically been rapidly overcome by virulent races of *Bg*[51,52]. Currently, several of the top-yielding hybrid rye cultivars examined in official Danish trials are susceptible to PM disease (Supplementary Table S1)[53].

Genomic-based breeding techniques have, however, dramatically accelerated the introgression and pyramiding of several *R* genes for enhanced resistance durability[54–56]. Recent advances in genomic resources available in rye, including the 600 K SNP array and chromosomal-scale reference genome of the German inbred winter rye line 'Lo7' and Chinese population rye variety 'Weining', constitute significant milestones in rye genomic breeding[50,57,58]. Continuous mining for the discovery of novel genetic variability in rye *R* genes is essential to expand the 'toolset' available for resistance breeding. The importance of genetic studies in rye is further supported by the possibility of introgressing novel *R* genes into the staple cereal wheat by chromosomal translocation lines[3].

In this paper, we investigate PM resistance in a less-prevalent elite Gülzow-type hybrid rye breeding germplasm. Our objective was to I) characterize PM resistance in the assayed germplasm, II) identify PM resistance-associated SNP markers to be implemented by marker-assisted selection for breeding novel resistant hybrid rye cultivars, III) investigate whether NLR genes residing in PM resistance-associated blocks on the 'Lo7' and 'Weining' reference genomes resemble known *Pm* genes, and IV) develop a marker map for the 600 K high-density SNP array and validate its performance in the assayed germplasm.

Results

600 K SNP genotyping of panel. To investigate the genetics underlying powdery mildew (PM) resistance, the assayed hybrid rye breeding germplasm was genotyped on the rye 600 K SNP array. With only scaffold positional data available for the array, SNP marker sequences were anchored to the recent 'Lo7' rye reference genome and stringently filtered to ensure its accuracy. In total, 591,196 markers were successfully mapped to the reference genome, and the developed marker map was made available at https://doi.org/10.5281/ZENODO.5086235. Quality filtration of markers for low minor allele frequency, missing markers, and missing individual scores across the panel led to the identification of 261,406 informative markers (Supplementary material 1). Characterization of fundamental performance-related metrics revealed a homogeneous inter- and intrachromosomal distribution of markers (Table 2). On average, each chromosome housed 32,676 markers with a mean

Chromosome	Chromosome length (Mbp)	Markers	Informative SNP markers			
			Markers	Mean inter- marker distance ± SD (kb)	Largest inter- marker distance (Mbp)	PIC ± SD
1R	727.33	72,089	33,854	21.47 ± 54.7	2.43	0.204 ± 0.134
2R	945.85	77,774	33,698	28.07 ± 96.5	9.95	0.240 ± 0.120
3R	965.54	69,428	31,493	30.65 ± 75.4	3.60	0.238 ± 0.111
4R	906.54	81,652	32,555	27.84 ± 75.5	2.98	0.230 ± 0.125
5R	876.06	81,842	37,073	23.63 ± 66.9	3.78	0.238 ± 0.123
6R	885.15	84,283	36,872	24.00 ± 62.2	3.31	0.249 ± 0.105
7R	889.76	86,994	38,918	23.12 ± 56.6	2.31	0.231 ± 0.120
Unmapped	–	35,245	16,943	–	–	0.239 ± 0.123
Mean	886.59	74,869	32,676	25,54 ± 69.7	4.05	0.234 ± 0.121

Table 2. Characteristics of informative 600 K SNP array on the Nordic Seed hybrid rye (*Secale cereale* L.) elite breeding germplasm (n = 180). Markers were positioned on the 'Lo7' reference genome. SD: Standard deviation, PIC: Polymorphism information content.

marker-to-marker distance of 25.54 kb. The largest marker-to-marker distance was 9.95 Mbp on chromosome 2R, with a mean of 4.05 Mbp across the chromosomes. As a quality parameter, the polymorphism information content (PIC) was calculated to estimate the ability of markers to detect polymorphisms within the assayed germplasm (Supplementary Table S2). Across the informative marker panel, a mean PIC of 0.234 was identified, with a mean interchromosomal PIC ranging from 0.204 to 0.249 (Table 2). Visualization of these array performance metrics along the rye genome using Circos revealed a drop in marker density and PIC across the pericentromeric region on all chromosomes (Fig. 1).

Phenotyping of hybrid rye breeding germplasm. To provide a comprehensive phenotypic dataset of PM resistance in the assayed Gülzow-type hybrid rye breeding germplasm, the lines were scored for their infection type (IT) against three distinct *Blumeria graminis* f. sp. *secalis* (*Bgs*) populations (Table 3, Supplementary Table S3). The lines scoring an IT below 1 were considered 'resistant' (Supplementary Fig. S1). Across the assayed germplasm, the N13 (Nienstädt, 2013) *Bgs* population yielded a mean IT of 2.40 ± 1.28 standard deviations (SD), with 47 resistant lines out of which 3 were restorers. The N18 (Nienstädt, 2018) *Bgs* population yielded a mean IT of 2.71 ± 1.11 SD with 29 resistant lines, out of which 1 was a restorer. D20 (Dyngby, 2020) *Bgs* population yielded a mean IT of 2.74 ± 1.10 SD with 20 resistant lines out of which 1 was a restorer. Across the assayed germplasms, 20 out of 88 non-restorer germplasm and 1 out of 92 restorer lines were consistently resistant to all three *Bgs* populations. Both controls, hybrid cv. KWS Binntto ('susceptible') and KWS Serafino ('resistant') were susceptible to all three *Bgs* populations. KWS Binntto had a mean IT of 3.56 ± 0.54 SD and KWS Serafino 3.01 ± 1.01 SD across *Bgs* populations.

To visualize the resistance spectrum of breeding lines, a circular neighbor-joining dendrogram was constructed, and concentric circles were added to integrate the scored IT (Fig. 2).

Statistical analysis of the line infection-type distribution across *Bgs* population by ANOVA showed that the N13 *Bgs* population differed significantly from the N18 and D20 *Bgs* populations ($p_{val} < 0.00016$) (Fig. 3). Nine non-restorer germplasm lines were found to exhibit a differential resistance response to the N13 *Bgs* population. These lines were categorized as 'partially resistant', with a mean infection type of 1.52 ± 0.58 SD; they exhibited a mean infection type of 2.74 ± 0.70 and 2.80 ± 0.77 against the N18 and D20 populations, respectively.

Genome-wide association study. For the identification of SNP markers associated with PM resistance, GWAS using MLM and the BLINK method was used for each of the three *Bgs* population trials and across trials for the entire germplasm and individual parental populations (Supplementary Figs. S2, S3). Isolated single SNPs in GWAS-MLM results were removed from the analysis even if they were significantly associated, as these were interpreted as having spurious associations or incorrect mapping positions. Instead, dense peaks comprising a large number of significant associated markers within a confined region in GWAS-MLM were selected for further analysis. GWAS-MLM on the entire panel led to the identification of a haplotype block on chromosome arm 7RL that was significantly associated ($-\log_{10} = 14.6$) with powdery mildew resistance spanning from 882 to 898 Mbp (Fig. 4B, C). The haplotype block harbored 244 markers exhibiting an association above the Bonferroni adjusted significance threshold ($-\log_{10} \geq 6.7$) (Supplementary Table S4. Successive GWAS-BLINK led to the identification of the top-most resistance-associated ($-\log_{10} = 37.9$) marker within the haplotype block on chromosome arm 7RL at 892.09 Mbp, which explained 16.8% of the phenotypic variance (Fig. 4A).

In the non-restorer germplasm population, lines were found to carry a highly conserved haplotype across the significantly associated markers on chromosome arm 7RL (Supplementary Table S6). A resistant haplotype was conserved in 42 out of 48 resistant lines and a susceptible haplotype was found in 28 out of 31 susceptible lines, and the remaining 9 lines showed a differential resistance response (Supplementary material S3).

GWAS using BLINK led to the disappearance of the marker position on the 'Unmapped' chromosome that was in linkage disequilibrium with the genomic region on chromosome arm 7RL (Fig. 4A, B). In the non-restorer germplasm, an additional nonsignificant ($-\log_{10} = 3.62$) peak was identified in GWAS-MLM spanning from

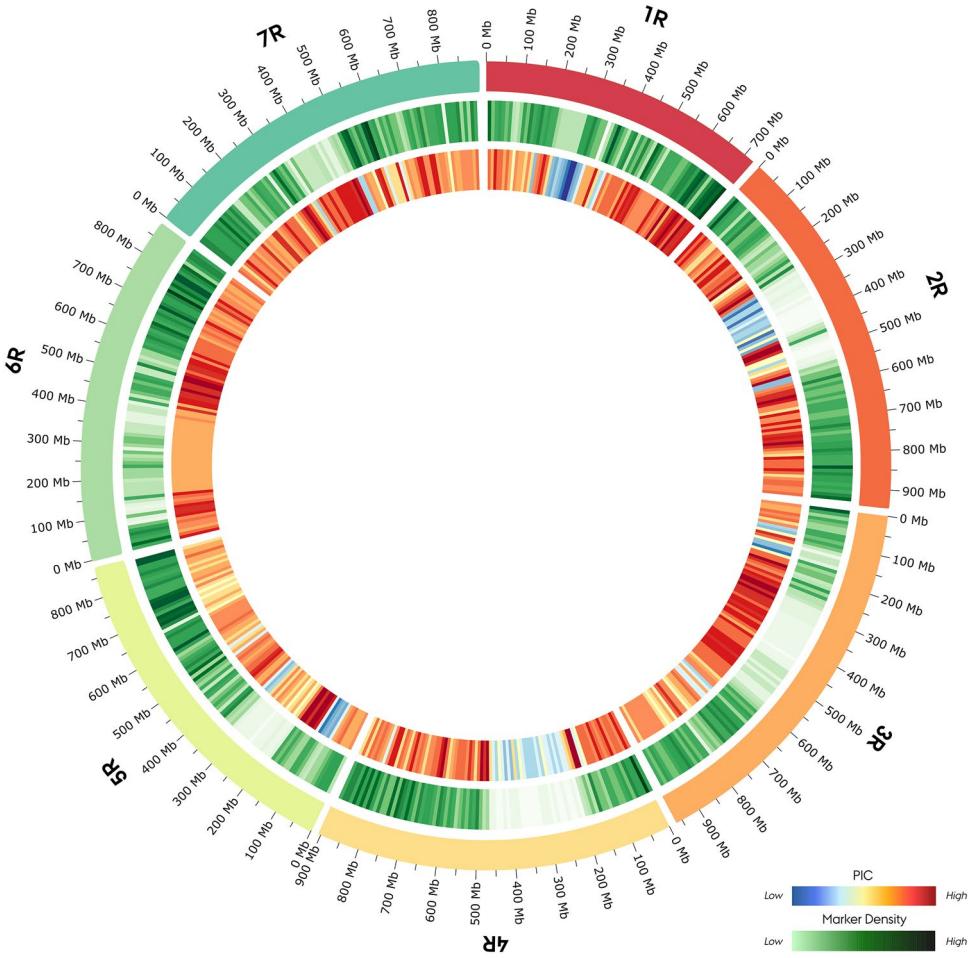


Figure 1. Distribution of 244,463 informative 600 K SNP array markers in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines along the 'Lo7' reference genome in 10 mb bins. The outer track depicts the marker density per bin. The inner track depicts the polymorphism-information content (PIC) per bin.

Phenotype	Infection-type	Mycelium growth	Sporulation	Development of chlorosis/necrosis
Resistant	0	None	None	No
	0–1	None	None	Yes
	1	Weak	None	Yes
Partially resistant	1–2	Weak	Weak	Yes
	2	Moderate	Weak	Yes
Partially susceptible	2–3	Moderate	Moderate	Yes
	3	Strong	Moderate	Yes
Susceptible	3–4	Strong	Strong	Yes
	4	Strong	Strong	No

Table 3. Nine-step 0–4 scale for scoring infection types in cereal powdery mildew, after Torp et al.[88].

743.9 to 754.4 Mbp on chromosome arm 5RL (Supplementary Fig. S2, Supplementary Table S4). Gene mining by GWAS in the restorer population alone was not performed due to the low number of resistant lines (n = 5).

Nucleotide-binding leucine-rich repeat proteins in powdery mildew resistance-associated blocks on chromosome arm 7RL. Annotation of NLR genes in the reference genomes 'Lo7' and 'Weining' led to the identification of 770 and 1027 full-length ('complete') NLR genes, respectively. Positional information, annotation, and NB-ARC and NLR sequences of all 'complete' and 'partial' NLR genes for both reference genomes have been made available at https://doi.org/10.5281/zenodo.5085854. The PM resistance-associated

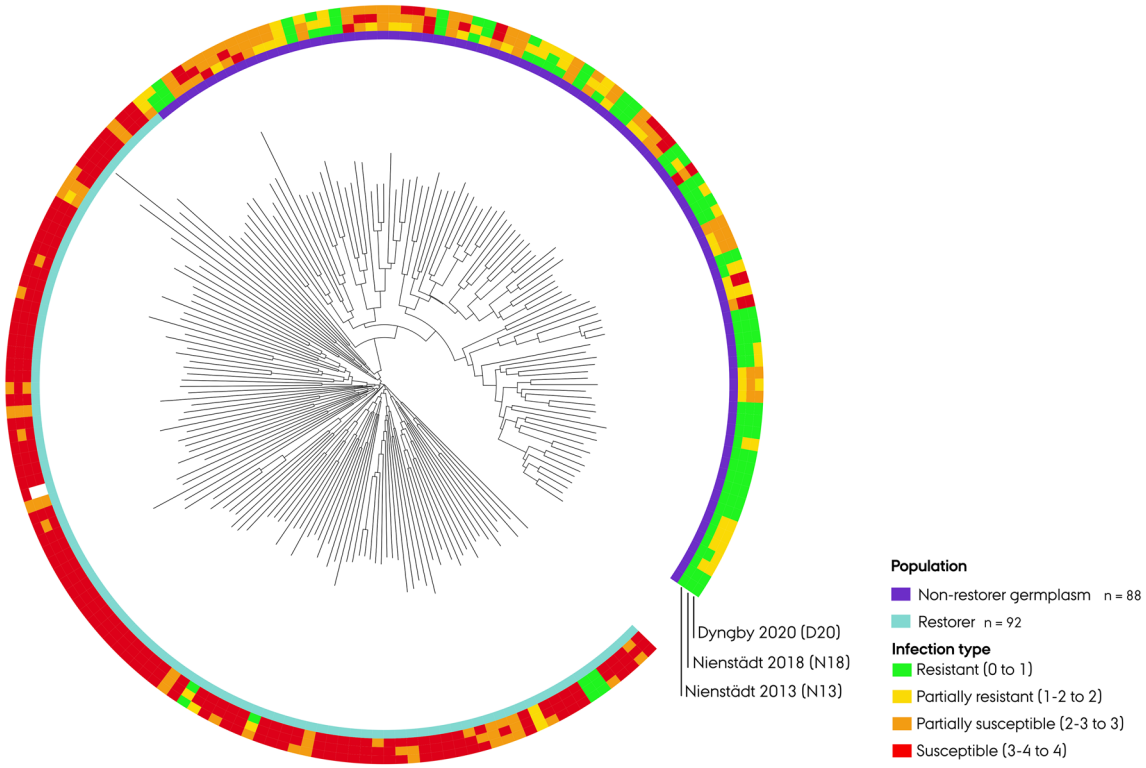


Figure 2. Circular neighbor-joining dendrogram of 180 hybrid rye (*Secale cereale* L.) breeding lines. Infection type (0–4) reaction against three powdery mildew populations displayed by concentric circles around the dendrogram.

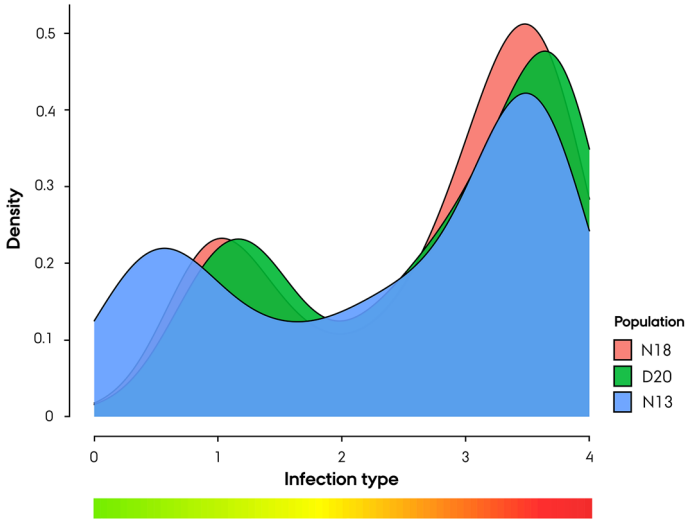


Figure 3. Density plot of the infection type distribution across 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines against three powdery mildew populations.

block residing in the subtelomeric region of chromosome arm 7RL spanned 17 Mbp on the 'Lo7' reference genome and harbored 25 NLR genes (Fig. 4C, Supplementary Table S8). Out of the 244 markers residing in the block significantly associated ($-\log_{10}(p) > 6.72$) with PM resistance, 155 were accurately positioned on the 'Weining' reference genome (Supplementary Table S7). The markers mapped to a site spanning 14 Mbp from 994.6 to 1,008.4 Mbp on chromosome arm 7RL and housed 16 NLR genes (Supplementary Table S8). A search in the protein database using the NCBI Blastx function showed that the majority of NLRs shared similarities with resistance gene analogs (RGAs) and *Pik-2-like* and *Pik6-NP*-like disease resistance proteins in the diploid parental species of wheat. In both reference genomes, two NLRs shared similarities with *Rpp13-like* disease resistance proteins in *Triticum urartu* and barley.

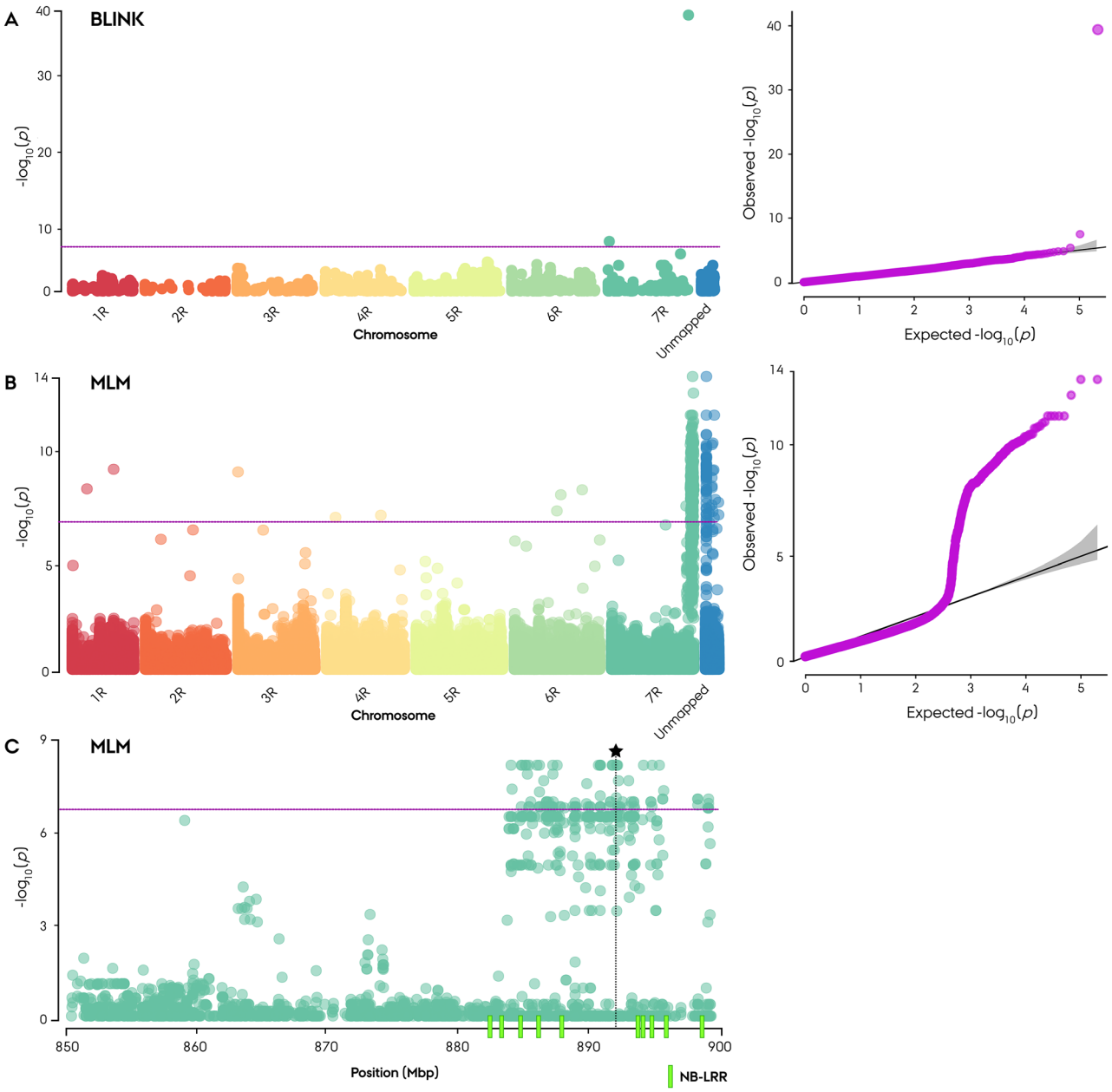


Figure 4. Manhattan plot for the genome-wide association study (GWAS) of powdery mildew disease resistance in Nordic Seed hybrid rye (*Secale cereale* L.) elite breeding germplasm (n = 180) using 261,406 informative SNP markers. (**A**) GWAS using BLINK including a Q–Q plot. (**B**) GWAS using the MLM method including a Q–Q plot. (**C**) GWAS using the MLM method for the PM resistance-associated region on chromosome arm 7RL. Asterisks denote the position of the top-associated SNP marker identified in BLINK GWAS. NLR genes on the 'Lo7' reference genome are represented by green bars. The purple line represents the Bonferroni adjusted significance threshold based on informative markers.

Phylogenetic analysis using the NB-ARC domain of NLRs residing in the PM resistance-associated haplotype block on chromosome arm 7RL led to the finding that the majority of the NLRs were represented in both reference genomes (Fig. 5). The 'Weining' reference genome exhibited one unique NLR not present in 'Lo7', while the latter exhibited eight unique NLRs, of which five formed a distinct clade. Eight homogeneously represented NLRs from each reference genome clustered in a large clade ('clade 1').

An additional phylogenetic analysis was performed using the entire reference genome NLR repertoire including NLR genes of characterized *R* genes as references (Fig. 6, Supplementary Fig. S4). The NLR genes residing within the haplotype block were found to span much of the reference NLR repertoire diversity. Clade 1 remained intact in both reference genome NLR repertoire trees positioned in a section harboring four out of five reference *Pm* genes included, with the closest being *Pm60*. In both reference genome NLR repertoire trees, a single

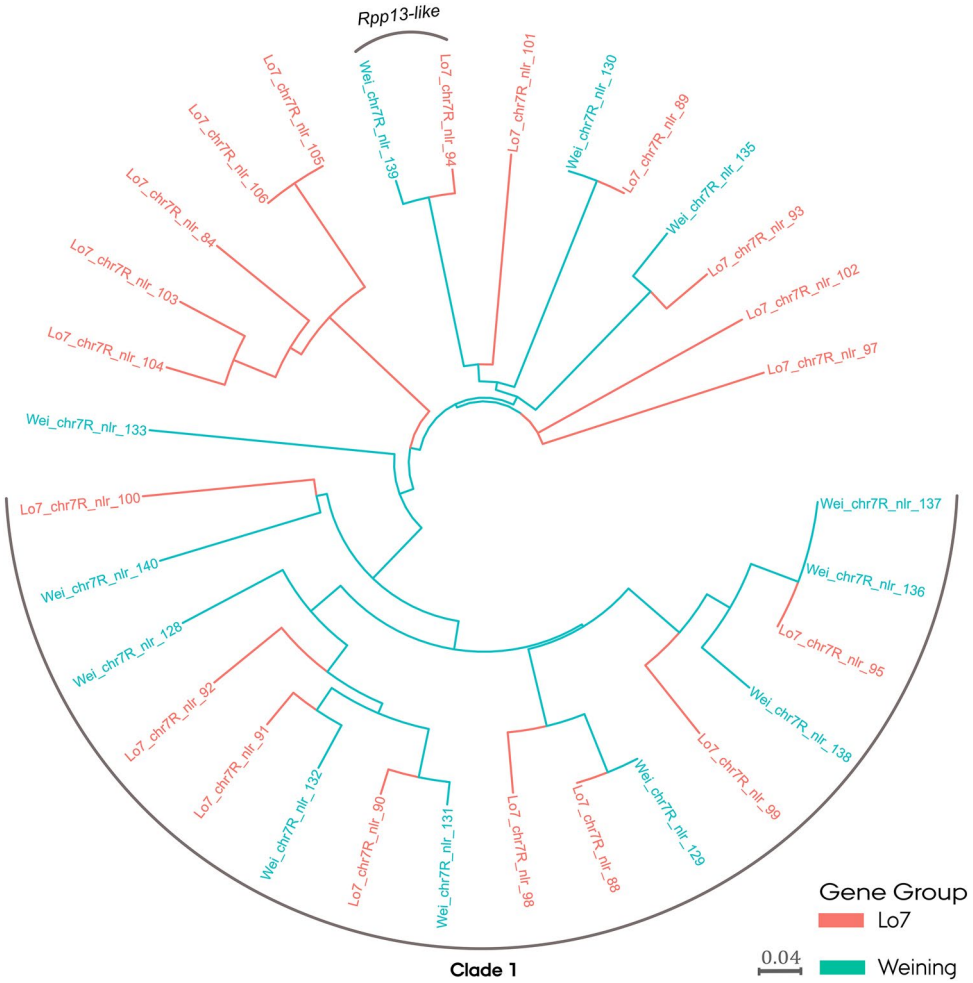


Figure 5. Phylogenetic tree of nucleotide-binding leucine-rich repeat (NLR) protein NB-ARC domains in the 'Weining' and 'Lo7' rye (*Secale cereale* L.) reference genomes. NLR genes are located at a powdery mildew resistance-associated site in the subtelomeric region of chromosome arm 7RL.

NLR (Lo7_chr7R_nlr_94, Wei_chr7R_nlr_139) showed an evolutionary relationship with an *Rpp13*-like disease resistance gene and four *Mla* alleles.

Characterization of an *Rpp13*-like NLR gene residing in the resistance-associated block on chromosome arm 7RL. While several of the NLR genes residing within the PM resistance-associated haplotype block on chromosome arm 7RL chromosome showed evidence of an evolutionary relationship with known *Pm* genes, we selected the *Rpp13*-like NLR gene for further investigation on the basis of its close proximity. The *Rpp13*-like NLR gene (Lo7_chr7R_nlr_94, Wei_chr7R_nlr_139) was the closest NLR gene, residing approximately 2 Mbp from the top-most PM resistance-associated marker identified by BLINK (Fig. 4C, Supplementary Table S8).

The homologous *Rpp13*-like NLR genes in the 'Lo7' and 'Weining' reference genomes encoded canonical NLR proteins of 922 to 947 aa, showing 97% sequence similarity, with sequences differing by three single amino acid variants and an indel of 25 aa in the NB-ARC domain (Supplementary material 2). Protein BLAST of the NLR gene led to 88% sequence similarity with disease resistance *Rpp13-like* protein 4 in *T. urartu*, which encodes a 924 aa NLR protein sequence.

Discussion

In Denmark, the top-yielding hybrid rye cultivars and population varieties were evaluated in Danish official trials as an advisory service for farmers[53]. In the last decade, a high level of powdery mildew infection in rye was recorded only once in the trials in 2017 at a site in southern Denmark (Supplementary Fig. S1). At the trial site, several hybrid rye cultivars displayed up to 18% leaf area covered by PM. While no recent studies have investigated the effect of powdery mildew in top-yielding hybrid rye cultivars, Matzen et al. (2019) reported a 16% yield reduction in triticale at a disease severity of ≤ 10% leaf area covered by powdery mildew under field conditions in 2017. It is, therefore, reasonable to presume that under certain conducive conditions, PM is capable

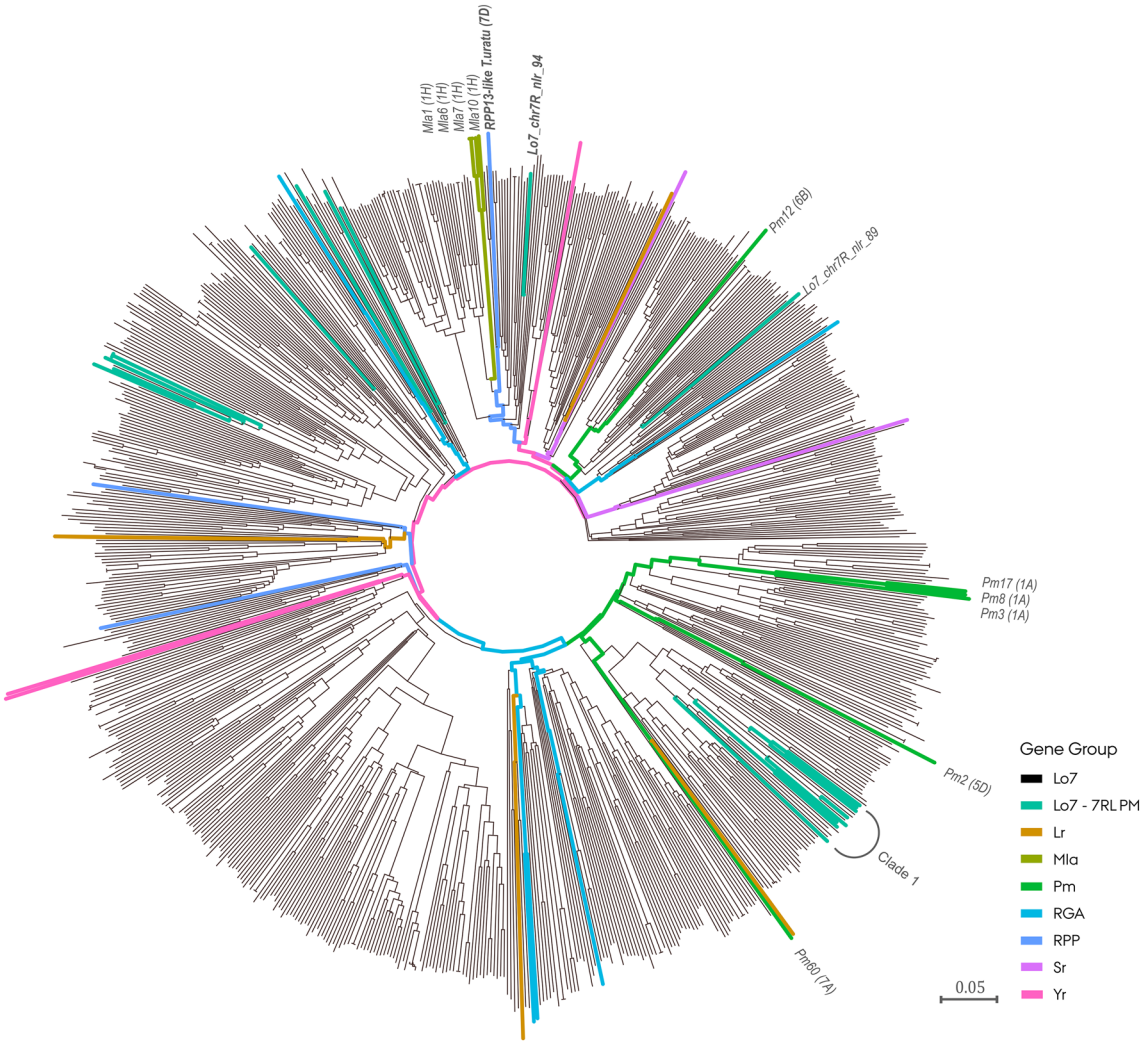


Figure 6. Phylogenetic tree of 770 nucleotide-binding leucine-rich repeat (NLR) protein NB-ARC domains in the 'Lo7' rye (*Secale cereale* L.) reference genome. The NB-ARC domains of known NLR genes have been included as references. NLR genes residing in a powdery mildew resistance block in the subtelomeric region of chromosome arm 7RL (Lo7–7RL PM) are colored teal.

of causing substantial grain yield and quality losses in rye. Furthermore, on the basis of the trial records, we selected a hybrid cv. KWS Serafino as a resistant control in our study, as it showed less than 1% leaf area covered by PM at the site in 2017. However, under a high disease pressure, cv. KWS Serafino was found to be susceptible in the current study, suggesting a potential lower level of resistance against PM in the evaluated top-yielding hybrid cultivars than evident from the official trials.

In this study, we investigated PM resistance in the less-prevalent Gülzow-type elite hybrid rye breeding germplasm against three distinct *Blumeria graminis* f. sp. *secalis* (*Bgs*) populations from Denmark and Northern Germany. We observed a moderate level of powdery mildew resistance in the non-restorer germplasm population, and by performing a genome-wide association study (GWAS) using 261,406 informative SNP markers, we identified a strong PM resistance-associated site on the distal region of chromosome arm 7RL.

Hybrid rye breeding germplasms are highly secluded with little exchange of material. The existing exchange is subsequently influenced by the deployed fertility control system, which determines the compatibility of foreign material for introgression into heterotic parental populations[59]. Additionally, the 600 K SNP array was developed using lines from a German hybrid rye breeding germplasm deploying the predominant Pampa-type cytoplasmic male sterility system (CMS)[57]. In contrast, the germplasm assayed in this study deploys the less-prevalent Gülzow-type CMS system[60]. Due to the distinct nature of the different hybrid rye breeding germplasms, we investigated the performance of the 600 K SNP array on the Gülzow-type germplasm. With no physical position data available for the 600 K SNP markers, we developed a marker map by anchoring the marker sequences to the 'Lo7' reference genome. In brief, we found that the 600 K SNP array yielded a dense panel of informative SNP markers in the Gülzow-type germplasm with markers homogeneously distributed across the rye genome. The level of marker informativeness measured by the polymorphism information content (PIC) was largely comparable with observations made in maize (*Zea mays* L.). Here, high-density SNP genotyping of 544 diverse CIMMYT

inbred lines yielded 362 K informative SNP markers with a mean PIC of 0.25[61]. In this study, however, we did observe a drop in marker informativeness and density across the pericentromeric region. Similar observations were made in the study by Rabanus-Wallace, et al.[50], who reported a considerable reduction in both genetic diversity and gene density within the pericentromeric region of the inbred rye line 'Lo7'. In barley, this region has furthermore been observed to display a 20-fold lower recombination rate[62]. Thus, we concluded that the 600 K SNP array performed satisfactorily on the Gülzow-type germplasm in relation to the fundamental characteristics investigated. The high-density array constitutes a milestone in rye genomic resources transitioning SNP-based genomic studies and replaces the previous rye 5 K SNP array by Haseneyer, et al.[57,63].

GWAS on SNP genotype data has become an effective tool in genome-based plant breeding for the study of oligogenic traits often governed by a few genes with large effects, such as the major monogenic inherited resistance (*R*) genes[63]. The use of high-density SNP typing has allowed whole-genome scans for the identification of often small haplotype blocks significantly associated with resistance[64]. The creation of the 600 K SNP array and the chromosomal-scale reference genomes 'Lo7' and 'Weining' in rye has significantly changed the genomic toolbox available for mining novel resistance genes[50,57]. Using these recent advances in rye genomic resources, we successfully managed to identify a site on chromosome arm 7RL that was significantly associated with PM resistance. The high level of resolution provided by the dense marker panel revealed that the PM resistance-associated haplotype block spanned 17 Mbp on the distal tip of the chromosome arm 7RL subtelomeric region.

Until recently, no PM resistance gene had been identified on the rye 7R chromosome. However, during the development of translocation lines for wheat improvement using a local Chinese variety of rye 'Baili', Ren et al.[42] discovered a PM-resistant 7BS:7RL translocation line. With the recipient wheat parent being susceptible, the PM resistance gene traced back to the rye donor chromosome arm 7RL. Resistance phenotyping of the translocation line demonstrated that it displayed a high level of PM resistance against prevalent *Bg. tritici* pathotypes in China, making it very promising for the development of novel PM-resistant wheat cultivars. Although a novel finding in rye, several PM resistance genes have been identified in wheat chromosomal segments syntenic to rye chromosome arm 7RL. During *Triticeae* speciation, a series of recurrent translocation events gave rise to major patterns of chromosomal rearrangements[65]. In rye, the distal region of chromosome arm 7RL, therefore, shows high homology with the wheat 2A/B/D chromosomes[58,66]. Currently, more than five *Pm* genes have been identified in wheat on the 2A/B/D chromosomes[67,68].

With the *Pm* gene discovered by Ren, et al.[42] originating from a forage-type population of rye in a gene pool distinct from the germplasm investigated in this study, it is reasonable to presume that the two *Pm* genes could be either distinct or allelic[69]. We provisionally denote the novel *Pm* locus residing in the subtelomeric region of chromosome arm 7RL as *PmNOS1*.

Comparative analysis of the three *Bgs* populations used in the study revealed that the N13 population significantly differed from the two more recent populations, showing a less virulent composition of pathotypes. In addition, nine non-restorer germplasm breeding lines displaying a differential resistance profile were identified, showing 'partial resistance' only to the N13 population. With none of these lines carrying the resistance haplotype associated with the resistance locus on chromosome arm 7RL, our findings suggest that these do not carry the *PmNOS1* locus. While this observation can be explained by the occurrence of recombination events between the PM resistance-associated marker and the causative gene, this result seems less likely due to the consistent divergence in the haplotype[70]. Instead, it seems more likely that the differentially resistant non-restorer germplasm lines carry a distinct *Pm* gene that lost its effect during the period 2013–2018 in northern Germany. As a result of host genetic uniformity and the high evolutionary capacity of *Bg* to acquire virulence and migrate rapidly over long distances, the effectiveness of *Pm* genes is often rapidly lost. An example of this is the *Mla13 Pm* gene introduced in a former Czechoslovakian barley cultivar, 'Koral', in 1980[17]. After years of providing effective resistance against PM disease, virulence was observed in England in 1988 in a pathotype believed to originate from Czechoslovakia, having migrated by wind across the European continent and North Sea[71]. Ongoing monitoring of the virulence structure of *Bg* was conducted to survey the effectiveness of deployed *Pm* genes in elite cultivars[72,73]. While the *Pm* gene present in the differentially resistant non-restorer germplasm lines has seemingly been overcome, our findings suggest that *PmNOS1* remains effective.

Enabled by recent advances in rye genomic resources, we investigated whether any of the NLR genes residing in the region harboring the *PmNOS1* locus resembled known *Pm* genes. In several crop species, including rye, NLR genes have been observed to accumulate at recombination hotspots in subtelomeric chromosomal regions[50,74,75]. Additionally, we identified several large clusters of NLR genes residing in the PM resistance-associated block harboring *PmNOS1* in both of the reference genomes. In addition to the PM resistance gene identified in our study, Fusarium head blight and leaf rust resistance genes have been mapped to the subtelomeric region of chromosome arm 7RL[76,77]. Stem and stripe rust resistance genes reported in the 7BS:7RL translocation line developed by Ren et al.[42] are furthermore likely to reside in the subtelomeric region. To investigate the likely diversity of *R* genes residing in the PM resistance-associated block, we conducted a phylogenetic analysis using the NB-ARC domain sequence of NLR genes residing in the block[78,79]. In contrast to the rapidly evolving LRR domain often exhibiting intraspecific polymorphism, the NB-ARC domain is largely conserved and suited for the study of evolutionary relationships among NLR genes[80,81]. As expected, the NLR genes residing in the block represented a large proportion of the NLR repertoire diversity in rye, accentuating the evolutionary plasticity of the NLRs residing in the subtelomeric region[75]. Intriguingly, guided by a panel of isolated NLR genes as a reference, we observed a predisposition of NLRs within the block to be in close proximity to known *Pm* genes in wheat and barley. This evolutionary relationship could hint at a common attribute among the NLRs[79].

Phylogenetic analysis led to the identification of an NLR gene with an evolutionary relationship and protein sequence similarity to *Rpp13*-like protein 4 in *T. urartu*. Based on its close proximity to the marker displaying the strongest association with the *PmNOS1* locus, the *Rpp13*-like NLR gene was selected for further characterization. In *A. thaliana*, *Rpp13* confers resistance against *Peronospora parasitica*, the causative agent of downy mildew

disease[82]. Recent studies have, however, identified *Rpp13-like* NLR genes associated with powdery mildew resistance in cereals. In wheat, Liu et al.[83] showed that silencing of the *Rpp13* homologous gene *TaRPP13-3* in resistant wheat cv. 'Brock' induced susceptibility to powdery mildew. In barley, Cheng et al.[84] found that the expression of an *Rpp13*-like NLR gene was highly upregulated after inoculation with powdery mildew.

In rye, while the rate of decay has only been determined in a few genes related to frost response, which showed a rapid linkage decay[85], these genes have been demonstrated to decay after 3.76 kb, on average, across the genome in a similar hybrid breeding germplasm in maize[61]. Due to the heterogenic nature of outcrossing species, their rate of decay is often rapid[86]. However, in a recent population study on assayed germplasm, non-restorer germplasm was found to exhibit relatively low genetic diversity, low effective population size and high linkage disequilibrium[59]. Consistent with their observations, we found a large conserved haplotype on chromosome arm 7RL harboring the *PmNOS1* locus[59]. The linkage decay in the non-restorer germplasm is likely considerably reduced by the low genetic diversity and effective population size, resulting in a similarly low frequency of effective recombination events. The large amount of linkage, while beneficial for trait discovery in GWAS at lower marker density, impedes the identification of a narrow and precise genomic region that may harbor the gene of interest, even at high marker resolution. In the case of the *PmNOS1* locus, the haplotype block on chromosome arm 7RL was found to span 17 Mbp, harboring between 17 and 25 potential candidate NLR genes in the 'Lo7' and 'Weining' reference genomes. For more accurate mapping of the *PmNOS1* locus, the development of multiparent mapping populations could be conducted, allowing several generations of potential effective recombination events in the region[87,88]. Identification of the causative gene could be performed by resistance gene enrichment sequencing (RenSeq) analysis followed by transformation of a susceptible non-restorer germplasm line to validate the gene[89]. This would in turn equally show whether the gene is present in the reference genomes and whether *PmNOS1* encodes an *Rpp13*-like NLR protein.

In conclusion, our study demonstrates the immediate value of recent advances in rye genomic resources for the mining of novel resistance genes. These resources now permit accurate identification of delimited resistance-associated haplotype blocks and scanning for trait-associated genes residing within. With pathogens such as *Bg* displaying a large evolutionary plasticity, shortening the process from identification of resistance-associated sites to isolation of the underlying *R* gene is important for the development of novel resistant cultivars. The relevance of studies in rye is accentuated by the possibility of introgressing novel *R* genes into the staple cereal wheat by chromosomal translocation lines. To aid further studies in the field, we have provided both a rye600K SNP array marker map anchored using the 'Lo7' reference genome and NLR repertoire information of the 'Lo7' and 'Weining' reference genomes in open-access data repositories.

Materials and methods

Plant material and DNA extraction. A panel of 180 inbred rye (*Secale cereale* L.) lines, 92 restorer and 88 non-restorer germplasm, belonging to the elite Gülzow-type hybrid rye breeding germplasm at Nordic Seed A/S (Dyngby, Denmark) were investigated. Population structure and information on the genetic characteristics of the germplasm were presented in a recent study by Vendelbo et al.[59]. The parental populations represent genetically secluded gene pools with restorer (paternal) lines carrying a dominant allele for the restoration of male fertility and non-restorer germplasm lines carrying a recessive allele and fertile cytoplasm, which is used to maintain cytoplasmic male sterile (maternal) lines. DNA extraction was performed using an adapted SDS-based method according to the USDA[90] after Pallotta et al.[91] on an equivalent of 75 mg of plant material collected from the coleoptiles and primary leaves of two seven-day-old seedlings per line. The DNA concentration and 260/280 nm ratio of the samples were measured using an Epoch™ microplate spectrophotometer (Biotek®), and evidence of fragmentation was obtained by size visualization on a 1.2% agarose gel.

Molecular marker resource and SNP genotyping. Samples of each line containing 200 ng of high molecular weight gDNA with a ≥ 1.8 260/280 nm ratio were sent for single nucleotide polymorphism (SNP) genotyping at Eurofins Genomics Europe Genotyping (Aarhus, Denmark). Genotyping was performed using a 600 K SNP array with 600,843 SNP markers on an Affymetrix GeneTitan™ Scanner platform[57].

Collection and multiplication of *Blumeria graminis* f. sp. *secalis* populations. As there was no unified information on germplasm resistance to powdery mildew, three field populations of *Bgs* were sampled in Northern Germany and Denmark in the period from 2013 to 2020 to screen the lines against a broad range of *Bgs* pathotypes prevalent in Northern Europe. Two *Bgs* populations were collected at the Nordic Seed rye multiplication site in Germany in 2013 (N13) and 2018 (N18) (52.29254°N, E9.14896°E). Ten leaves exhibiting PM disease were carefully collected and rinsed in 0.5 L of water to release *Bgs* spores. Six pots containing 15–20 susceptible 12-day-old seedlings of the restorer line R277 were then inoculated by spraying a fine mist of the spore solution using an atomizer bottle. Pots were transferred to a separate climate chamber for each population and incubated at 18 °C with 12 h of light using 400 W high-pressure Phillips SON-T Agro lamps. Populations were continuously multiplied in an overlapping two-week cycle. Each week, the tray of 2-week-old seedlings was substituted with a tray of fresh 12-day-old seedlings. Inoculation was achieved by passive dissemination of spores from the 'older' tray to the tray with new plants through a steel-grid shelf.

An additional *Bgs* population was collected in autumn 2020 (D20) at Nordic Seed (Dyngby, Denmark) (55.94944°N, E10.25414°E) using a mixture of hybrid cvs. KWS Binntto, KWS Bono and KWS Florano. Pots with 15 to 20 seedlings of the susceptible mixture were placed outdoors 12 days after sowing (DAS) in August-October 2020. Plants were controlled regularly, and all leaves showing PM disease within the period were collected. Prior to inoculation, leaves were placed in Petri dishes with moist filter paper and set to sporulate at room temperature in light for four to eight hours. Then, a pot containing 15–20 seedlings of the susceptible mixture was inoculated

at 12 DAS by horizontally stroking the collected leaves across the seedlings. Inoculated pots were sprayed with a fine mist of water, placed in a container with transparent lids to ensure 100% RH and incubated in the dark at 10–15 °C for 24 h. After incubation in the dark, the pots were transferred to an isolated greenhouse cabin and incubated under 16 h of daylight at 18–24 °C and 8 h of dark at 14–16 °C.

Multiplication of trial inoculum for the three *Bgs* populations was performed using the same procedure as described above. For each population, two trays containing 35 pots of the susceptible mixture were inoculated by brushing 3–4 highly infected pots inoculated 20 days earlier across the tray.

Infection and scoring. All lines were phenotyped for the infection-type response to the *Bgs* populations. In each greenhouse trial, eight seeds per line were sown in a 28-hole tray using a completely randomized design with two repetitions for each of the two trial replicates. For each tray, a positive ('susceptible') control consisting of hybrid cv. KWS Binntto and negative ('resistant') control cv. KWS Serafino was included. At 14 DAS, trays were inoculated as described above by brushing 3–4 highly infected pots inoculated 20 days earlier across the tray. After 14 days of incubation, the lines were phenotyped by scoring the infection response on the first and second leaves for each of the eight seedlings per repetition in accordance with a 9-step 0–4 scale by Torp et al.[88] (Table 3).

Data analysis. Bioinformatic analysis of SNP marker data was performed with the R studio (v. 1.3.959) interface in R statistical software (v. 4.0.1) by applying various predesigned packages[92,93].

Mapping of 600 K SNP array markers to the 'Lo7' reference genome. Positional data of the 600 K SNP markers were obtained by mapping each of the 600,843 SNP marker sequences to the rye reference genome 'Lo7' using the NCBI blastn (v. 2.9.0 +) function[50,94]. The mapping positions of SNPs were hereafter stringently filtered for I) complete SNP sequence alignment and II) a maximum of 1 mismatch to ensure accurate positioning.

Molecular markers and characterization of 600 K SNP array performance. Prior to analysis, markers were filtered for a marker allele frequency ≥ 0.005, missing individual score ≤ 0.2 and missing marker score ≤ 0.1. Fundamental characteristics of SNP marker informativeness, including polymorphism information content (PIC), were calculated using the SnpReady (v. 0.9.6) R package[95]. The interchromosomal distribution of the informative marker PIC, marker-to-marker distance and marker density in 10 mb bins on the 'Lo7' rye genome were visualized using Circos (v. 0.69.8) in the Galaxy online interface[96,97]. A Circos plot was constructed using the pipeline developed by Hiltemann, et al.[98].

Analysis of phenotypic data. The distribution of infection types against the three respective PM populations was visualized by density plots using ggplot2 (v. 3.3.3) R package[99]. To determine whether the population infection-type distribution differed significantly, ANOVA was conducted using R.

To correct the resistance phenotype for the effects of replication and population, we fitted the data to a linear mixed model using the lme4 (v. 1.1.26) package in R:

$$y = \mu + P + R + l + \varepsilon$$

where μ is the general mean, P is the population, R represents the replications, *l* is the line id, and ε is the residuals. P and R were set as fixed effects, and *l* was set as a random effect. The random effect and residuals were assumed to be independent normally distributed variables described as follows: $l \sim N(0, I\sigma^2_l)$, and $\varepsilon \sim N(0, I\sigma^2\varepsilon)$. The BLUP solutions for the line effect were used in GWAS. Data were also fitted to a linear mixed model for each of the populations to correct for the effect of replication. In this model 'P' was removed.

Genome-wide association study. Discovery of PM resistance-associated SNP markers was performed by a genome-wide association study (GWAS) using the genomic association and prediction integration tool (GAPIT) (v.3) package in R[100]. The Manhattan plot was colorized using the RColorBrewer (v.1.1–2) R package color palette[99]. GWAS using a mixed linear model (MLM) was performed to identify discrete haplotype blocks associated with powdery mildew resistance. Additionally, the Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) method was performed to identify the top-most powdery mildew resistance-associated marker within each haplotype block[101]. BLINK uses a multiple loci test for MLM by combining a fixed effects model, Bayesian information content and linkage disequilibrium information to collectively improve the statistical power while simultaneously reducing the computational run time. Markers that are in linkage disequilibrium with the top-most significant marker at a site are excluded in BLINK. A standard Bonferroni-corrected threshold of $\alpha = 0.05$ was used as the significance threshold.

Annotation of nucleotide-binding leucine-rich repeat proteins in the reference genome. To investigate whether PM-associated sites in the 'Lo7' and 'Weining' reference genomes housed nucleotide-binding leucine-rich repeat (NLR) genes, annotation was performed using NLR-parser (v.3) and NLR-annotator (https://github.com/steuernb/NLR-Annotator)[79,102].

Characterization of nucleotide-binding leucine-rich repeat proteins. The gene structure, coding sequence and NLR protein sequence of candidate genes were extracted from RNA-seq and de novo protein data provided in 'Weining' and 'Lo7' reference genome data repositories[50,58]. To investigate whether NLR genes residing in

PM resistance-associated sites resembled known genes, the NCBI blastx function was used for protein–protein searches in the online database[94]. For functional analysis and the prediction of protein domains, InterPro Scan was used[103]. Identification of sequence divergence between reference genome homologs was performed by multiple sequence alignment using the multiple sequence comparison by log-expectation (MUSCLE) method for coding sequences and the 'Clustal Omega' method for NLR protein sequences in Geneious Prime (v. 2020.2.3).

Phylogenetic analysis. Neighbor-joining clustering analysis of breeding lines was performed with Euclidean genetic distance measurement using the ape (v. 5.3) R package[104]. The tree was constructed after 10,000 bootstrapping iterations with weak nodes (≤80% recurrence) collapsed into multifurcations. A circular neighbor-joining tree was generated using the iTOL (v. 5) online tool (http://itol.embl.de/), enabling a colorful visualization of each line infection type spectrum against the three *Bgs* populations[105].

To investigate the sequence similarity between nucleotide-binding leucine-rich repeat (NLR) genes at PM resistance-associated sites in the 'Lo7' and 'Weining' reference genomes, neighbor-joining clustering analysis using NB-ARC (nucleotide-binding adaptor shared by APAF-1, R proteins, and CED-4) domain sequences was performed. Identification of PM resistance-associated sites on 'Weining' was performed by mapping PM resistance-associated markers using the same procedure as described previously for 'Lo7'. As references, NB-ARC domain sequences of available leaf rust (*Lr1, Lr10, Lr21, Lr22a*), stem rust (*Sr13, Sr22*), yellow rust (*Yr5, Yr10, Yr28*), powdery mildew (*Pm2, Pm3, Pm8, Pm12, Pm17, Pm60, Mla1, Mla6, Mla7, Mla10, RPP13, RPP-like-T. urartu*) and resistance gene analogs (*RGA1, RGA2, RGA3, RGA4, RGA5*) were included. NB-ARC sequences were obtained from the UniProt online database[106]. Phylogenetic analysis was conducted using a pipeline developed by Toparslan, et al.[107] in R. Multiple sequence alignment of NB-ARC domain sequences was performed via msa (v. 1.20.1) using the 'Clustal Omeage' method and pairwise genetic distance based on identity calculated with the seqinr (4.2–8) package in R[108,109]. A tree was constructed for the 'Lo7' and 'Weining' NLR repertoires separately and visualized using ggtree (v. 2.2.4) R package[110].

Graphical editing. Graphs and figures were outputted from R in .svg format and manually curated using Inkscape (v. 1.1) (https://inkscape.org/).

Ethical statement. Experimental research and field studies on plants (either cultivated or wild), including the collection of plant material,complies with relevant institutional, national, and international guidelines and legislation.

Received: 15 July 2021; Accepted: 16 November 2021
Published online: 29 November 2021

References

1. Schlegel, R. H. J. *Rye genetics, breeding and cultivation* (CRC Press Taylor & Francis Group, 2014).
2. Schlegel, R. H. J. & Korzun, V. Genes, markers and linkage data of rye (Secale cereale L.), 11th updated inventory, V. 0.1.21. 1–115. http://www.rye-gene-map.de (2021).
3. Crespo-Herrera, L. A., Garkava-Gustavsson, L. & Åhman, I. A systematic review of rye (*Secale cereale* L.) as a source of resistance to pathogens and pests in wheat (*Triticum aestivum* L.). *Hereditas* **154**, 1–9. https://doi.org/10.1186/s41065-017-0033-5 (2017).
4. Savary, S. *et al.* The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evolut.* **3**, 430–439 (2019).
5. Savary, S., Ficke, A., Aubertot, J.-N. & Hollier, C. Crop losses due to diseases and their implications for global food production losses and food security. *Food Secur.* **4**, 519–537. https://doi.org/10.1007/s12571-012-0200-5 (2012).
6. Griffey, C. A., Das, M. K. & Stromberg, E. L. Effectiveness of adult-plant resistance in reducing grain yield loss to powdery mildew in winter wheat. *Plant Dis.* **77**, 618–622 (1993).
7. Conner, R. L., Kuzyk, A. D. & Su, H. Impact of powdery mildew on the yield of soft white spring wheat cultivars. *Can. J. Plant Sci.* **83**, 725–728 (2003).
8. Matzen, N., Heick, T. M. & Jørgensen, L. N. Control of powdery mildew (*Blumeria graminis* spp.) in cereals by Serenade® ASO (Bacillus amyloliquefaciens (former subtilis) strain QST 713). *Biol. Control* **139**, 104067 (2019).
9. Te Beest, D. E., Paveley, N. D., Shaw, M. W. & Van Den Bosch, F. Disease-weather relationships for powdery mildew and yellow rust on winter wheat. *Phytopathology* **98**, 609–617 (2008).
10. Troch, V. *et al.* Formae speciales of cereal powdery mildew: Close or distant relatives?. *Mol. Plant Pathol.* **15**, 304–314. https://doi.org/10.1111/mpp.12093 (2014).
11. Jorgensen, J. H. Erysiphe graminis, powdery mildew of cereals and grasses. *Genet. Plant Pathogenic Fungi Adv. Plant Pathol.* **6**, 137–157 (1988).
12. Jankovics, T. *et al.* New insights into the life cycle of the wheat powdery mildew: direct observation of ascosporic Infection in *Blumeria graminis* f. sp. tritici. *Phytopathology* **105**, 797–804. https://doi.org/10.1094/PHYTO-10-14-0268-R (2015).
13. Cowger, C., Parks, R. & Kosman, E. Structure and migration in U.S. *Blumeria graminis* f. sp. tritici populations. *Phytopathology* **106**, 295–304. https://doi.org/10.1094/PHYTO-03-15-0066-R (2016).
14. Dreiseitl, A. Great pathotype diversity and reduced virulence complexity in a Central European population of *Blumeria graminis* f. sp. hordei in 2015–2017. *Europ. J. Plant Pathol.* **153**, 801–811. https://doi.org/10.1007/s10658-018-1593-6 (2018).
15. Zhu, J. *et al.* Genetic evidence of local adaption and long distance migration in *Blumeria graminis* f. sp. hordei populations from China. *J. General Plant Pathol.* **82**, 69–81. https://doi.org/10.1007/s10327-016-0643-1 (2016).
16. Limpert, E., Godet, F. & Müller, K. Dispersal of cereal mildews across Europe. *Agric. Forest Meterol.* **97**, 293–308 (1999).
17. Wolfe, M. S., Brändle, U., Koller, B., Limpert, E. & McDermott, J. Barley mildew in Europe: population biology and host resistance. *Euphytica* **63**, 125–139 (1992).
18. Brown, J. K. & Hovmoller, M. S. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297** (2002).
19. Kast, W. K. & Geiger, H. H. Studies on the inheritance of mildew resistance in rye. I. Results from inbred lines and F1 crosses. (1982).

20. Miedaner, T., Schmidt, H. & Geiger, H. Components of variation for quantitative adult-plant resistance to powdery mildew in winter rye. *Phytopathology* **83**, 1071–1075 (1993).
21. Nelson, R., Wiesner-Hanks, T., Wisser, R. & Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **19**, 21–33. https://doi.org/10.1038/nrg.2017.82 (2018).
22. Frieb, B., Heun, M. & Bushuk, W. Cytological characterization, powdery mildew resistance and storage protein composition of tetraploid and hexaploid 1BL/1RS wheat-rye translocation lines. *Theor. Appl. Genet.* **78** (1989).
23. Hsam, S. L. K. & Zeller, F. J. Evidence of allelism between genes pm8 and pm17 and chromosomal location of powdery mildew and leaf rust resistance genes in the common wheat cultivar "Amigo". *Plant Breed.* **116**, 119–122 (1997).
24. Ren, T.-H. *et al.* Development and characterization of a new 1BL.1RS translocation line with resistance to stripe rust and powdery mildew of wheat. *Euphytica* **169**, 207–213. https://doi.org/10.1007/s10681-009-9924-5 (2009).
25. Han, G. *et al.* Identification of an elite wheat-Rye T1RS.1BL translocation line conferring high resistance to powdery mildew and stripe rust. *Plant Dis.* **104**, 2940–2948. https://doi.org/10.1094/PDIS-02-20-0323-RE (2020).
26. He, H. *et al.* Characterization of a new gene for resistance to wheat powdery mildew on chromosome 1RL of wild rye Secale sylvestre. *Theor. Appl. Genet.* **134**, 887–896. https://doi.org/10.1007/s00122-020-03739-1 (2021).
27. Koebner, R. M. D. Introduction into wheat of disease resistance from related species. *Annual report of the Plant Breeding Institute*, 69–70 (1986).
28. Driscoll, C. J. & Jensen, N. F. A genetic method for detecting introduced intergenetic translocations. *Genetics* **48**, 459–468 (1963).
29. Zhuang, L. F., Sun, L., Li, A. X., Chen, T. T. & Qi, Z. J. Identification and development of diagnostic markers for a powdery mildew resistance gene on chromosome 2R of Chinese rye cultivar Jingzhouheimai. *Mol. Breed.* **27**, 455–465. https://doi.org/10.1007/s11032-010-9443-z (2010).
30. Hysing, S.-C. *et al.* Agronomic performance and multiple disease resistance in T2BS.2RL wheat-rye translocation lines. *Crop Sci.* **47**, 254–260. https://doi.org/10.2135/cropsci2006.04.0269 (2007).
31. Heun, M. & Friebe, B. Introgression of powdery mildew resistance from rye into wheat. *Phytopathology* **80**, 242–245 (1990).
32. Driscoll, C. J. & Jensen, N. F. Release of a wheat-rye translocation stock involving leaf rust and powdery mildew resistances 1. *Crop Sci.* **5**, 279–280 (1965).
33. An, D. G., Li, L. H., Li, J. M., Li, H. J. & Zhu, Y. G. Introgression of resistance to powdery mildew conferred by chromosome 2R by crossing wheat nullisomic 2D with rye. *J. Integr. Plant Biol.* **48**, 838–847 (2006).
34. Lind, V. Analysis of the resistance of wheat-rye addition lines to powdery mildew of wheat *Tagungsbericht Akademie der Landwirtschaftswissenschaften der DDR* (1982).
35. Ma, P. *et al.* Development of novel wheat-rye chromosome 4R translocations and assignment of their powdery mildew resistance. *Plant Dis.* **104**, 260–268. https://doi.org/10.1094/PDIS-01-19-0160-RE (2020).
36. Fu, S. *et al.* New wheat-rye 5DS-4RS.4RL and 4RS-5DS.5DL translocation lines with powdery mildew resistance. *J. Plant Res.* **127**, 743–753. https://doi.org/10.1007/s10265-014-0659-6 (2014).
37. An, D. *et al.* Molecular cytogenetic characterization of a new wheat-rye 4R chromosome translocation line resistant to powdery mildew. *Chromosome Res.* **21**, 419–432. https://doi.org/10.1007/s10577-013-9366-8 (2013).
38. Riley, R. & Macer, R. C. F. The chromosomal distribution of genetic resistance of rye to wheat pathogens. *Canad. J. Genet. Cytol.* **8** (1966).
39. Hao, M. *et al.* Introgression of powdery mildew resistance gene Pm56 on rye chromosome arm 6RS into wheat. *Front. Plant Sci.* **9**, 1–8. https://doi.org/10.3389/fpls.2018.01040 (2018).
40. Friebe, B., Heun, M., Tuleen, N., Zeller, F. J. & Gill, B. S. Cytogenetically monitored transfer of powdery mildew resistance from rye into wheat. *Crop Sci.* **34**, 621–625 (1994).
41. Miller, T. E. The homoeologous relationship between the chromosomes of rye and wheat. *Current status. Canad. J. Genet. Cytol.* **26**, 578–589 (1984).
42. Ren, T. *et al.* Molecular and cytogenetic characterization of a wheat-rye 7BS7RL translocation line with resistance to stripe rust, powdery mildew, and fusarium head blight. *Phytopathology* **110**, 1713–1720. https://doi.org/10.1094/PHYTO-02-20-0061-R (2020).
43. Elmore, J. M., Lin, Z. J. & Coaker, G. Plant NB-LRR signaling: Upstreams and downstreams. *Curr. Opin. Plant Biol.* **14**, 365–371. https://doi.org/10.1016/j.pbi.2011.03.011 (2011).
44. Hurni, S. *et al.* Rye Pm8 and wheat Pm3 are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J.* **76**, 957–969. https://doi.org/10.1111/tpj.12345 (2013).
45. Xie, J. *et al.* A rare single nucleotide variant in Pm5e confers powdery mildew resistance in common wheat. *New Phytol.* **228**, 1011–1026 (2020).
46. Marone, D., Russo, M. A., Laido, G., De Leonardis, A. M. & Mastrangelo, A. M. Plant nucleotide binding site-leucine-rich repeat (NBS-LRR) genes: Active guardians in host defense responses. *Int. J. Mol. Sci.* **14**, 7302–7326. https://doi.org/10.3390/ijms14047302 (2013).
47. Wang, J., Han, M. & Liu, Y. Diversity, structure and function of the coiled-coil domains of plant NLR immune receptors. *J. Integr. Plant Biol.* **63**, 283–296. https://doi.org/10.1111/jipb.13032 (2021).
48. Takken, F. L. & Goverse, A. How to build a pathogen detector: Structural basis of NB-LRR function. *Curr. Opin. Plant Biol.* **15**, 375–384. https://doi.org/10.1016/j.pbi.2012.05.001 (2012).
49. Krasileva, K. V., Dahlbeck, D. & Staskawicz, B. J. Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell* **22**, 2444–2458. https://doi.org/10.1105/tpc.110.075358 (2010).
50. Rabanus-Wallace, M. T. *et al.* Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. *Nat. Genet.* **53**, 564–573. https://doi.org/10.1101/2019.12.11.869693 (2021).
51. Ryan, P., Carbone, L., Marshall, D. & Cowger, C. Virulence structure of eastern US wheat powdery mildew population. *Plant Dis.* **92**, 1074–1082 (2008).
52. Cowger, C., Mehra, L., Arellano, C., Meyers, E. & Murphy, J. P. Virulence differences in Blumeria graminis f. sp. tritici from the Central and Eastern United States. *Phytopathology* **108**, 402–411. https://doi.org/10.1094/PHYTO-06-17-0211-R (2018).
53. Sortsinfo. *The Danish official trial records*, <https://sortinfo.dk/> (2021).
54. Miedaner, T. & Korzun, V. Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* **102**, 560–566 (2012).
55. Mundt, C. C. Pyramiding for resistance durability: Theory and practice. *Phytopathology* **108**, 792–802. https://doi.org/10.1094/PHYTO-12-17-0426-RVW (2018).
56. Koller, T., Brunner, S., Herren, G., Hurni, S. & Keller, B. Pyramiding of transgenic Pm3 alleles in wheat results in improved powdery mildew resistance in the field. *Theor. Appl. Genet.* **131**, 861–871. https://doi.org/10.1007/s00122-017-3043-9 (2018).
57. Bauer, E. *et al.* Towards a whole-genome sequence for rye (*Secale cereale* L.). *Plant J.* **89**, 853–869. https://doi.org/10.1111/tpj.13436 (2017).
58. Li, G. *et al.* A high-quality genome assembly highlights rye genomic characteristics and agronomically important genes. *Nat. Genet.* **53**, 574–584. https://doi.org/10.1038/s41588-021-00808-z (2021).
59. 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* **15**, e0239541. https://doi.org/10.1371/journal.pone.0239541 (2020).

60. Vendelbo, N. M. *et al.* Genomic scan of male fertility restoration genes in a 'Gülzow' type hybrid 1 breeding system of rye (*Secale cereale* L.). *Int. J. Mol. Sci.*, Submitted.
61. Wu, Y. *et al.* Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theor. Appl. Genet.* **129**, 753–765. https://doi.org/10.1007/s00122-016-2664-8 (2016).
62. Baker, K. *et al.* The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. *Plant J.* **79**, 981–992. https://doi.org/10.1111/tpj.12600 (2014).
63. Liu, H. J. & Yan, J. Crop genome-wide association study: A harvest of biological relevance. *Plant J.* **97**, 8–18. https://doi.org/10.1111/tpj.14139 (2019).
64. Brachi, B., Moris, G. P. & Borevitz, J. O. Genome-wide association studies in plants: The missing heritability is in the field. *Genome Biol.* **12**, 1–8 (2011).
65. Genes, G., Li, W., Challa, G. S., Zhu, H. & Wei, W. Recurrence of chromosome rearrangements and reuse of dna breakpoints in the evolution of the triticeae genomes. *G3 Genetics* **6**, 3837–3847. https://doi.org/10.1534/g3.116.035089 (2016).
66. Martis, M. M. *et al.* Reticulate evolution of the rye genome. *Plant Cell* **25**, 3685–3698. https://doi.org/10.1105/tpc.113.114553 (2013).
67. Huang, X. Q. & Röder, M. S. Molecular mapping of powdery mildew resistance genes in wheat: A review. *Euphytica* **137**, 203–223 (2004).
68. Zhang, R. *et al.* Pm62, an adult-plant powdery mildew resistance gene introgressed from Dasypyrum villosum chromosome arm 2VL into wheat. *Theor. Appl. Genet.* **131**, 2613–2620. https://doi.org/10.1007/s00122-018-3176-5 (2018).
69. Ma, R., Yli-Mattila, A. & Pulli, S. Phylogenetic relationships among genotypes of worldwide collection of spring and winter ryes (*Secale cereale* L.) determined by RAPD-PCR markers. *Hereditas* **140**, 210–221 (2004).
70. Andersen, J. R. & Lubberstedt, T. Functional markers in plants. *Trends Plant Sci.* **8**, 554–560. https://doi.org/10.1016/j.tplants.2003.09.010 (2003).
71. Brown, J. K., Jessop, A. C. & Rezanoor, N. H. Genetic uniformity in barley and its powdery mildew pathogen. *Proc. Royal Soc. Lond. Ser. B Biol. Sci.* **246**, 83–90 (1991).
72. Xue, A. G. *et al.* Virulence structure of Blumeria graminis f sp tritici, the causal agent of wheat powdery mildew, in Ontario, Canada, in 2018 and 2019. *Canad. J. Plant Pathol.* https://doi.org/10.1080/07060661.2021.1915876 (2021).
73. Liu, N. *et al.* Virulence structure of blumeria graminis f sp tritici and Its genetic diversity by ISSR and SRAP profiling analyses. *PLoS ONE* **10**, e0130881. https://doi.org/10.1371/journal.pone.0130881 (2015).
74. Nieri, D., Di Donato, A. & Ercolano, M. R. Analysis of tomato meiotic recombination profile reveals preferential chromosome positions for NB-LRR genes. *Euphytica* https://doi.org/10.1007/s10681-017-1982-5 (2017).
75. Chen, N. W. G. *et al.* Common bean subtelomeres are hot spots of recombination and favor resistance gene evolution. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2018.01185 (2018).
76. Gruner, P., Schmitt, A. K., Flath, K., Piepho, H. P. & Miedaner, T. Mapping and validating stem rust resistance genes directly in self-incompatible genetic resources of winter rye. *Theor. Appl. Genet.* https://doi.org/10.1007/s00122-021-03800-7 (2021).
77. Wehling, P. *et al.* Leaf-rust resistance in rye (*Secale cereale* L.) 1 Genetic analysis and mapping of resistance genes Pr1 and Pr2. *Theor. Appl. Genet.* **107**, 432–438. https://doi.org/10.1007/s00122-003-1263-7 (2003).
78. Van de Weyer, A. L. *et al.* A species-wide inventory of NLR genes and alleles in Arabidopsis thaliana. *Cell* **178**, 1260–1272. https://doi.org/10.1016/j.cell.2019.07.038 (2019).
79. Steuernagel, B. *et al.* The NLR-annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiol.* **183**, 468–482. https://doi.org/10.1104/pp.19.01273 (2020).
80. Mondragon-Palomino, M., Meyers, B. C., Michelmore, R. W. & Gaut, B. S. Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. *Genome Res.* **12**, 1305–1315. https://doi.org/10.1101/gr.159402 (2002).
81. Florence, J., Vernaldi, S. & Maekawa, T. Evolution and conservation of plant NLR functions. *Front. Immunol.* **4**, 297. https://doi.org/10.3389/fimmu.2013.00297 (2013).
82. Bittner-Eddy, P. D., Crute, I. R., Holub, E. B. & Beynon, J. L. RPP13 is a simple locus in Arabidopsis thaliana for alleles that specify downy mildew resistance to different avirulence determinants in Peronospora parasitica. *Plant J.* **21**, 177–188 (2000).
83. Liu, X. *et al.* TaRPP13-3, a CC-NBS-LRR-like gene located on chr 7D, promotes disease resistance to wheat powdery mildew in Brock. *J. Phytopathol.* **168**, 688–699. https://doi.org/10.1111/jph.12949 (2020).
84. Cheng, J. *et al.* Genome-wide Identification and expression analyses of RPP13-like genes in barley. *BioChip J.* **12**, 102–113. https://doi.org/10.1007/s13206-017-2203-y (2018).
85. Li, Y. *et al.* High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (*secale cereale* L.) genes involved in frost response. *BMC Plant Biol.* **11**, 6 (2011).
86. Gupta, P. K., Rustgi, S. & Kulwal, P. L. Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* **57**, 461–485. https://doi.org/10.1007/s11103-005-0257-z (2005).
87. Ren, Y. *et al.* QTL analysis and nested association mapping for adult plant resistance to powdery mildew in two bread wheat populations. *Front. Plant Sci.* **8** (2017).
88. Novakazi, F. *et al.* You had me at "MAGIC"!: Four barley MAGIC populations reveal novel resistance QTL for powdery mildew. *Genes* **11**, 1512 (2020).
89. Arora, S. *et al.* Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* **37**, 139–143. https://doi.org/10.1038/s41587-018-0007-9 (2019).
90. USDA. *United States Department of Agriculture: wheat and barley DNA extraction protocol (96-well plate format)*, <https://www.ars.usda.gov/ARSUserFiles/60701500/SmallGrainsGenotypingLaboratory/Protocols/wheat%20and%20barleyDNA%20extraction_original.pdf> (2006).
91. Pallotta, M. A. *et al.* Marker assisted wheat breeding in the southern region of Australia. *Proceedings of the Tenth International Wheat Genetics Symposium*, 789–791 (2003).
92. RStudio Team. *Rstudio: integrated development for R. RStudio, Inc., Boston*, <http://www.rstudio.com > (2015).
93. R Core Team. *R a language and environment for statistical computing. R foundation for statistical Computing, Vienna, Austria*, <https://www.R-project.org/> (2021).
94. NCBI. *National Center for Biotechnology Information*, <https://www.ncbi.nlm.nih.gov> (2021).
95. Granato, I. S. C. *et al.* snpReady: A tool to assist breeders in genomic analysis. *Mol. Breed.* **38**, 102. https://doi.org/10.1007/s11032-018-0844-8 (2018).
96. Afghan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analysis. *Nucleic Acids Res.* **46**, 537–544 (2018).
97. Rasche, H. & Hiltemann, S. Galactic circos: User-friendly circos plots within the galaxy platform. *GigaScience* **9**, gaa06565 (2020).
98. Hiltemann, S., Rasche, H. & Gallardi, C. *Visualization with circos (galaxy training materials)* <https://training.galaxyproject.org/training-material/topics/visualisation/tutorials/circos/tutorial.html> (2021).
99. Wickham, H. ggplot2. *Wiley Interdisciplin. Rev. Comput. Stat.* **3**, 180–185. https://doi.org/10.1002/wics.147 (2011).
100. Lipka, A. E. *et al.* GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **28**, 2397–2399. https://doi.org/10.1093/bioinformatics/bts444 (2012).
101. Huang, M., Liu, X., Zhou, Y., Summers, R. M. & Zhang, Z. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* https://doi.org/10.1093/gigascience/giy154 (2019).

102. Steuernagel, B., Jupe, F., Witek, K., Jones, J. D. & Wulff, B. B. NLR-parser: Rapid annotation of plant NLR complements. *Bioinformatics* **31**, 1665–1667. https://doi.org/10.1093/bioinformatics/btv005 (2015).
103. Jones, P. *et al.* InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
104. Paradis, E. & Schliep, K. ape 50: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528. https://doi.org/10.1093/bioinformatics/bty633 (2019).
105. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, 256–259. https://doi.org/10.1093/nar/gkz239 (2019).
106. UniProt, C. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, 480–489. https://doi.org/10.1093/nar/gkaa1100 (2021).
107. Toparslan, E., Karabag, K. & Bilge, U. A workflow with R: Phylogenetic analyses and visualizations using mitochondrial cytochrome b gene sequences. *PLoS ONE* **15**, e0243927. https://doi.org/10.1371/journal.pone.0243927 (2020).
108. Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C. & Hochreiter, S. msa: An R package for multiple sequence alignment. *Bioinformatics* **31**, 3997–3999 (2015).
109. Charif, D. & Lobry, J. R. SeqinR 1.0–2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. *Struct. Approach. Sequence Evolut.* 207–232 (2007).
110. Yu, G. Using ggtree to visualize data on tree-like structures. *Current Protocols Bioinform.* **69**, e96 (2020).

Acknowledgements

We would like to thank laboratory technician Hanne Svenstrup at Nordic Seed A/S for contributing to the genotypic data collection and Anette Deterding and Marlene Walbrodt at Nordic Seed Germany GmbH (Nienstädt, Germany) for multiplying the N13 and N18 *Bgs* populations. We thank Professor Mogens Hovmøller and senior researchers Dr. Annemarie Fejer Justesen and Dr. Chris Khadgi Sørensen at the Institute of Agroecology Aarhus University (Flakkebjerg, Denmark) for their intellectual input for the study. The plant material included in this study was provided by Nordic Seed Germany GmbH (Nienstädt, Germany). 600K SNP genotyping was provided by Eurofind Genomics (Skejby, Denmark).

Author contributions

All authors were involved in the study design. N.M.V. performed the phenotyping, bioinformatic analysis, and visual output and wrote the manuscript. K.M. assisted in the DNA extraction of the germplasm. P.S.K oversaw the multiplication of *Bgs* populations. J.O. was responsible for all communication with Trait Genetics and Eurofins Genomics conducting the SNP genotyping. K.M., P.S., J.O. and A.J. were involved in the intellectual input for the study, including the interpretation of the results. All authors were involved in the conceptualization of the study and revision of the manuscript.

Funding

The research was funded by Innovation Fund Denmark (Grant No. 8053-00085B) and the Pajbjerg Foundation.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-02488-5.

Correspondence and requests for materials should be addressed to N.M.V.

Reprints and permissions information is available at www.nature.com/reprints.

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

Discovery of a novel powdery mildew (*Blumeria graminis*) resistance gene in rye (*Secale cereale* L.)

Nikolaj M. Vendelbo, Khalid Mahmood, Pernille Sarup, Peter S. Kristensen,
Jihad Orabi, and Ahmed Jahoor

Supplementary material

Caption of all supplementary files has been given below and a subset of supplementary figures and tables included in the dissertation. The extended supplementary materials has been provided at an open-access Zenodo data repository (https://zenodo.org/record/5510294).

Supplementary Table S1: Powdery mildew disease severity (% leaf area affected) of top-yielding hybrid rye (*Secale cereale* L.) and hybrid rye-population mixtures tested in the Danish official trials in 2017 at the Tinglev trial site in Southern Denmark.

Supplementary Table S2: Characteristics of informative 600K SNP markers on each chromosome in Nordic Seed hybrid rye (*Secale cereale* L.) breeding germplasm (n=180).

Supplementary Table S3: Infection type (0-4) of 180 Nordic seed hybrid rye (*Secale cereale* L.) breeding lines and hybrid controls phenotyped in greenhouse using three distinct powdery mildew populations.

Supplementary Table S4: Top 1000 powdery mildew resistance associated 600K SNP markers from MLM Genome-wide association study on 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines using BLUP estimated resistance value across three *Bgs* populations.

Supplementary Table S5: Top 1000 powdery mildew resistance associated 600K SNP markers from BLINK Genome-wide association study on 180 Nordic Seed hybrid rye (*Secale cereale* L.) elite restorer breeding lines using BLUP estimated resistance value across three *Bgs* populations.

Supplementary Table S6:

- **A)** Genotype of powdery mildew resistance associated SNP in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines.
- **B)** Genotype of powdery mildew resistance associated markers on 7RL in 88 Nordic Seed hybrid rye (*Secale cereale* L.) non-restorer germplasm breeding lines.

Supplementary Table S7: Mapping position of 600K SNP markers significantly associated with powdery mildew resistance in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines on 'Lo7' and 'Weining' reference genome.

Supplementary Table S8: Nucleotide-binding leucine-rich repeat proteins in the distal region of chromosome arm 7RL chromosome in the 'Lo7' and 'Weining' rye (*Secale cereale* L.) reference genomes.

Supplementary Figure S1: Two inbred hybrid rye (*Secale cereale* L.) breeding lines displaying complete susceptibility (left) and resistance (right) towards powdery mildew

Supplementary Figure S2: Manhattan plots for discovery of powdery mildew resistance associated SNP markers by genome-wide association study (GWAS) MLM in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines using 261,406 informative markers. Lines were phenotyped for resistance towards three distinct field populations from Northern Germany in 2013 (N13), 2018 (N18), and Denmark 2020 (D20) and a resistance value estimated using BLUP. The purple line represents the Bonferroni adjusted significance threshold based on informative markers.

Supplementary Figure S3: Manhattan plots for discovery of powdery mildew resistance associated SNP markers by genome-wide association study (GWAS) BLINK in 180 Nordic Seed hybrid rye (*Secale cereale* L.) breeding lines using 261,406 informative markers. Lines were phenotyped for resistance towards three distinct field populations from Northern Germany in 2013 (N13), 2018 (N18), and Denmark 2020 (D20) and a resistance value estimated using BLUP. The purple line represents the Bonferroni adjusted significance threshold based on informative markers.

Supplementary Figure S4: Phylogenetic tree of 1,027 nucleotide-binding leucine-rich repeat (NLR) proteins NB-ARC domain in the 'Weining' rye (*Secale cereale* L.) reference genome. NB-ARC domains of known NLR-type *R* genes have been included as reference. NLR genes residing in a powdery mildew resistance associated block on chromosome arm 7RS subteleomeric region (7RL PM) have been colored in teal.

Supplementary material S1: NB-ARC sequence of nucleotide-binding leucine-rich repeat (NLR) proteins in powdery mildew resistance associated region on chromosome arm 7RL in the rye reference genome 'Lo7' and 'Weining'.

Project overview

Key findings

In the pursuit to address the project objectives and answer to formulated hypothesis, the following key findings were made

Manuscript I:

- Parental populations in the Nordic Seed hybrid rye breeding germplasm display a large genetic seclusion with potential of attaining a high level of heterosis in hybrids.
- The restorer population is characterized by a relative high level of genetic diversity, effective population size and a low linkage disequilibrium.
- The NRG&CMS population is characterized by a relative low level of genetic diversity, effective population size and a high linkage disequilibrium.
- The fertility control system constitutes a population determining parameter by influencing the rate of introducing novel genetic variation to the parental populations.
- In order to expand the genetic reservoir within the non-restorer germplasm population, a broadening of the genetic base is recommended through guided introgression of novel material (population varieties, landraces, genetic resources).

Manuscript II

- A novel major G-type male-fertility restoration gene, provisionally denoted *RfNOS1*, was discovered on chromosome arm 3RL in the Nordic Seed hybrid rye breeding germplasm.
- Absence of RFL-PPR genes on chromosome arm 3RL and the close proximity of *RfNOS1* associated marker to a mTERF gene co-expressed in Nordic Seed hybrid cv. Stannos and cv. Helltop suggest that *RfNOS1* might be a non-PPR *Rf* gene belonging to the mTERF gene family.
- With no previous record of a major *Rf* gene in fertility control systems of rye on chromosome arm 3RL nor on syntenic regions in barley and wheat our finding constitute a novel discovery with potential value for implementation in other hybrid breeding systems in Triticeae.
- We found evidence of a potential minor G-type *Rf* gene distal to *RfNOS1* on chromosome arm 3RL.

Manuscript III

- The assayed subset of Nordic Seed hybrid rye breeding germplasm was found to display a high level and diverse spectra of leaf rust resistance.
- A novel leaf rust *R* gene, provisionally denoted *PrNOS1*, was discovered on chromosome arm 7RS.
- The *PrNOS1* gene resided in close proximity to a large cluster of *Lr1* orthologous NLR genes. With wheat leaf rust *R* gene *Lr1* situated on chromosome arm 5DL syntenic to rye 7RS our findings suggest that *PrNOS1* might be a *Lr1* ortholog in rye.
- Gene mining by GWAS in the restorer population is strongly influenced by several population genetic characteristics identified in Manuscript I preventing *R* gene and QTL discovery at the assayed sample size.
- Breeding lines displaying quantitative resistance were identified in both parental populations.

Manuscript IV

- Detailed protocol for recovery, multiplication, single pustule isolation, large scale inoculation, and infection typing was developed for leaf rust in rye and made publicly available on an open-access data repository
- Five intermediate effect leaf rust resistance QTLs were identified on chromosome arms 1RS, 1RL, 2RL, 5RL and 7RS in the restorer population using 600K GWAS.
- *In silico* validation of the 60K Triticeae NLR bait library revealed a near-complete capture of the NLR repertoire in the 'Lo7' reference genome.
- *In vitro* assessment of the 60K Triticeae NLR bait library performance in rye confirmed that a 80% sequence similarity is sufficient for bait target enrichment.
- SMRT RenSeq led to an incomplete target capture of the NLR repertoire in some of the selected lines.
- Four candidate leaf rust *R* genes in rye were identified using SMRT AgRenSeq-GLM.
- Most promising candidate leaf rust *R* gene physically co-localized with molecular markers delimiting *Pr3* on chromosome arm 1RS and top-most resistance associated QTL in the panel.
- The candidate leaf rust *R* gene on chromosome arm 1RS encoded a canonical NLR gene distinguished from the susceptible allele by a large deletion and C-terminal extension of the LRR domain.

Manuscript V

- A marker-map was developed for the rye 600K high-density SNP array using the 'Lo7' reference genome and made publicly available on an open-access data repository
- Assessment of the 600K high-density SNP arrays performance in the Nordic Seed hybrid rye breeding germplasm led to finding of a homogeneous marker distribution, moderate level of marker informativeness, and an average marker to marker distance of 25.54 kb.
- The assayed subset of Nordic Seed hybrid rye breeding germplasm diverged in resistance to powdery mildew with the NRG&CMS population characterized by a moderate level of resistance and the restorer population a low level
- Using GWAS a powdery mildew *R* gene, provisionally denoted *PmNOS1*, was discovered on chromosome arm 7RL.

Table 1 Physical position of major and minor quantitative trait loci associated with agronomic important traits in the Gülzow based hybrid rye (*Secale cereale* L.) elite breeding germplasm at Nordic Seed on the 'Lo7' reference genome. The traits, leaf rust resistance (*Pr*), powdery mildew resistance (*Pm*), and restoration of male fertility (*Rf*), were identified by genome wide association study using a 600K or 30K SNP array.

Chromosome	Arm	Position (Mbp)	Trait	Phenotypic variance explained	Population	Marker	SNP array	Manuscript
1R	S	115.55	*Pr*	13.13%	R	AX-99251803	600K	IV
1R	L	625.54	*Pr*	9.34%	R	AX-99805135	600K	IV
2R	L	818.90	*Pr*	9.62%	R	AX-99478491	600K	IV
3R	L	747[1]	*Rf*	20.24%	F_2	AX-158558079	30K	II
3R	L	807.11	*RfNOS1*	7.45%	F_2	AX-99346444	30K	II
5R	L	770.21	*Pr*	9.26%	R	AX-99776626	600K	IV
7R	S	3.41	*PrNOS1*	21.06%	NRG&CMS	AX-99555960	600K	III, IV
7R	S	26.93	*Pr*	12.12%	R	AX-99684185	600K	IV
7R	L	892.01	*PmNOS1*	15.48%	NRG&CMS	AX-99886178	600K	V

[1] Marker derived from the Wheat 90K SNP array with no mapping position on the 'Lo7' reference genome. Mapped using pairwise linkage disequilibrium analysis to markers with accurate mapping position on the 'Lo7' reference genome

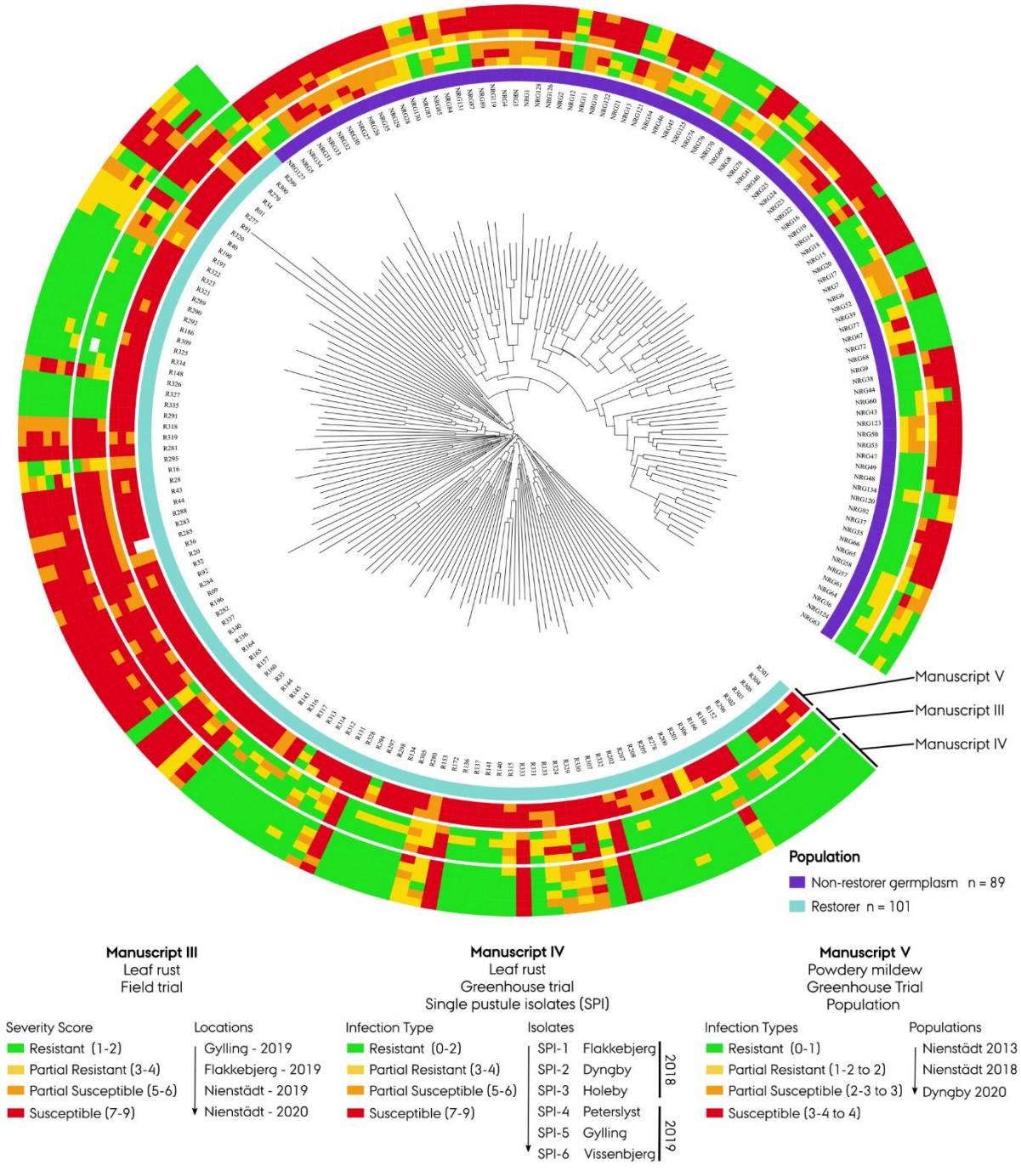


Figure 6 Circular neighbor joining dendrogram of the Gülzow based hybrid rye (*Secale cereale* L.) elite breeding germplasm at Nordic Seed. Concentric tracks depict phenotypic scorings on leaf rust (Manuscript III, IV), and powdery mildew (V). Arrow indicates the organization of tracks, from inside to outside of the concentric circle.

General discussion

Despite decades of declining rye acreage in Northern and Eastern Europe, rye cultivation seems to be stabilizing around 2000 hectares, remaining an important crop in the European agriculture (Figure 1.4, FAO, 2021). In Germany and Denmark the prevalence of hybrids derived from the Carsten and Petkus gene pool has had a notable effect on rye breeding and cultivation. Since the expansion of rye hybrids during the 1990s the primary emphasis in resistance breeding has been to address the inherent issue of ergot susceptibility related to the impartial male-fertility restoration in the P-type CMS used in these systems (Mielke, 1993; Mirdita et al., 2008; Mirdita and Miedaner, 2009; Miedaner et al., 2010; Miedaner et al., 2021). Meanwhile the foliar diseases leaf rust and powdery mildew were given a lower priority (Miedaner et al., 2002). Accentuated by the prevalence of susceptibility towards particularly leaf rust in released hybrid rye cultivars we set out to investigate the genetics underlying resistance towards the two foliar diseases in a Gülzow hybrid rye breeding germplasm distinct from the Petkus and Carsten gene pools (Figure 4.2, Figure 4.3, Sortsinfo, 2021). In brief summary we identified a high level of leaf rust resistance and moderate level of powdery mildew resistance in the germplasm, constituting a promising genetic resource for development of resistant hybrid rye cultivars (Table 1, Figure 6). Here we discuss the implications of our findings and observations in the context of state-of-the-art literature within the field.

Powdery mildew in rye, a bygone threat or latent menace?

While the powdery mildew pathogen in rye, *Blumeria graminis* f. sp. *secalis*, is prolific across rye growing regions in Northern Europe, for which reason or reasons does it not constitute a more important biotic stress factor in rye production? (Figure 4.3A, Sortsinfo, 2021). To address this question each component of the disease triangle underlying powdery mildew disease in rye will be discussed.

In the global populations of powdery mildew on barley and wheat, high levels of virulence complexity and pathotype diversity have been recurrently demonstrated (Cowger et al., 2018; Dreiseitl, 2018; Xue et al., 2021; Zhou et al., 2021). Assuming that the *Bgs* population in Northern Europe exhibit similar population characteristics, associated with a high evolutionary plasticity and adaptive capacity, a wide variety of virulent *Bgs* pathotypes exist in the environment. In accordance, we confirmed that the assayed hybrid rye breeding germplasm segregated for resistance towards powdery mildew, with a large subset of lines (particularly

restorer lines) highly susceptible towards the disease (Figure 6). The loss of resistance observed in nine NRG lines during the period 2013 to 2018 in Northern Germany, further support the presence and emergence of novel virulent *Bgs* pathotypes in the Northern European population (Manuscript V).

The rare occurrence of severe powdery mildew outbreaks and low to intermediate levels of disease severity observed in the Danish official trials during the past 25 years could suggest a high level of inherent resistance in the top-yielding hybrid rye cultivars (Figure 4.3A, Sortsinfo, 2021). On basis of these records we selected cv. KWS Serafino as 'resistant' control for the greenhouse trial in which the assayed hybrid rye breeding germplasm was phenotyped for resistance towards powdery mildew (Manuscript V). In contrast we found that cv. KWS Serafino was highly susceptible at the seedling stage towards all three powdery mildew populations tested. This observation could be explained by the presence of an adult plant resistance (APR) *R* gene or QTLs in cv. KWS Serafino, conferring a broad-spectrum resistance towards powdery mildew in the later plant stages (Li et al., 2014). However, while both major APR *R* genes and intermediate-effect QTLs have been identified in wheat, none of the *Pm R* genes identified in rye have been shown to confer an APR-type resistance (Li et al., 2014; Mohler and Stadlmeier, 2019). Intriguingly a collection of inbred rye lines at the University of Hohenheim belonging to the Petkus and Carsten gene pool was found to exhibit a broad spectrum of resistance towards powdery mildew with majority of lines displaying a quantitative type of resistance (Kast and Geiger, 1982; Miedaner et al., 1993). With majority of released hybrid rye cultivars derived from the Petkus and Carsten gene pool this might indicate an inherent level of quantitative resistance towards powdery mildew in these hybrids (Bolibok-Bragoszewska et al., 2014; Targonska et al., 2016). Prevalence of quantitative resistance amongst release hybrids in Northern Europe could be an important factor modulating the importance of powdery mildew in rye (Figure 4.3A). Across the assayed hybrid rye breeding germplasm distinct from the Petkus and Carsten pools we also identified lines displaying a quantitative resistance response at seedling stage. As a consequence of a low level of natural powdery mildew infection in the field trials we could, however, not confirm whether any lines displayed a APR-type resistance.

In terms of the environmental factor the powdery mildews on Triticeae, *Blumeria graminis* (*Bg*), are highly dependent on certain environmental conditions and cues during specific life stages, determining their ability to cause severe epidemics (Aust and Hoyningen-Huene, 1986). The

sporulation and germination of *Bg* spores is highly temperature sensitive with an optimum around 20 degrees at a high relative humidity, and ascospores dependent on a prolonged period of surface moisture to differentiate and release (Ward and Manners, 1974; Adams Jr et al., 1986; Jankovics et al., 2015). Growing predominantly extracellular as a conspicuous mycelial mat *Bg* is, furthermore, directly exposed to fluctuations in the environment. Overwintering as dormant mycelia on the aerial parts of volunteer- or newly established autumn sown crops, harsh winters with prolonged periods of bare-surface frost is likely to have a detrimental effect on the survival rate of *Bg* (Jorgensen, 1988). The importance of these environmental factors in powdery mildew epidemiology are emphasized in developed forecasting models incorporating precipitation, temperature, sun radiation and humidity to predict the spatiotemporal occurrence and scale of powdery mildew (Zhang et al., 2014; Zhao et al., 2018; Gu et al., 2020). During the past 25 years in Denmark the only incidence of severe powdery mildew outbreak in rye in 2010 coincided with the coldest year during the measured period, with an average surface temperature nearly 2ºC below the yearly average measured during the period (Figure 4.3A, DMI, 2021; Sortsinfo, 2021). In Denmark the year 2010 was characterized by a warm spring followed by a colder May and June with frequent precipitation and moderate temperatures around or below 20ºC (DMI, 2011). These observations could suggest that the rarity of prolonged periods with precipitation and moderate-to-low temperature co-occurring during spring and early summer in Denmark are important factors in modulating the importance of powdery mildew in rye.

Accentuated by the reliance of *Bgs* on certain environmental conditions climate changes are likely to have a profound effect on powdery mildew in rye. Since the onset of the 20th century the global mean surface air temperature has increased by 0.74 ± 0.14ºC and is predicted to increase by an additional 1.0 to 3.7ºC by the end of the 21th century (Delworth and Knutson, 2000; Anderson et al., 2016). Consensus amongst climate projections indicate an increase of extreme weather events, including heat waves, droughts and heavy precipitation (Rajczak and Schär, 2017; Spinoni et al., 2018). In Denmark the annual precipitation is predicted to increase by 15% during the 21th century in a seasonal pattern, with a decrease in precipitation and rain fall events during summer (Olesen, 2020).

In 2019, Denmark experienced the fourth warmest year since the first national temperature recording in 1873 with a mean surface air temperature 0.7ºC above average during the period 1991 to 2020 (Rubek, 2019). Precipitation in 2019 followed the projected seasonal pattern with

a wet spring and autumn, and prolonged periods between rainfall events during the summer. In accordance to the climate projections the weather conditions observed in 2019 in Denmark are likely to constitute a near-future standard. The field trial site in 2019 near Odder (Gylling) in Central Jutland, therefore, provide a glimpse into the potential occurrence of powdery mildew in rye under these climatic conditions in Denmark. Across the hybrid rye breeding germplasm assayed at the site we observed a low level of natural powdery mildew infection. Powdery mildew was observed on 52 out of 190 lines with an 1% average leaf area covered by the disease. Several environmental conditions were, however, conducive during the trial including an warm autumn in 2018 with windows of high humidity and precipitation allowing the establishment of powdery mildew in the autumn sown crop, a mild winter and short periods with moderate temperature, humidity and frequent rainfall during late April, mid-May and mid-June (Figure 7).

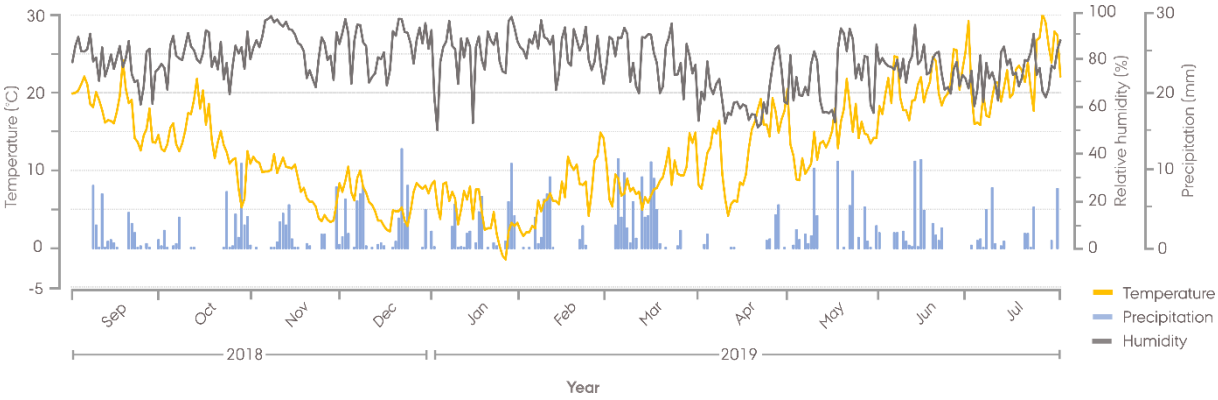


Figure 7 Highest daily temperature, precipitation and humidity measured at a Danish Meteorological Institute weather station near Odder in Central Jutland, Denmark, from September 2018 to August 2019 (DMI, 2021).

The absence of co-occurring periods with moderate temperature around 20ºC, frequent precipitation and high humidity during April to June are likely important factors underlying the low level of powdery mildew infection at the trial (Figure 7). While a momentary glimpse these observations could infer that powdery mildew in Northern Europe is becoming a less important biotic stress factor in hybrid rye as a result of climatic changes, predominantly due to warmer and dryer summers. However, moderate temperatures in autumn and mild winters favors an increased establishment and survival rate of *Bgs*. Years with high temperatures during spring co-occurring with the expected increase in precipitation during the period could, likewise give cause to severe epidemics early in the growing season. Furthermore, rye is predominantly grown on sandy soils in Denmark, and despite ryes inherent tolerance to drought conditions the

necessity of irrigation will likely increase. Irrigation of hybrid rye field crops might lead to high incidence in powdery mildew as observed for irrigated wheat crops in China (Tang et al., 2017). During the past 25 years in Denmark powdery mildew has constituted a less important biotic stress factor in rye despite our observations of a virulent *Bgs* population in Northern Europe with evidence of novel pathotypes emerging. A potential high level of inherent quantitative resistance and/or APR in released hybrid rye cultivars could explain the low level of powdery mildew disease severity observed in the Danish official trials. The recent climatic changes in Northern Europe is, furthermore, an important factor contributing to a recurrently unfavorable environment for the proliferation of powdery mildew during the early summer months. However, with increasing weather instability as a product of climate change, powdery mildew can still become an important biotic stress factor in certain years. Increase in irrigation practices can likewise lead to an increase in powdery mildew incidence. Conclusively, while powdery mildew is not a 'bygone threat' it is likely to become confined as an important biotic stress factor under certain uncommon climatic conditions and agricultural practices.

Leaf rust in rye, a forgotten detriment

Recordings of leaf rust in top-yielding hybrid rye cultivars and population varieties tested in Denmark during the past 25 years provide extensive documentation on the importance of leaf rust as an important biotic stress in rye (Figure 4.2A, Sortsinfo, 2021). During the period several outbreaks of leaf rust has occurred (2000, 2005, 2009, 2014), in addition to a high yearly mean disease severity across the evaluated panel as observed in 2019 (Figure 4.2A-B). The low level of inherent resistance in the predominant hybrid rye breeding gene pools Petkus and Carsten, and the low effectiveness of leaf rust *R* genes *Pr3*, *Pr4*, and *Pr5* was reported already in early 2000s (Miedaner et al., 2002; Roux et al., 2004). Despite inadequate documentation of leaf rusts impact on yielding components in modern released hybrid rye cultivars, it is reasonable to presume that leaf rust is causing considerable grain yield losses and quality reductions on a yearly basis in Denmark (Miedaner and Sperling, 1995; Morgounov et al., 2015). Nonetheless, the incentive of breeding for leaf rust resistance has until now been eclipsed by breeding for environmental stable and complete male-fertility restoration, ergot resistance, lodging resistance and yield (Falke et al., 2009; Miedaner et al., 2010; Bernal-Vasquez et al., 2017; Laidig et al., 2017; Miedaner et al., 2018). With recent years breeding effort having considerably improved male-fertility restoration and ergot susceptibility in released hybrid rye

cultivars, this may provide new opportunities to address the inherent susceptibility to leaf rust (Miedaner et al., 2017; Wilde et al., 2019; Miedaner et al., 2021).

The emphasis for commercial breeding programs to prioritize the selection and improvement of leaf rust resistance is accentuated by several factors. Unlike powdery mildew, projected to display a fluctuating positive or negative regional pattern related the effect of climatic changes on disease incidence, all current projections concur that incidence, frequency and duration of leaf rust in cereals will likely increase in Northwestern and Eastern Europe as a result of climatic changes (Bregaglio et al., 2013; Racca et al., 2015; Junk et al., 2016; Caubel et al., 2017; Wójtowicz et al., 2017; Launay et al., 2020; Wójtowicz et al., 2020). The increase in average surface air temperature has been related to several stages in the disease cycle of cereal leaf rusts. In wheat and triticale the latency period, an important epidemiological parameter of leaf rust for prediction of epidemic progression, has been predicted to decrease in according to rising temperatures (de Vallavieille-Pope et al., 2000; Wójtowicz et al., 2017; Wójtowicz et al., 2020). As a result of a shorter latency period the cereal rusts will be capable of completing a larger number of clonal infection cycles during periods of conducive conditions, accelerating the disease progression in the field exponentially (Van den Berg et al., 2013; Suffert and Thompson, 2018). The future climatic conditions during autumn and winter will likewise favor leaf rust infection of the autumn sown crop, facilitating an earlier establishment and survival of the leaf rust pathogen (Roelfs, 1989; Launay et al., 2020). However, projections also emphasize that a more frequent absence of precipitation in extended periods during spring and summer can likewise lead to a haltered progression of leaf rust (Launay et al., 2020). To envision the impact of leaf rust under standard near-future climatic conditions the trial site in 2019 near Odder (Gylling) in Central Jutland, as discussed previously for powdery mildew, constitute an ideal example (DMI, 2021). Despite a cold and dry April and early May in 2019, we observed a high disease pressure of leaf rust at the non-inoculated trial site with susceptible lines displaying between 10 to 50% leaf area covered by the disease (Figure 7). At the site the ‘resistant’ control, hybrid cv. KWS Serafino, was likewise found to displayed a moderate level of leaf rust severity with an average of 10% leaf area covered. The high temperatures and moderate level of humidity during the summer provided ideal conditions for extensive dew formation, facilitating the repetitive infections of leaf rust by clonal reproduction While fluctuation in the occurrence of leaf rust in rye is to be expected as a result of a more unstable future climate, the steady increase in surface air temperature has a general elevating effect on the disease (Roelfs, 1989).

During the last 25 years the average yearly disease severity of leaf rust across the evaluated panel in the Danish official trials has increased by 44% despite yearly fluctuations in climatic conditions (Figure 4.2A). While this increase is influenced by a multitude of factors, including loss of *R* gene effectiveness (Roux et al., 2004), the period was characterized by an increase of 0.3ºC in the average yearly surface temperature in Denmark (Rubek et al., 2020).

Another factor emphasizing the importance of addressing leaf rust in rye are the inherent properties of the causative agent *Puccinia recondita* f. sp. *secalis* (*Prs*). Unlike the clonal reproducing leaf rust fungi on wheat (Kolmer et al., 2007; Goyeau et al., 2012), evidence suggest a mixed reproduction mode in the *Prs* population (Miedaner et al., 2011). The potential of frequent sexual reproduction, due to a widespread alternate host *Anchusa arvensis* L. in the Danish flora (Anikster et al., 1997; Frederiksen et al., 2012)., likely induce an evolutionary plasticity enabling the pathogen to adapt and overcome deployed host resistance and chemical control (Kema et al., 2018; Drenth et al., 2019; Pereira et al., 2020). Consistent with the observations made by Miedaner et al. (2011) in the German *Prs* population, we observed that five out of the six Danish *Prs* SPI tested across a subset of restorer breeding lines displayed a distinct virulence phenotype (Manuscript IV, Figure 4B). The combination of a high pathotype diversity, virulence complexity, potential frequent sexual recombination and capacity to disperse spores over long distances underlines the potential threat residing in the Northern European *Prs* population. This threat is likely further exacerbated in the context of near-future climate change (Launay et al., 2020; Wójtowicz et al., 2020).

The control of leaf rust in rye predominantly depend on the application of chemical control agents in Denmark as a consequence of the low level of inherent resistance in released hybrid rye cultivars (Figure 4.2, Jørgensen et al., 2013; Sortsinfo, 2021). In the European Union (EU) a stringent system is imposed for authorizing and controlling the use of pesticides as a legislative basis for the safe and sustainable use of pesticides (EC, 2009). Political incentive in the EU is an accelerated transition of the European agriculture towards more sustainable practices including means to minimize the current dependency on pesticides. In the recently published EC Farm to Fork Strategy, part of the European Green Deal, a reduction target of 50% in overall use of pesticide in 2030 was presented (EC, 2020; Silva et al., 2021). In Denmark a further restrictive authorization scheme has been imposed on pesticides as a precautionary mean to minimize potential non-target effect on environment and human health (MVFM, 2017; Thorling et al., 2021). The prohibition of three broad-spectrum fungicides in 2021 due to their

content of epoxiconazole has further narrowed the list of authorized pesticides in Denmark for the control of fungal disease in cereals, including leaf rust in rye (Nielsen, 2021). The decreasing number of authorized pesticides in Denmark, rising environmental taxation and restriction on their use underline the necessity of improving the leaf rust resistance in released hybrid rye cultivars as a sustainable alternative to pesticides. This is further emphasized by the considerable organic production of rye in Denmark relying on host resistance for the control of leaf rust, a production system that is likely to expand in near-future (LBST, 2020b).

The low level of resistance and increase in disease severity across released hybrid rye cultivars and population varieties towards leaf rust during the past 25 years in Denmark underlines the growing importance of the foliar disease as a biotic stress factor in rye. While the control of leaf rust currently depend on the application of pesticides the strict authorization process, increase in prohibitions, restrictive use, and 50% targeted reduction of pesticide use in the EU by 2030 emphasizes the need of alternative control methods. This need is further emphasized by (i) the projected increase in leaf rust incidence, frequency and duration under near-future climatic conditions in Northern and Eastern Europe, (ii) recent insight into the *Prs* population in Northern Europe indicating a highly diverse, virulent and adaptive pathogen, and (iii) a likely expansion of the rye area under organic farming. In order to sustain rye productivity under near-future conditions resistance towards leaf rust should be prioritized in commercial breeding programs in order to improve the level of resistance in released hybrid cultivars as a sustainable alternative to pesticides.

You got me, now what? Breeding of leaf rust resistant hybrids in rye

While the future importance of powdery mildew disease in rye in Northern Europe is equivocal, current evidence presented above underlines the evident need for developing leaf rust resistant hybrid cultivars (Figure 4.2A, Figure 4.3A). In stark contrast to the low level of leaf rust resistance reported in the Petkus and Carsten gene pools (Miedaner and Sperling, 1995; Solodukhina, 2002), we observed a high level- and diverse spectra of resistance in the assayed Gülzow based germplasm. In both parental populations we identified breeding lines displaying qualitative and quantitative resistance towards leaf rust (Figure 6, Manuscript III Figure 1). Through GWAS we identified a novel *R* gene, provisionally denoted *PrNOS1*, on chromosome arm 7RS and five QTLs on chromosome arms 1RS, 1RL, 2RL, 5RL and 7RS (Table 1). These findings suggest a broad genetic base of leaf rust resistance in the Gülzow germplasm,

constituting a valuable genetic resource for addressing the evident need of resistant hybrid rye cultivars.

In the 'short term' development of a leaf rust resistance hybrid rye cultivar could be initiated directly, involving four steps, (i) selection of multiple resistant lines in each parental pool guided by phenotypic data and/or markers to select parental lines with complimenting *R* genes, (ii) guided crossing of parental lines using genomic prediction models (Auinger et al., 2016), (iii) screening of three-way hybrids in field under high leaf rust disease pressure and in non-treated trials to assess yielding characteristics and agronomic important traits, (iv) submission of selected hybrid for evaluation in the Danish national trials (Tystofte, 2021).

The identified leaf rust resistance associated markers can be implemented for (i) marker assisted backcrossing (MABC) of single or multiple leaf rust *R* genes into parental lines of elite hybrid rye cultivars within each population to accelerate recurrent parent genome recovery (Yadav et al., 2015; Yang et al., 2019), (ii) MABC of *R* genes into the opposing parental population, and (iii) MAS in existing breeding programs for development of new resistant parental lines (Beukert et al., 2020). The success of implemented MAS or MABC strategies, however, depend on the quality of markers (Cobb et al., 2019). In our study several biological and statistical factors are expected to have influenced the quality of identified markers (Table 1, Platten et al., 2019), of which the primary is likely the inherent genetic characteristics of the individual parental populations (Manuscript I). The large divergence in LD between parental populations is possibly diminishing the statistical power of GWAS to ascertain a marker-trait association in collective analysis of the germplasm (Manuscript III Figure 3, Sorkheh et al., 2008; Korte and Farlow, 2013). Furthermore, the divergence in LD is likewise impeding the application of markers identified in the high LD NRG&CMS population for MAS in the low LD restorer population due to decay of linkage between non-functional markers and the causative *R* gene or QTL. The combination of a potential broad genetic base of leaf rust resistance with the high genetic differentiation of parental populations could infer a distinctive genetic architecture underlying leaf rust resistance in each population characterized by multiple population specific *R* genes and QTLs. The collective analysis of the germplasm would, therefore, lead to a further decrease in frequency of the individual *R* genes and QTLs, reducing their effect and phenotypic variation explained needed for GWAS to establish a reliable marker-trait association (Platten et al., 2019).

In order to address these limitations and improve marker quality for successful MAS of identified leaf rust *R* genes and QTLs, we hypothesize that GWAS should be conducted on the individual parental populations instead with consideration of their genetic characteristics. In the NRG&CMS population GWAS could likely be done successfully on a low sample size of 100 to 150 individuals genotyped on the low-density custom Illumina Infinium 30K SNP array due to the high LD and narrow genetic profile (Haseneyer et al., 2011; Wang et al., 2014). In the restorer population, GWAS for the identification of widespread *R* genes likely requires a larger sample size of more than 200 individuals genotyped on the high-density 600K SNP array (Bauer et al., 2017). For the identification of less prevalent, potentially valuable, *R* genes and intermediate effect QTLs in the restorer population, development of multi-parent populations (MPP) could be a suitable strategy (Scott et al., 2020). In MPPs, rare *R* genes and QTLs present in the founders are increased to an intermediate level improving the statistical power of GWAS to ascertain a marker-trait association (Septiani et al., 2019). Common MPP designs used in the similar hybrid breeding system of maize for the discovery of resistance QTLs include multi-parent advanced generation inter-cross (MAGIC) (Butrón et al., 2019; Septiani et al., 2019) and nested association mapping (NAM) (Figure 8, Li et al., 2018).

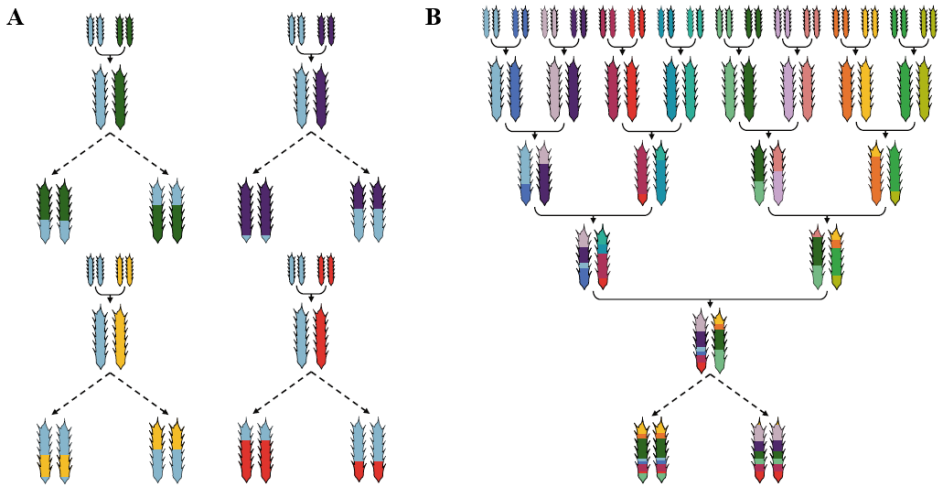


Figure 8 Multi-parent populations, **A**) nested association mapping (NAM) panel developed from series of biparental crosses against a common founder, and **B**) 'funnel type' multi-parent advances generation inter-cross (MAGIC) panel developed from equally balanced crosses of different founders. Dotted lines represents generations of selfing and single seed descent for generation of recombinant inbred lines (RILs). After (Scott et al., 2020).

The MAGIC design distinguishes itself from NAM by several features including a higher haplotype diversity and increased recombination events due to several generations of intercrossing (Ladejobi et al., 2016; Scott et al., 2020). While the establishment of MAGIC population is time requiring, it may in addition to its inherent value for detection of *R* genes and

QTLs represent an elite material for improvement of the parental populations (Arrones et al., 2020). Some of the developed recombination inbred lines (RILs) may, given a sufficient population size, exhibit a desirable combination of pyramided *R* genes and QTLs, hence saving the need for time consuming MAS or MABC (Huynh et al., 2018; Zaw et al., 2019). For the establishment of an restorer-based MAGIC population, selected founder lines, often 4, 8 or 16 founders, should span the entire populations genetic and phenotypic diversity of the germplasm to optimize the statistical power of GWAS (Scott et al., 2020). In order to capture the diversity of leaf rust resistance spectra within the restorer population resistant founders should both comprise lines displaying a qualitative resistance, for detection of *R* genes, and quantitative resistance, for detection of slow rusting *R* genes and QTLs (Figure 6). For augmenting the breeding value of developed RILs, selected founder lines susceptible to leaf rust should instead exhibit other agronomic traits of interest, *e.g.*, strong yielding characteristics, resistance to ergot or fusarium, straw strength, feeding value, etc. (Novakazi et al., 2020). After establishment of the MAGIC population the lines should be phenotyped for resistance towards leaf rust both at seedling stage for detection of *R* genes, and in field at later plant stages under high leaf rust disease pressure to identify slow rusting *R* genes and QTLs. While an expensive endeavor, the restorer MAGIC population should be genotyped using the 600K high-density SNP array to facilitate the detection of *R* genes and QTLs by GWAS and identification of high quality markers tightly linked to these (Cobb et al., 2019). Hereafter, prior to implementation of markers in MAS or MABC in the hybrid rye breeding program, markers should be validated on a distinct training set to ensure their reliability (Mourad et al., 2018). The validation process should conclude by calculation of the five core marker quality metrics defined by Platten et al. (2019). A similar approach could be conducted for the NRG&CMS population, however, it would require the establishment of two genetically distinct MAGIC populations since the hybrid breeding scheme of rye involves a double hybrid seed mother.

The high level and diverse spectra of leaf rust resistance in the Gülzow germplasm provides a valuable genetic resource for the direct development of leaf rust resistant hybrid rye cultivars. Identification of high quality makers for MAS and MABC strategies is, however, compromised by the large LD divergence between parental populations, suggesting the need for separation of populations in future GWAS. For the discovery of less prevalent and potential valuable *R* genes and slow rusting resistance QTLs establishment of MAGIC populations could be a suitable strategy, likewise providing an elite material for germplasm improvement. In addition to the

high quality markers identified through this pursuit, potential developed RILs with pyramided *R* genes and slow rusting resistance QTLs could constitute a valuable genetic resource for enhancing the durability of leaf rust resistance in developed cultivars.

Conclusion and perspectives

The current evidence suggest that powdery mildew in Northern Europe might constitute a less important biotic stress factor in rye, likely to become confined to sporadic years with conducive conditions and certain agricultural practices under near-future climate conditions. In contrast, leaf rust is becoming a biotic stress factor of growing importance in rye with current projections concurring on an augmented effect of climatic changes on leaf rust in Northern and Eastern Europe. The current sustainable transition of agricultural practices in the EU emphasizes the need for breeding initiatives to address the low level of leaf rust resistance in released hybrid rye cultivars as a sustainable alternative to the current dependency on pesticides. In contrast to the predominant hybrid rye breeding gene pools Petkus and Carsten we identified a high level and diverse spectra of leaf rust resistance in the Gülzow germplasm, constituting a valuable genetic resource for the development of resistance hybrid cultivars. In order to improve the quality of identified molecular markers for MAS and MABC strategies implemented to improve leaf rust resistance in the hybrid rye breeding program we propose (i) separation of parental populations for GWAS due to considerable divergence in LD, (ii) increase in sample size of the restorer population due to a higher level of genetic diversity and low LD, and (iii) establishment of MPPs and/or MAGIC population(s). The MAGIC population(s) would serve multiple purposes, (i) detection of less prevalent *R* genes and QTLs, (ii) identification of tightly linked high quality markers, (iii) development of elite material for germplasm improvement, and (iv) potential development of RILs with pyramided *R* genes and QTLs. If resistance to leaf rust is prioritized, hybrid rye cultivars with a high level of leaf rust resistance can be developed from the Gülzow germplasm within few years using obtained phenotypic and molecular data.

The genetic characteristics of the *Prs* population in Northern Europe indicates a high evolutionary capacity to overcome deployed resistance such as observed for the leaf rust *R* genes *Pr1, Prs3*, *Prs4*, and *Prs5* (Roux et al., 2004; Roux and Wehling, 2010; Miedaner et al., 2011). In order to enhance the durability of deployed *R* genes, the potential RILs developed in the established MAGIC population(s) displaying various combinations of pyramided *R* genes and QTLs could prove a convenient and valuable genetic resource (Pilet-Nayel et al., 2017).

Currently pyramiding strategies using MAS and MABC is attracting a lot of scientific attention for the improvement of disease resistance in crops and has been successfully deployed for stripe (Liu et al., 2020), stem (Zhang et al., 2019), and leaf rust (Sharma et al., 2021) resistance in wheat. Broadening the distribution of potential slow rusting resistance QTLs and *R* genes in the germplasm could likewise contribute to an inherent level of leaf rust resistance in parental lines and enhancement of resistance durability (Huerta-Espino et al., 2020). While quantitative resistant lines was observed in both parental populations, one restorer line R190 was found to display a stable and strong quantitative resistance at seedling stage and high level of resistance at adult stage in the 2019 and 2020 field trial (Figure 6, Figure 9). The restorer line R190 could provide a valuable genetic resource for improving the level of slow rusting resistance in the germplasm (Skowrońska et al., 2020).

Figure 9 Quantitative resistance response of a inbred hybrid rye (*Secale cereale* L.) restorer breeding line R190 at seedling stage under high leaf rust disease pressure.

Genetic engineering of crops constitute a promising technique for the global agricultural production to meet the unprecedented challenges; exponential population growth, detrimental climate changes, sustainable agricultural transition, etc., in the 21st century (Mackelprang and Lemaux, 2020). The clustered regularly interspaced short palindromic repeats (CRISPR) - CRISPR-associated protein (Cas) system constitutes a high-potential gene editing technology, allowing a precise genetic manipulation (*e.g.*, insertion or deactivation of gene) for improvement of disease resistance in crop (Schenke and Cai, 2020; Zhu et al., 2020). This would allow the transformation of multiple *R* genes and QTLs without the need for exhaustive MABC to a recurrent elite parent with zero linkage drag (Tripathi et al., 2020). While the CRISPR-Cas technology is currently subject to the same stringent regulation that is imposed on conventual genetically modified organisms (GMO) in the EU (EC, 2019), recent EC report

infer the need for modernizing the GM term and a differentiated legislation for cisgenic and transgenic gene editing (EC, 2021b). If sanctioned the sequence data of candidate leaf rust *R* genes identified using SMRT AgRenSeq constitutes a valuable genetic resource for engineering of leaf rust resistant hybrids in rye using CRISPR-Cas technology. For the initial 'crude' validation of identified candidate *R* genes, KASP markers should be designed using resistance specific SNV and tested in available biparental mapping populations already phenotyped for leaf rust resistance (Słomińska-Durdasiak et al., 2020). This would allow the potential development of high quality functional markers for practical breeding purposes. For state-of-the-art validation a transformation of candidate *R* genes into a leaf rust susceptible line would be required followed by extensive infection typing using distinct *Prs* SPI to evaluate the NLR genes resistance specificity (Arora et al., 2019; Wang et al., 2019b).

Appendices

Protocol

Recovery, multiplication and infection typing of *Puccinia recondita* f. sp. *secalis*, causing leaf rust in rye (*Secale cereale* L.)

Nikolaj M. Vendelbo, Industrial PhD, Nordic Seed (Dyngby, Denmark) and Institute of Agroecology, Aarhus University (Slagelse, Denmark). E-mail: nive@nordicseed.com

Jihad Orabi, Head of Molecular Breeding, Nordic Seed (Dyngby, Denmark)

Ellen Jørgensen, Laboratory technician, Global Rust Reference Center, Department of Agroecology, Aarhus University (Slagelse, Denmark)

Mogens Hovmøller, Professor, Plant Pathology, Global Rust Reference Center, Department of Agroecology, Aarhus University (Slagelse, Denmark)

Abbreviations:

DAI Days after inoculation
SPI Single pustule isolate
IT Infection type
Prs *Puccinia recondita* f. sp. *secalis*
RH Relative humidity

Overview

1. Collection of rust samples from field crops
2. Recovery and, multiplication of samples from field crops
3. Recovery and, multiplication of stock isolates from -80ºC frozen liquid N_2
4. Procedure for isolation and multiplication of single pustule isolates (SPI)
5. Large-scale inoculation using air brush spray system
6. Large-scale multiplication
7. Leaf rust phenotyping at seedling stage

1. Collection of rust samples from field crops

Step 1: Identify a green leaf with vigorous uredinia pustules and no other foliar diseases

Step 2: Detach leaf at basis using a scissor

Step 3: Tape the stretched leaf onto a clean piece of paper and fold to create a closed envelope

Step 4: Repeat above depending on number of samples required

Step 5: Dry the leaf segments (on paper) in a desiccator at least 24h and place weight on top to prevent leaf curling

2. Recovery and, multiplication of samples from field crops

Modified after Thach., *et al* (2015)

11 Days before inoculation:

Step 1: Sow 12-14 seeds of a susceptible cultivar or cultivar mixture at 1 cm depth in pots (7x7x8 cm) of Pindstrup Substrate peat mix containing slow-release plant nutrients (Pindstrup Mosebrug A/S, Ryomgaard, Denmark)

Note: Rye may be prone to establish poorly in heavy ('fine') soil types and if sown too deep (> 3-4 cm)

Step 2: Water the pot from above and place in a tray with sufficient water from below to cover 1/3 of the pot

Step 3: Place a lid on the tray to ensure 100% RH

Step 4: Cultivate under 16 hours of light at 18ºC using a cultivation lamp (high-pressure sodium or LED) and 8 hours of dark at 12ºC

7-8 Days before inoculation: Remove the lid at emergence of coleoptiles (2-3 cm plant height)

5 Days before inoculation: Add 3.5 mL $4.8x10^{-3}$ M Antergon MH180 per pot, a growth regulator enhancing plant susceptibility to *Puccinia* spp. and increasing spore productivity

Day of inoculation: At half-emergence of 2nd leaf of seedlings

Step 1: Revitalize dried/fresh samples of infected rye, by placing leaf sections on a moist filter paper in a Petri dish, leave for 3-8 hours (dried samples) / up to 24h for fresh samples, at 100% RH until new urediniospores emerges
Note: If leaf segments are curled, place a clean object glass on top

Step 2: Carefully swipe seedling leaves in vertical direction for 10 seconds using presterilized gloves to 'break' the cuticula wax layer

Step 3: Inoculate the seedlings by repeatedly rubbing a revitalized leaf segment across the leaves in a horizontical direction to transfer the uredinio spores

Step 4: Carefully swipe the seedlings in a vertical direction to spread the spores uniformly

Step 5: Spray a fine mist of water across the seedlings from all angles

Step 6: Place the pot in a closed container with surface water to ensure 100% RH, and incubate at ≈ 12.5ºC in dark for 24 hours

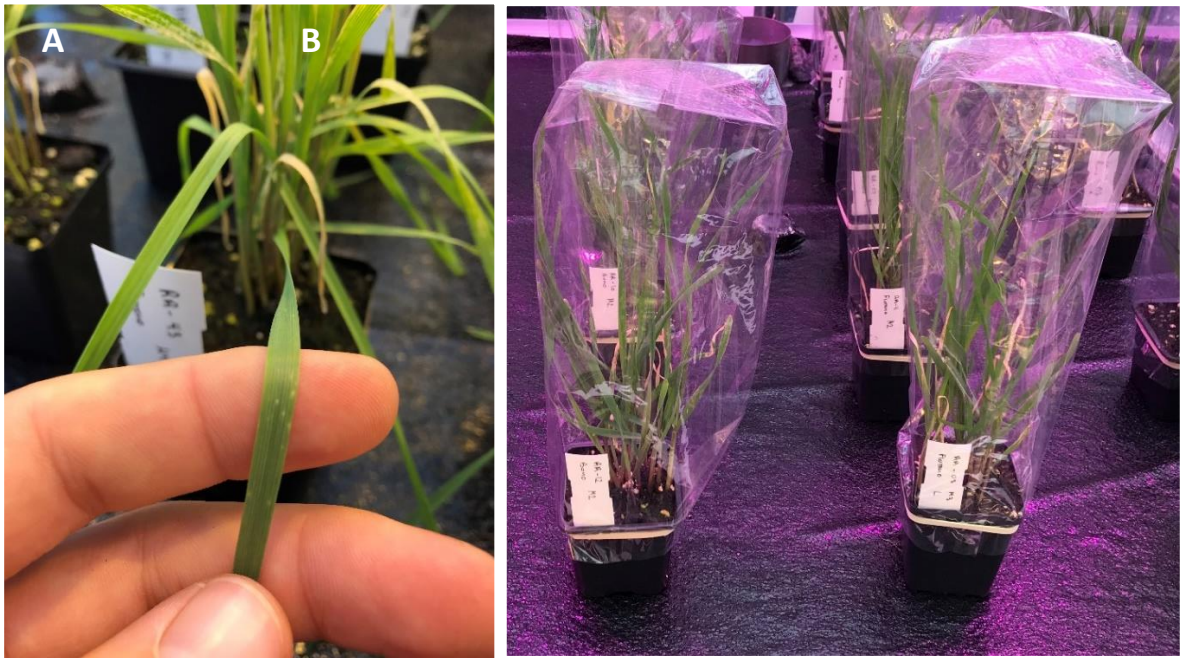
1 DAI: Transfer the multiplication pot to the greenhouse

Note: High humidity (>70% RH) during the latent period positively enhances the infection success

7 DAI: At first emergence of chlorotic spots on inoculated plants, cover each pot with a cellophane bag to prevent cross contamination by spores from neighboring pots (Figure 1A-B)

Note: Cellophane bags may delay/reduce rust spore production on rye. In case large amount of spores are required, it is recommend to multiply spores without cellophane bags, in single isolation cabins per isolate to avoid cross contamination. (See 'Large-scale multiplication')

Figure 1: Multiplication of rye (*Secale cereale* L.) leaf rust, **A**) Emergence of chlorotic spots 7 days after inoculation, **B**) Pots placed in cellophane isolation bags to prevent cross-contamination.



11-14 DAI: 1st spore harvest

Step 1: Lift the multiplication pot slowly using the left hand, while putting a careful pressure to the seedling stalks, making a bent V shape so that loosened spores will collect in the cellophane bag (Figure 2A)

Note: The V-bent is to prevent soil from contaminating the spores when tapping the leaves

Step 2: Pot is laid down horizontally and transported to a flow cabinet

Step 3: Place a piece of weighing paper on the table, pre folded on the end to create a 'bottle neck'

Step 4: Cut the corner of the bag where spores are accumulated, and pour the spores onto the weighing paper by tapping the cellophane bag (Figure 2B)

Step 5: Remove the cut off cellophane bag corner using a pincer

Step 6: Transfer the spores to a pre-labeled tube

Step 7: Place the tube in a desiccator and loosen the lid, let dry for 3 days at 18-20ºC

Note: Temperatures (≥ 25ºC) may dramatically reduce spore vitality

Step 8: Place the tube in a -80ºC freezer for mid-long term storage

Note: Freezing of spores at -80ºC (or -20ºC) may dramatically reduce spore vitality, in cases of, insufficient drying

Step 9: Repeat spore harvest 2-3 times until plant and rust vigor decreases

Figure 2: Multiplication of leaf rust spores on rye (*Secale cereale* L.), **A**) Harvest of spores 14 days after inoculation by tapping leaves, releasing spores to collect in the bottom corners, **B**) Spores poured onto weighing paper before transfer into tubes for mid-long term storage.

3. Recovery and, multiplication of stock isolates from -80ºC frozen or liquid N_2

Modifications to above described protocol 'Multiplication of field samples'

Day 0: At half-emergence of 2nd leaf of seedlings

Step 1: Revitalize spores kept at -80ºC by heat shock at 42ºC for 2 minutes

Note: Spores are stable at this stage for several days when kept at 5ºC under dry conditions

Step 3.1: Pour the spores onto a piece of weighing paper

Step 3.2: Using presterilized gloves fixate the spores onto the index finger

Step 3.3: Inoculate the seedlings by stroking the leaves in a vertical direction for 30 seconds to transfer the uredinia spores and spread it uniformly

4. Procedure for isolation and multiplication of single pustule isolates

As above with the following modifications: Single pustule isolates can either be isolated from field samples at low disease incidence or alternatively by low-density inoculation of seedlings in pots as described above.

Figure 3: Revitalization of rye (*Secale cereale* L.) leaf segment with leaf rust uredinia pustules. In this case, the fungus has expanded by fungal growth from the initial infection site on a seedling leaf, forming a circular ring of new pustules

5. Large scale inoculation using air brush spray system

Modified after Thach., *et al* (2015)

At time of inoculation (half-emergence of 2nd leaf of seedlings)

Step 1.1: Pour spores into an air brush glass container

Note: If using spores from a frozen stock, revitalize as previously described

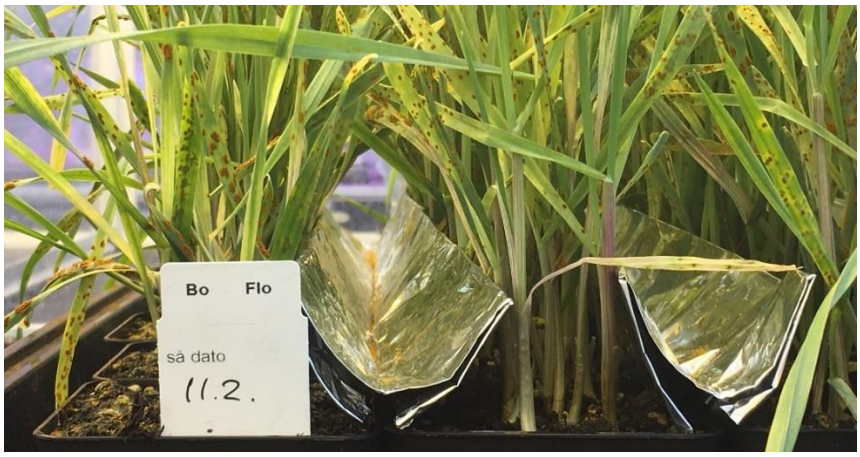
Step 1.2: Solubilize spores in 2 mL 3M™ Novec™ 7100 engineering fluid

Step 3: Inoculate the seedlings by spraying at 30-40 cm distance

Large-scale multiplication can be done using the two following procedures

1. Place support grinds around the pots at 2 days after inoculation by gently ‘forcing’ the leaves through the upper (‘second’) grid window to get leaves hanging in a 45º angle

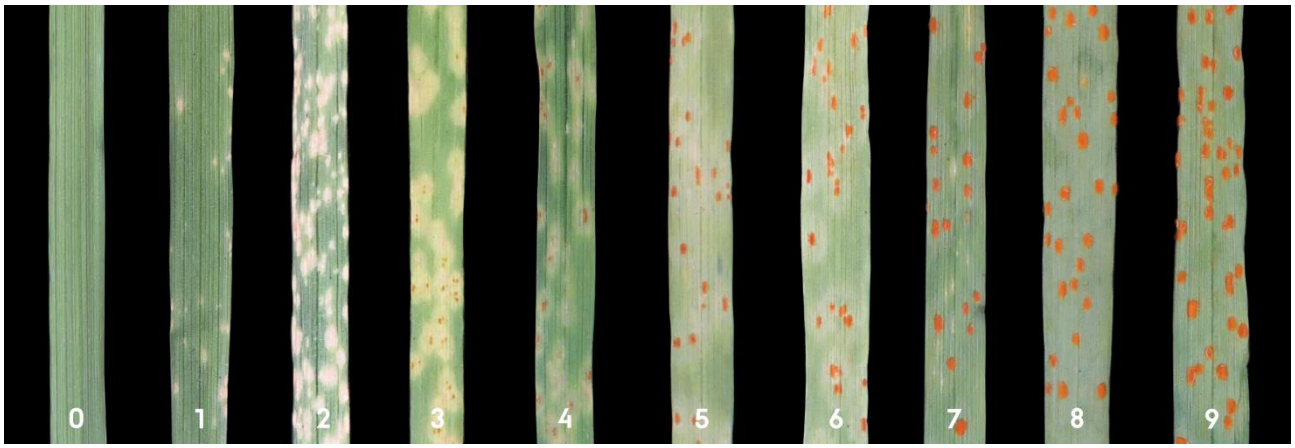
2. Place aluminum foil trays between each row in the tray 2 days after spray inoculation using air brush. At harvest, filter the spores through a sieve.



6. Leaf rust phenotyping at seedling stage

Seedlings are inoculated at the 2nd leaf stage allowing a scoring of infection types (IT) on both 1 and 2. Methodology for scoring of *Prs* SPI IT was adapted from Hovmøller., *et al.* (2017) about race typing of yellow rust in wheat. High-quality resolution picture of the ten IT in rye presented in (Figure 4), is available at https://zenodo.org/record/5478060.

Figure 4: Infection type response (0-9) for leaf rust in rye (*Secale cereale* L.) caused by the fungal pathogen *Puccinia recondita* f. sp. *secalis* after Hovmøller., *et al.* (2017) and McNeal., *et al*. (1971). IT 0-2 are considered to represent 'Resistant', IT 3-4 'Partial resistant', IT 5-6 'Partial susceptible', and IT 7-9 'Susceptible' host plants. In terms of virulence/avirulence, IT 0-6 are considered 'avirulent' and 7-9 'virulent'.



References:

Hovmøller., M. S., Rodriguez-Algaba, J., Thach, T., & Sørensen, C. K. (2017). Race typing of *Puccinia striiformis* on wheat. In *Wheat rust diseases* (pp. 29-40). Humana Press, New York, NY.

McNeal, F. H., Konzak, C. F., Smith, E. P., Tate, W. S., & Russell, T. S. (1971). A uniform system for recording and processing cereal research data (No. REP-10904. CIMMYT.).

Thach, T., Ali, S., Justesen, A. F., Rodriguez-Algaba, J., & Hovmøller, M. S. (2015). Recovery and virulence phenotyping of the historic 'Stubbs collection' of the yellow rust fungus *Puccinia striiformis* from wheat. *Annals of Applied Biology*, *167*(3), 314-326.

References

Acevedo-Garcia, J., Kusch, S., and Panstruga, R. (2014). Magical mystery tour: MLO proteins in plant immunity and beyond. *New Phytologist* 204**,** 273-281.

Acosta-Pech, R., Crossa, J., De Los Campos, G., Teyssèdre, S., Claustres, B., Pérez-Elizalde, S., and Pérez-Rodríguez, P. (2017). Genomic models with genotype× environment interaction for predicting hybrid performance: an application in maize hybrids. *Theoretical and Applied Genetics* 130**,** 1431-1440.

Adams Jr, G., Gottwald, T., and Leach, C. (1986). Environmental factors initiating liberation of conidia of powdery mildews. *Phytopathology* 76**,** 1239-1245.

Albert, I., Zhang, L., Bemm, H., and Nürnberger, T. (2019). Structure-Function Analysis of Immune Receptor At RLP23 with Its Ligand nlp20 and Coreceptors At SOBIR1 and At BAK1. *Molecular Plant-Microbe Interactions* 32**,** 1038-1046.

Alqudah, A.M., Sallam, A., Stephen Baenziger, P., and Borner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate cereals: lessons from barley - A review. *Journal of Advanced Research* 22**,** 119-135.

Andersen, E.J., Nepal, M.P., Purintun, J.M., Nelson, D., Mermigka, G., and Sarris, P.F. (2020). Wheat disease resistance genes and their diversification through integrated domain fusions. *Frontiers in Genetics* 11.

Anderson, J.P., Gleason, C.A., Foley, R.C., Thrall, P.H., Burdon, J.B., and Singh, K.B. (2010). Plants versus pathogens: an evolutionary arms race. *Functional Plant Biology* 37**,** 499-512.

Anderson, T.R., Hawkins, E., and Jones, P.D. (2016). CO2, the greenhouse effect and global warming: from the pioneering work of Arrhenius and Callendar to today's Earth System Models. *Endeavour* 40**,** 178-187.

Andersson, U., Rosen, L., Ostman, E., Strom, K., Wierup, N., Bjorck, I., and Holm, C. (2010). Metabolic effects of whole grain wheat and whole grain rye in the C57BL/6J mouse. *Nutrition* 26**,** 230-239.

Anikster, Y., Bushnell, W., Roelfs, A., Eilam, T., and Manisterski, J. (1997). Puccinia recondita causing leaf rust on cultivated wheats, wild wheats, and rye. *Canadian Journal of Botany* 75**,** 2082-2096.

Anikster, Y., Eilam, T., Bushnell, W., and Kosman, E. (2005). Spore dimensions of Puccinia species of cereal hosts as determined by image analysis. *Mycologia* 97**,** 474-484.

Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C.J., Choulet, F., Distelfeld, A., and Poland, J. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361.

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., Johnson, R., Enk, J., Periyannan, S., Singh, N., Asyraf Md Hatta, M., Athiyannan, N., Cheema, J., Yu, G., Kangara, N., Ghosh, S., Szabo, L.J., Poland, J., Bariana, H., Jones, J.D.G., Bentley, A.R., Ayliffe, M., Olson, E., Xu, S.S., Steffenson, B.J., Lagudah, E., and Wulff, B.B.H. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature Biotechnology* 37**,** 139-143.

Arrones, A., Vilanova, S., Plazas, M., Mangino, G., Pascual, L., Díez, M.J., Prohens, J., and Gramazio, P. (2020). The dawn of the age of multi-parent MAGIC populations in plant breeding: novel powerful next-generation resources for genetic analysis and selection of recombinant elite material. *Biology* 9**,** 229.

Arseniuk, E., and Oleksiak, T. (2003). Rye production and breeding in Poland. *Plant Breeding and Seed Science* 47.

Au (2021). *Powdery mildew of barley* [Online]. Available: https://extensionaus.com.au/FieldCropDiseases Vic/docs/identification-management-of-field-crop-diseases-in-victoria/foliar-diseases-of-barley/powdery-mildew-of-barley/ [Accessed 14/8 2021].

Auinger, H.-J., Schönleben, M., Lehermeier, C., Schmidt, M., Korzun, V., Geiger, H.H., Piepho, H.-P., Gordillo, A., Wilde, P., and Bauer, E. (2016). Model training across multiple breeding cycles significantly improves genomic prediction accuracy in rye (Secale cereale L.). *Theoretical and Applied Genetics* 129**,** 2043-2053.

Aust, H., and Hoyningen-Huene, J.V. (1986). Microclimate in relation to epidemics of powdery mildew. *Annual Review of Phytopathology* 24**,** 491-510.

Baggs, E., Dagdas, G., and Krasileva, K. (2017). NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Current Opinion in Plant Biology* 38**,** 59-67.

Balestre, M., Machado, J.C., Lima, J.L., Souza, J.C., and Nóbrega-Filho, L. (2008). Genetic distance estimates among single cross hybrids and correlation with specific combining ability and yield in corn double cross hybrids. *Genetic and Molecular Research* 7**,** 65-73.

Balint-Kurti, P. (2019). The plant hypersensitive response: concepts, control and consequences. *Molecular Plant Pathology* 20**,** 1163-1178.

Barbey, C.R., Lee, S., Verma, S., Bird, K.A., Yocca, A.E., Edger, P.P., Knapp, S.J., Whitaker, V.M., and Folta, K.M. (2019). Disease resistance genetics and genomics in octoploid strawberry. *G3: Genes, Genomes, Genetics* 9**,** 3315-3332.

Bauer, E., Schmutzer, T., Barilar, I., Mascher, M., Gundlach, H., Martis, M.M., Twardziok, S.O., Hackauf, B., Gordillo, A., Wilde, P., Schmidt, M., Korzun, V., Mayer, K.F., Schmid, K., Schon, C.C., and Scholz, U. (2017). Towards a whole-genome sequence for rye (Secale cereale L.). *The Plant Journal* 89**,** 853-869.

Becker, H., Geiger, H., and Morgenstern, K. (1982). Performance and phenotypic stability of different hybrid types in winter rye 1. *Crop Science* 22**,** 340-344.

Bederska-Łojewska, D., Arczewska-Włosek, A., Świątkiewicz, S., Orczewska-Dudek, S., Schwarz, T., Puchała, M., Krawczyk, J., Boros, D., Fraś, A., and Micek, P. (2019). The effect of different dietary levels of hybrid rye and xylanase addition on the performance and egg quality in laying hens. *British Poultry Science* 60**,** 423-430.

Bederska-Łojewska, D., Świątkiewicz, S., Arczewska-Włosek, A., and Schwarz, T. (2017). Rye non-starch polysaccharides: their impact on poultry intestinal physiology, nutrients digestibility and performance indices – a review. *Annals of Animal Science* 17**,** 351-369.

Behre, K.-E. (1992). The history of rye cultivation in Europe. *Vegetation History and Archaeobotany* 1**,** 141-156.

Bernal-Vasquez, A.-M., Gordillo, A., Schmidt, M., and Piepho, H.-P. (2017). Genomic prediction in early selection stages using multi-year data in a hybrid rye breeding program. *BMC genetics* 18**,** 1-17.

Beug, H.-J. (2011). Vegetation changes during the Slavic period, shown by a high resolution pollen diagram from the Maujahn peat bog near Dannenberg, Hanover Wendland, Germany. *Vegetation History and Archaeobotany* 20**,** 199-206.

Beukert, U., Thorwarth, P., Zhao, Y., Longin, C.F.H., Serfling, A., Ordon, F., and Reif, J.C. (2020). Comparing the potential of marker-assisted selection and genomic prediction for improving rust resistance in hybrid wheat. *Frontiers in Plant Science* 11.

Bi, G., Zhou, Z., Wang, W., Li, L., Rao, S., Wu, Y., Zhang, X., Menke, F.L., Chen, S., and Zhou, J.-M. (2018). Receptor-like cytoplasmic kinases directly link diverse pattern recognition receptors to the activation of mitogen-activated protein kinase cascades in Arabidopsis. *The Plant Cell* 30**,** 1543-1561.

Bolibok-Bragoszewska, H., Targonska, M., Bolibok, L., Kilian, A., and Rakoczy-Trojanowska, M. (2014). Genome-wide characterization of genetic diversity and population structure in secale. *BMC Plant Biology* 14**,** 184.

Bolton, M.D., Kolmer, J.A., and Garvin, D.F. (2008). Wheat leaf rust caused by *Puccinia triticina*. *Molecular Plant Pathology* 9**,** 563-575.

Borgen, A., and Grupe, P. (2012). Gamle danske kornsorter, en rapport o mderes historie og om anvendelsen i økologisk landbrug.

Borrelli, G.M., Mazzucotelli, E., Marone, D., Crosatti, C., Michelotti, V., Vale, G., and Mastrangelo, A.M. (2018). Regulation and evolution of NLR genes: A close interconnection for plant immunity. *International Journal of Molecular Sciences* 19.

Bourras, S., Mcnally, K.E., Ben-David, R., Parlange, F., Roffler, S., Praz, C.R., Oberhaensli, S., Menardo, F., Stirnweis, D., and Frenkel, Z. (2015). Multiple avirulence loci and allele-specific effector recognition control the Pm3 race-specific resistance of wheat to powdery mildew. *The Plant Cell* 27**,** 2991-3012.

Bozkurt, T.O., and Kamoun, S. (2020). The plant–pathogen haustorial interface at a glance. *Journal of Cell Science* 133**,** jcs237958.

Brar, G.S., Pozniak, C.J., Kutcher, H.R., and Hucl, P.J. (2019). Evaluation of Fusarium head blight resistance genes Fhb1, Fhb2, and Fhb5 introgressed into elite Canadian hard red spring wheats: effect on agronomic and end-use quality traits and implications for breeding. *Molecular Breeding* 39**,** 1-12.

Braun, U., Cook, R., Inman, A., and Shin, H. (2001). The taxonomy of the powdery mildew fungi. *The Powdery Mildews: A Comprehensive Treatise***,** 13-55.

Bregaglio, S., Donatelli, M., and Confalonieri, R. (2013). Fungal infections of rice, wheat, and grape in Europe in 2030–2050. *Agronomy for Sustainable Development* 33**,** 767-776.

Brewer, M.T., Cadle-Davidson, L., Cortesi, P., Spanu, P.D., and Milgroom, M.G. (2011). Identification and structure of the mating-type locus and development of PCR-based markers for mating type in powdery mildew fungi. *Fungal Genetics and Biology* 48**,** 704-713.

Brown, J., and Rant, J. (2013). Fitness costs and trade-offs of disease resistance and their consequences for breeding arable crops. *Plant Pathology* 62**,** 83-95.

Brown, J.K., and Hovmøller, M.S. (2002). Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* 297.

Butrón, A., Santiago, R., Cao, A., Samayoa, L., and Malvar, R. (2019). QTLs for resistance to fusarium ear rot in a multiparent advanced generation intercross (MAGIC) maize population. *Plant Disease* 103**,** 897-904.

Calmes, B., N'guyen, G., Dumur, J., Brisach, C.A., Campion, C., Iacomi, B., Pigné, S., Dias, E., Macherel, D., and Guillemette, T. (2015). Glucosinolate-derived isothiocyanates impact mitochondrial function in fungal cells and elicit an oxidative stress response necessary for growth recovery. *Frontiers in Plant Science* 6.

Carver, T., Thomas, B., and Ingerson-Morris, S. (1995). The surface of Erysiphe graminis and the production of extracellular material at the fungus–host interface during germling and colony development. *Canadian Journal of Botany* 73**,** 272-287.

Caubel, J., Launay, M., Ripoche, D., Gouache, D., Buis, S., Huard, F., Huber, L., Brun, F., and Bancal, M.O. (2017). Climate change effects on leaf rust of wheat: Implementing a coupled crop-disease model in a French regional application. *European Journal of Agronomy* 90**,** 53-66.

Chairi, F., Elazab, A., Sanchez-Bragado, R., Araus, J.L., and Serret, M.D. (2016). Heterosis for water status in maize seedlings. *Agricultural Water Management* 164**,** 100-109.

Chang, Z., Chen, Z., Wang, N., Xie, G., Lu, J., Yan, W., Zhou, J., Tang, X., and Deng, X.W. (2016). Construction of a male sterility system for hybrid rice breeding and seed production using a nuclear male sterility gene. *Proceedings of the National Academy of Sciences* 113**,** 14145-14150.

Chen, J., Ullah, C., Reichelt, M., Beran, F., Yang, Z.-L., Gershenzon, J., Hammerbacher, A., and Vassão, D.G. (2020). The phytopathogenic fungus Sclerotinia sclerotiorum detoxifies plant glucosinolate hydrolysis products via an isothiocyanate hydrolase. *Nature Communications* 11**,** 1-12.

Chen, N.W.G., Thareau, V., Ribeiro, T., Magdelenat, G., Ashfield, T., Innes, R.W., Pedrosa-Harand, A., and Geffroy, V. (2018). Common bean subtelomeres are hot spots of recombination and favor resistance gene evolution. *Frontiers in Plant Science* 9.

Chen, X. (2013a). High-temperature adult-plant resistance, key for sustainable control of stripe rust. *American Journal of Plant Sciences* 4**,** 1-20.

Chen, Z.J. (2013b). Genomic and epigenetic insights into the molecular bases of heterosis. *Nature Reviews Genetics* 14**,** 471-482.

Chhetri, M., Bariana, H., Wong, D., Sohail, Y., Hayden, M., and Bansal, U. (2017). Development of robust molecular markers for marker-assisted selection of leaf rust resistance gene Lr23 in common and durum wheat breeding programs. *Molecular Breeding* 37**,** 21.

Chitwood-Brown, J., Vallad, G.E., Lee, T.G., and Hutton, S.F. (2021). Characterization and elimination of linkage-drag associated with Fusarium wilt race 3 resistance genes. *Theoretical and Applied Genetics***,** 1-12.

Cobb, J.N., Biswas, P.S., and Platten, J.D. (2019). Back to the future: revisiting MAS as a tool for modern plant breeding. *Theoretical and Applied Genetics* 132**,** 647-667.

Compant, S., Samad, A., Faist, H., and Sessitsch, A. (2019). A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research* 19**,** 29-37.

Corwin, J.A., and Kliebenstein, D.J. (2017). Quantitative resistance: more than just perception of a pathogen. *The Plant Cell* 29**,** 655-665.

Cowger, C., Mehra, L., Arellano, C., Meyers, E., and Murphy, J.P. (2018). Virulence differences in *Blumeria graminis* f. sp. *tritici* from the Central and Eastern United States. *Phytopathology* 108**,** 402-411.

Cowger, C., Parks, R., and Kosman, E. (2016). Structure and migration in U.S. *Blumeria graminis* f. sp. *tritici* populations. *Phytopathology* 106**,** 295-304.

Crespo-Herrera, L.A., Garkava-Gustavsson, L., and Åhman, I. (2017). A systematic review of rye (Secale cereale L.) as a source of resistance to pathogens and pests in wheat (*Triticum aestivum* L.). *Hereditas* 154**,** 1-9.

Cuevas, J., Crossa, J., Montesinos-López, O.A., Burgueño, J., Pérez-Rodríguez, P., and De Los Campos, G. (2017). Bayesian genomic prediction with genotype× environment interaction kernel models. *G3: Genes, Genomes, Genetics* 7**,** 41-53.

Czeczuga, B., and Kossacka, W. (1973). Nasiona roslin uprawnych i chwastow z osady jacwieskiej w Osowej, pow. Suwalki (III-V wne)(Seeds of cultivated plants and weeds from a AD 3rd–5th c. Yotvingian settlement in Osowa, county Suwałki). *Acta Baltico-Slavica* 8**,** 151-166.

Dab (2021). *myBaits. Hybridization capture for targeted next-generation sequencing. User manual version 5.01 march 2021. Daicel Arbor Biosciences* [Online]. Available: https://arborbiosci.com/wp-content/uploads/2021/03/myBaits_v5.01_Manual.pdf [Accessed 12/9 2021].

De Toledo Thomazella, D.P., Seong, K., Mackelprang, R., Dahlbeck, D., Geng, Y., Gill, U.S., Qi, T., Pham, J., Giuseppe, P., and Lee, C.Y. (2021). Loss of function of a DMR6 ortholog in tomato confers broad-spectrum disease resistance. *Proceedings of the National Academy of Sciences* 118.

De Vallavieille-Pope, C., Giosue, S., Munk, L., Newton, A., Niks, R., Østergård, H., Pons-Kühnemann, J., Rossi, V., and Sache, I. (2000). Assessment of epidemiological parameters and their use in epidemiological and forecasting models of cereal airborne diseases. *Agronomie* 20**,** 715-727.

Delworth, T.L., and Knutson, T.R. (2000). Simulation of early 20th century global warming. *Science* 287**,** 2246-2250.
Devos, K.M. (2005). Updating the 'crop circle'. *Current Opinion in Plant Biology* 8**,** 155-162.
Dmi (2011). *Teknisk rapport 11-01. Danmarks klima 2010 med Tórshavn, Færøerne og Nuuk, Grønland. Danish Meteorological Institute* [Online]. Minestry of Climate and Energy. Available: https://www.dmi.dk/fileadmin/Rapporter/TR/tr11-01.pdf [Accessed 5/9 2021].
Dmi (2021). *Danish Meteorological Institute* [Online]. Available: www.dmi.dk [Accessed 4/9 2021].
Dreiseitl, A. (2018). Great pathotype diversity and reduced virulence complexity in a Central European population of *Blumeria graminis* f. sp. *hordei* in 2015–2017. *European Journal of Plant Pathology* 153**,** 801-811.
Drenth, A., Mctaggart, A.R., and Wingfield, B.D. (2019). Fungal clones win the battle, but recombination wins the war. *IMA Fungus* 10**,** 1-6.
Dst (2021). *Danmarks Statistik* [Online]. Available: http://www.dst.dk [Accessed 20/8 2021].
Dvorak, J., and Fowler, D. (1978). Cold hardiness potential of triticale and tetraploid rye. *Crop Science* 18.
EC (2002). Directive 2002/32/EC of the European parliament and of the council of 7 May 2002 on undesirable substances in animal feed. *The Official Journal of the European Union* L140**,** 10-22.
EC (2009). *Directive 2009/128/EC of the European Parliament and of the council of 21 October 2009. Establishing a framework for community action to achieve the sustainable use of pesticides.* [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02009L0128-20091125&from=EN [Accessed 14/9 2021].
EC (2019). *A scientific perspective on the regulatory status of products derived from gene editing and the implications for the GMO Directive. Statement by the Group of Chief Scientific Advisors* [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/a9100d3c-4930-11e9-a8ed-01aa75ed71a1/language-en [Accessed 14/9 2021].
EC (2020). *Communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - A Farm to Fork Strategy for a fair, healthy and environmentally-friendly food system. Brussels, 20.5.2020 - COM(2020) 381* [Online]. Available: https://ec.europa.eu/food/horizontal-topics/farm-fork-strategy_en [Accessed 9/9 2021].
EC (2021a). *Communication from the commission to the European Parliament, the Council, the European Econimic and Social Committee and the Committee of the Regions. Brussels, 25.3.2021 - COM(2021) 141 final* [Online]. Available: https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/farming/documents/com2021_141_act_organic-action-plan_en.pdf [Accessed 9/9 2021].
EC (2021b). *Study on the status of new genomic techniques under Union law and in light of the Court of Justice ruling in Case C-528/16. Commission staff working document. Brussels, 9.4.2021 - SWD(2021) 92 final* [Online]. Available: https://ec.europa.eu/food/system/files/2021-04/gmo_mod-bio_ngt_eu-study.pdf [Accessed 9/9 2021].
Eriksen, L.B. (2014). Hybrid i Korn - "hot or not"? Optimér dyrkningen af hybridrug. *Plante Kongress***,** 140-141.
EU (2012). *Scientific Opinion on Ergot Alkaloids in Food and Feed. EFSA Panel on Contaminants in the Food Chain (CONTAM). 2012. EFSA 10: 2798* [Online]. Available: http://onlinelibrary.wiley.com/doi/10.2903/j.efsa.2012.2798/pdf [Accessed 21/8 2021].
Eversmeyer, M.G., Kramer, C., and Burleigh, J. (1973). Vertical spore concentrations of three wheat pathogens above a wheat field. *Phytopathology* 63**,** 211-218.
Falke, K.C., Wilde, P., and Miedaner, T. (2009). Rye introgression lines as source of alleles for pollen-fertility restoration in pampa cms. *Plant Breeding* 128**,** 528-531.
FAO (2021). *Food and Agriculture Organization of the United Nations. FAOSTAT statistical database* [Online]. Available: http://www.fao.org/faostat [Accessed 20/8 2021].
Feehan, J.M., Castel, B., Bentham, A.R., and Jones, J.D. (2020). Plant NLRs get by with a little help from their friends. *Current Opinion in Plant Biology* 56**,** 99-108.
Feng, J., Wang, F., Hughes, G.R., Kaminskyj, S., and Wei, Y. (2011). An important role for secreted esterase in disease establishment of the wheat powdery mildew fungus *Blumeria graminis* f. sp. *tritici*. *Canadian Journal of Microbiology* 57**,** 211-216.
Ferrão, L.F.V., Marinho, C.D., Munoz, P.R., and Resende Jr, M.F. (2020). Improvement of predictive ability in maize hybrids by including dominance effects and marker× environment models. *Crop Science* 60**,** 666-677.
FiBL (2021). *The world of organic agriculture, statistics & emerging trends 2021* [Online]. Available: https://www.fibl.org/fileadmin/documents/shop/1150-organic-world-2021.pdf [Accessed 9/9 2021].
Figueroa, M., Dodds, P.N., and Henningsen, E.C. (2020). Evolution of virulence in rust fungi - multiple solutions to one problem. *Current Opinion in Plant Biology* 56**,** 20-27.

Finkel, O.M., Castrillo, G., Paredes, S.H., González, I.S., and Dangl, J.L. (2017). Understanding and exploiting plant beneficial microbes. *Current Opinion in Plant Biology* 38**,** 155-163.
Flint-Garcia, S.A., Buckler, E.S., Tiffin, P., Ersoz, E., and Springer, N.M. (2009). Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS One* 4**,** e7433.
Flor, H.H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology* 9**,** 275-296.
Frederiksen, S., and Petersen, G. (1998). A taxonomic revision of Secale (Triticea, Poaceae). *Nordic Journal of Botany* 18**,** 399-420.
Frederiksen, S., Rasmussen, F., and Seberg, O. (2012). *Dansk flora.* København: Gyldendal A/S.
Friebe, B., Jiang, J., Raupp, W.J., Mcintosh, R.A., and Gill, B.S. (1996). Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91**,** 59-87.
Frolich, W., Aman, P., and Tetens, I. (2013). Whole grain foods and health - a Scandinavian perspective. *Food & Nutrition Research* 57**,** 18503.
Fuong, F., Voylokov, A., and Smirnov, V. (1993). Genetic studies of self-fertility in rye (Secale cereale L.). 2. The search for isozyme marker genes linked to self-incompatibility loci. *Theoretical and Applied Genetics* 87**,** 619-623.
Gaborieau, L., Brown, G.G., and Mireau, H. (2016). The propensity of pentatricopeptide repeat genes to evolve into restorers of cytoplasmic male sterility. *Frontiers in Plant Science* 7.
Gabur, I., Chawla, H.S., Lopisso, D.T., Von Tiedemann, A., Snowdon, R.J., and Obermeier, C. (2020). Gene presence-absence variation associates with quantitative Verticillium longisporum disease resistance in Brassica napus. *Scientific Reports* 10**,** 1-11.
Gałka, M., Miotk-Szpiganowicz, G., Marczewska, M., Barabach, J., Van Der Knaap, W.O., and Lamentowicz, M. (2015). Palaeoenvironmental changes in Central Europe (NE Poland) during the last 6200 years reconstructed from a high-resolution multi-proxy peat archive. *The Holocene* 25**,** 421-434.
Geiger, H. (1972). Wiederherstellung der Pollenfertilität in cytoplasmatisch männlich sterilem Roggen. *Theoretical and Applied Genetics* 42**,** 32-33.
Geiger, H.H., and Miedaner, T. (1999). "Hybrid rye and heterosis," in *Genetics and Exploitation of Heterosis in Crops*. (Madison, WI, USA: American Society of Agronomy and Crop Science Society of America), 439-450.
Geiger, H.H., and Miedaner, T. (2009). "Rye breeding," in *Handbook of Plant Breeding, Cereals,* ed. M.J. Carena. (New York: Springer Publishing), 157-181.
Geiger, H.H., and Schnell, F.W. (1970). Cytoplasmic male sterility in rye (Secale cereale L.). *Crop Science* 10**,** 590-593.
Gilbert, G.S. (2002). Evolutionary ecology of plant diseases in natural ecosystems. *Annual Review of Phytopathology* 40**,** 13-43.
Glawe, D.A. (2008). The powdery mildews: a review of the world's most familiar (yet poorly known) plant pathogens. *Annual Reviews Phytopathology* 46**,** 27-51.
Gold, R., Littlefield, L., and Statler, G. (1979). Ultrastructure of the pycnial and aecial stages of Puccinia recondita. *Canadian Journal of Botany* 57**,** 74-86.
Gold, R.E., and Mendgen, K. (1991). "Rust basidiospore germlings and disease initiation," in *The Fungal Spore and Disease Initiation in Plants and Animals*. Springer), 67-99.
Gong, B.-Q., Guo, J., Zhang, N., Yao, X., Wang, H.-B., and Li, J.-F. (2019). Cross-microbial protection via priming a conserved immune co-receptor through juxtamembrane phosphorylation in plants. *Cell Host & Microbe* 26**,** 810-822. e817.
Goyeau, H., Berder, J., Czerepak, C., Gautier, A., Lanen, C., and Lannou, C. (2012). Low diversity and fast evolution in the population of *Puccinia triticina* causing durum wheat leaf rust in France from 1999 to 2009, as revealed by an adapted differential set. *Plant Pathology* 61**,** 761-772.
Goyeau, H., Park, R., Schaeffer, B., and Lannou, C. (2006). Distribution of pathotypes with regard to host cultivars in French wheat leaf rust populations. *Phytopathology* 96**,** 264-273.
Grabowski, R. (2013). Cereal cultivation in east-central Jutland during the Iron Age, 500 BC–AD 1100. *Danish Journal of Archaeology* 2**,** 164-196.
Green, J., Carver, T.L., and Gurr, S. (2002). The formation and function of infection and feeding structures. *The Powdery Mildews: A Comprehensive Treatise***,** 66-82.
Grikpėdis, M., and Matuzevičiūtė, G.M. (2016). The beginnings of rye (Secale cereale) cultivation in the East Baltics. *Vegetation History and Archaeobotany* 25**,** 601-610.
Gu, Y., Chu, B., Wang, C., Li, L., Zhou, Y., Luo, Y., and Ma, Z. (2020). Spore concentrations of *Blumeria graminis* f. sp. *tritici* in relation to weather factors and disease development in Gansu, China. *Canadian Journal of Plant Pathology* 42**,** 52-61.

Gökgöl, D. (1944). Samen, die bei den Ausgrabungen in Alaca Höyük im Jahre 1936 gefunden worden sind. *Ausgrabungen von Alaca Höyük. Ankara*, 185-186.
Götz, M., and Boyle, C. (1998). Haustorial function during development of cleistothecia in *Blumeria graminis* f. sp. *tritici*. *Plant Disease* 82, 507-511.
Hackauf, B., Bauer, E., Korzun, V., and Miedaner, T. (2017). Fine mapping of the restorer gene Rfp3 from an Iranian primitive rye (Secale cereale L.). *Theoretical and Applied Genetics* 130, 1179-1189.
Hackauf, B., Korzun, V., Wortmann, H., Wilde, P., and Wehling, P. (2012). Development of conserved ortholog set markers linked to the restorer gene Rfp1 in rye. *Molecular Breeding* 30, 1507-1518.
Hackauf, B., and Wehling, P. (2005). Approaching the self-incompatibility locus Z in rye (Secale cereale L.) via comparative genetics. *Theoretical and Applied Genetics* 110, 832-845.
Hansjakob, A., Riederer, M., and Hildebrandt, U. (2012). Appressorium morphogenesis and cell cycle progression are linked in the grass powdery mildew fungus *Blumeria graminis*. *Fungal Biology* 116, 890-901.
Hao, M., Liu, M., Luo, J., Fan, C., Yi, Y., Zhang, L., Yuan, Z., Ning, S., Zheng, Y., and Liu, D. (2018). Introgression of powdery mildew resistance gene Pm56 on rye chromosome arm 6RS into wheat. *Frontiers in Plant Science* 9, 1-8.
Hardzei, S., and Urban, E. (2003). Prospects and problems of hybrid rye breeding in Belarus. *Plant Breeding and Seed Science* 47.
Haseneyer, G., Schmutzer, T., Seidel, M., Zhou, R., Mascher, M., Schon, C.C., Taudien, S., Scholz, U., Stein, N., Mayer, K.F., and Bauer, E. (2011). From RNA-seq to large-scale genotyping - genomics resources for rye (Secale cereale L.). *BMC Plant Biology* 11, 131.
He, H., Du, H., Liu, R., Liu, T., Yang, L., Gong, S., Tang, Z., Du, H., Liu, C., Han, R., Sun, W., Wang, L., and Zhu, S. (2021). Characterization of a new gene for resistance to wheat powdery mildew on chromosome 1RL of wild rye Secale sylvestre. *Theoretical and Applied Genetics* 134, 887-896.
He, H., Ji, J., Li, H., Tong, J., Feng, Y., Wang, X., Han, R., Bie, T., Liu, C., and Zhu, S. (2020). Genetic diversity and evolutionary analyses reveal the powdery mildew resistance gene Pm21 undergoing diversifying selection. *Frontiers in Genetics* 11.
Hedrick, P.W., and Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology & Evolution* 31, 940-952.
Heitman, J. (2015). Evolution of sexual reproduction: a view from the fungal kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biology Reviews* 29, 108-117.
Helfer, S. (2014). Rust fungi and global change. *New Phytologist* 201, 770-780.
Henriksen, P.S. (2003). Rye cultivation in the Danish Iron Age–some new evidence from iron-smelting furnaces. *Vegetation History and Archaeobotany* 12, 177-185.
Hermansen, J.E., Torp, U., and Prahm, L.P. (1978). Studies of transport of live spores of cereal mildew and rust fungi across the North Sea. *Grana* 17, 41-46.
Hickey, L.T., Germán, S.E., Pereyra, S.A., Diaz, J.E., Ziems, L.A., Fowler, R.A., Platz, G.J., Franckowiak, J.D., and Dieters, M.J. (2017). Speed breeding for multiple disease resistance in barley. *Euphytica* 213, 64.
Hillman, G. (1978). On the origins of domestic rye—secale cereale: the finds from aceramic Can Hasan III in Turkey. *Anatolian Studies* 28, 157-174.
Hoecker, N., Keller, B., Piepho, H.-P., and Hochholdinger, F. (2006). Manifestation of heterosis during early maize (Zea mays L.) root development. *Theoretical and Applied Genetics* 112, 421-429.
Hovmøller, M.S., Rodriguez-Algaba, J., Thach, T., and Sorensen, C.K. (2017). "Race typing of *Puccinia striiformis* on wheat," in *Wheat Rust Diseases: Methods and Protocols*. (New York, NY: Humana Press), 29-40.
Hovmøller, M., Justesen, A., and Brown, J. (2002a). Clonality and long-distance migration of *Puccinia striiformis* f. sp. *tritici* in north-west Europe. *Plant Pathology* 51, 24-32.
Hovmøller, M.S., Justesen, A.F., and Brown, J.K.M. (2002b). Clonality and long-distance migration of *Puccinia striiformis* f.sp. tritici in North-West Europe. *Plant Pathology* 51, 24-32.
Hovmøller, M.S., Walter, S., Bayles, R.A., Hubbard, A., Flath, K., Sommerfeldt, N., Leconte, M., Czembor, P., Rodriguez-Algaba, J., Thach, T., Hansen, J.G., Lassen, P., Justesen, A.F., Ali, S., and De Vallavieille-Pope, C. (2016). Replacement of the European wheat yellow rust population by new races from the centre of diversity in the near-Himalayan region. *Plant Pathology* 65, 402-411.
Hu, G., and Rijkenberg, F. (1998a). Scanning electron microscopy of early infection structure formation by Puccinia recondita f. sp. tritici on and in susceptible and resistant wheat lines. *Mycological Research* 102, 391-399.
Hu, G., and Rijkenberg, F. (1998b). Ultrastructural studies of the intercellular hypha and haustorium of Puccinia recondita f. sp. tritici. *Journal of Phytopathology* 146, 39-50.

Huang, S., Sirikhachornkit, A., Faris, J., Su, X., Gill, B.S., Haselkorn, R., and Gornicki, P. (2002). Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. *Plant Molecular Biology* 48**,** 805-820.

Huang, Y., Yin, L., Sallam, A.H., Heinen, S., Li, L., Beaubien, K., Dill-Macky, R., Dong, Y., Steffenson, B.J., and Smith, K.P. (2021). Genetic dissection of a pericentromeric region of barley chromosome 6H associated with Fusarium head blight resistance, grain protein content and agronomic traits. *Theoretical and Applied Genetics***,** 1-19.

Huerta-Espino, J., Singh, R., Crespo-Herrera, L.A., Villasenor-Mir, H.E., Rodriguez-Garcia, M.F., Dreisigacker, S., Barcenas-Santana, D., and Lagudah, E. (2020). Adult plant slow rusting genes confer high levels of resistance to rusts in bread wheat cultivars from Mexico. *Frontiers in Plant Science* 11.

Huerta-Espino, J., Singh, R.P., Germán, S., Mccallum, B.D., Park, R.F., Chen, W.Q., Bhardwaj, S.C., and Goyeau, H. (2011). Global status of wheat leaf rust caused by *Puccinia triticina*. *Euphytica* 179**,** 143-160.

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., Wicker, T., Yahiaoui, N., Mago, R., and Keller, B. (2013). Rye Pm8 and wheat Pm3 are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *The Plant Journal* 76**,** 957-969.

Huynh, B.L., Ehlers, J.D., Huang, B.E., Muñoz-Amatriaín, M., Lonardi, S., Santos, J.R., Ndeve, A., Batieno, B.J., Boukar, O., and Cisse, N. (2018). A multi-parent advanced generation inter-cross (MAGIC) population for genetic analysis and improvement of cowpea (Vigna unguiculata L. Walp.). *The Plant Journal* 93**,** 1129-1142.

Inuma, T., Khodaparast, S.A., and Takamatsu, S. (2007). Multilocus phylogenetic analyses within *Blumeria graminis*, a powdery mildew fungus of cereals. *Molecular Phylogenetics and Evolution* 44**,** 741-751.

Jahns, S. (2000). Late-glacial and Holocene woodland dynamics and land-use history of the Lower Oder valley, north-eastern Germany, based on two, AMS 14 C-dated, pollen profiles. *Vegetation History and Archaeobotany* 9**,** 111-123.

Jahns, S. (2007). Palynological investigations into the Late Pleistocene and Holocene history of vegetation and settlement at the Löddigsee, Mecklenburg, Germany. *Vegetation History and Archaeobotany* 16**,** 157-169.

Jamieson, P.A., Shan, L., and He, P. (2018). Plant cell surface molecular cypher: Receptor-like proteins and their roles in immunity and development. *Plant Science* 274**,** 242-251.

Jankovics, T., Komaromi, J., Fabian, A., Jager, K., Vida, G., and Kiss, L. (2015). New insights into the life cycle of the wheat powdery mildew: direct observation of ascosporic Infection in *Blumeria graminis* f. sp. *tritici*. *Phytopathology* 105**,** 797-804.

Janmohammadi, M., Sabaghnia, N., and Mahfoozi, S. (2018). Frost tolerance and metabolite changes of rye (Secale cereale) during the cold hardening and overwintering. *Acta Physiologiae Plantarum* 40**,** 1-11.

Jonsson, K., Andersson, R., Bach Knudsen, K.E., Hallmans, G., Hanhineva, K., Katina, K., Kolehmainen, M., Kyro, C., Langton, M., Nordlund, E., Laerke, H.N., Olsen, A., Poutanen, K., Tjonneland, A., and Landberg, R. (2018). Rye and health - Where do we stand and where do we go? *Trends in Food Science & Technology* 79**,** 78-87.

Jorgensen, J.H. (1988). Erysiphe graminis, powdery mildew of cereals and grasses. *Genetics of Plant Pathogenic Fungi, Advances in Plant Pathology* 6**,** 137-157.

Junk, J., Kouadio, L., Delfosse, P., and El Jarroudi, M. (2016). Effects of regional climate change on brown rust disease in winter wheat. *Climatic Change* 135**,** 439-451.

Jupe, F., Pritchard, L., Etherington, G.J., Mackenzie, K., Cock, P.J., Wright, F., and Hein, I. (2012). Identification and localization of the NB-LRR gene family within the potato genome. *BMC Genomics* 13**,** 75.

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G.J., Maclean, D., Cock, P.J., Leggett, R.M., Bryan, G.J., Cardle, L., Hein, I., and Jones, J.D. (2013). 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. *The Plant Journal* 76**,** 530-544.

Jørgensen, L.N., Kristjansen, H.S., Kirkegaard, S., and Dahlgaard, A.A. (2013). Bekæmpelse af svampedygdomme i korn. *Anvendelsesorienteret Planteværn 2013***,** 1-49.

Kang, Y., Zhou, M., Merry, A., and Barry, K. (2020). Mechanisms of powdery mildew resistance of wheat–a review of molecular breeding. *Plant Pathology* 69**,** 601-617.

Kast, W.K., and Geiger, H.H. (1982). Studies on the inheritance of mildew resistance in rye. I. Results from inbred lines and F1 crosses.

Kema, G.H., Gohari, A.M., Aouini, L., Gibriel, H.A., Ware, S.B., Van Den Bosch, F., Manning-Smith, R., Alonso-Chavez, V., Helps, J., and M'barek, S.B. (2018). Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance. *Nature Genetics* 50**,** 375-380.

Kim, B., Baier, A., Somers, D., and Gustafson, J. (2002). "Aluminum tolerance in triticale, wheat, and rye," in *Mutations, In Vitro and Molecular Techniques for Environmentally Sustainable Crop Improvement*. Springer), 101-111.
Kim, Y.J., and Zhang, D. (2018). Molecular control of male fertility for crop hybrid breeding. *Trends in Plant Science* 23**,** 53-65.
Kirk, P.M., Canon, P.F., Minter, D.W., and Stalpers, J.A. (2008). *Dictionary of the Fungi.* Wallingford, UK: CABI.
Kobylianskii, V.D. (1971). Production of sterile analogues of winter rye varieties, sterile mantainers and fertile restorers. *Trudy po Prikladnoi Botanike Genetike i Selektsii*.
Kolmer, J. (2013). Leaf rust of wheat: pathogen biology, variation and host resistance. *Forests* 4**,** 70-84.
Kolmer, J. (2015). Collections of *Puccinia triticina* in different provinces of China are highly related for virulence and molecular genotype. *Phytopathology* 105**,** 700-706.
Kolmer, J., Hanzalova, A., Goyeau, H., Bayles, R., and Morgounov, A. (2013). Genetic differentiation of the wheat leaf rust fungus *Puccinia triticina* in Europe. *Plant Pathology* 62**,** 21-31.
Kolmer, J., Long, D., and Hughes, M. (2007). Physiologic specialization of *Puccinia triticina* on wheat in the United States in 2005. *Plant Disease* 91**,** 979-984.
Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9**,** 1-9.
Kosman, E., Chen, X., Dreiseitl, A., Mccallum, B., Lebeda, A., Ben-Yehuda, P., Gultyaeva, E., and Manisterski, J. (2019). Functional variation of plant–pathogen interactions: new concept and methods for virulence data analyses. *Phytopathology* 109**,** 1324-1330.
Kourelis, J., and Van Der Hoorn, R.A. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *The Plant Cell* 30**,** 285-299.
Krattinger, S.G., Lagudah, E.S., Wicker, T., Risk, J.M., Ashton, A.R., Selter, L.L., Matsumoto, T., and Keller, B. (2011). Lr34 multi-pathogen resistance ABC transporter: molecular analysis of homoeologous and orthologous genes in hexaploid wheat and other grass species. *The Plant Journal* 65**,** 392-403.
Kyro, C., Skeie, G., Dragsted, L.O., Christensen, J., Overvad, K., Hallmans, G., Johansson, I., Lund, E., Slimani, N., Johnsen, N.F., Halkjaer, J., Tjonneland, A., and Olsen, A. (2012). Intake of whole grain in Scandinavia: intake, sources and compliance with new national recommendations. *Scandinavian Journal of Public Health* 40**,** 76-84.
Kyro, C., Tjonneland, A., Overvad, K., Olsen, A., and Landberg, R. (2018). Higher whole-grain intake is associated with lower risk of type 2 diabetes among middle-aged men and women: The Danish diet, cancer, and health cohort. *Journal of Nutrition* 148**,** 1434-1444.
Köhler, E., and Lange, E. (1979). A contribution to distinguishing cereal from wild grass pollen grains by LM and SEM. *Grana* 18**,** 133-140.
Labroo, M.R., Studer, A.J., and Rutkoski, J.E. (2021). Heterosis and hybrid crop breeding: A multidisciplinary review. *Frontiers in Genetics* 12.
Ladejobi, O., Elderfield, J., Gardner, K.A., Gaynor, R.C., Hickey, J., Hibberd, J.M., Mackay, I.J., and Bentley, A.R. (2016). Maximizing the potential of multi-parental crop populations. *Applied & Translational Genomics* 11**,** 9-17.
Laflamme, B., Dillon, M.M., Martel, A., Almeida, R.N., Desveaux, D., and Guttman, D.S. (2020). The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. *Science* 367**,** 763-768.
Laidig, F., Piepho, H.-P., Rentel, D., Drobek, T., Meyer, U., and Huesken, A. (2017). Breeding progress, variation, and correlation of grain and quality traits in winter rye hybrid and population varieties and national on-farm progress in Germany over 26 years. *Theoretical and Applied Genetics* 130**,** 981-998.
Lal, N.K., Nagalakshmi, U., Hurlburt, N.K., Flores, R., Bak, A., Sone, P., Ma, X., Song, G., Walley, J., and Shan, L. (2018). The receptor-like cytoplasmic kinase BIK1 localizes to the nucleus and regulates defense hormone expression during plant innate immunity. *Cell Host & Microbe* 23**,** 485-497. e485.
Lanver, D., Tollot, M., Schweizer, G., Presti, L.L., Reissmann, S., Ma, L.-S., Schuster, M., Tanaka, S., Liang, L., and Ludwig, N. (2017). Ustilago maydis effectors and their impact on virulence. *Nature Reviews Microbiology* 15**,** 409-421.
Lapinski, M. (1972). Cytoplasmic-genic type of male sterility in Secale montanum Guss. *Wheat Information Service* 35.
Larièpe, A., Moreau, L., Laborde, J., Bauland, C., Mezmouk, S., Décousset, L., Mary-Huard, T., Fiévet, J., Gallais, A., and Dubreuil, P. (2017). General and specific combining abilities in a maize (Zea mays L.) test-cross hybrid panel: relative importance of population structure and genetic divergence between parents. *Theoretical and Applied Genetics* 130**,** 403-417.

Launay, M., Zurfluh, O., Huard, F., Buis, S., Bourgeois, G., Caubel, J., Huber, L., and Bancal, M.-O. (2020). Robustness of crop disease response to climate change signal under modeling uncertainties. *Agricultural Systems* 178**,** 102733.

LBST (2020a). *Statistik over økologiske jordbrugsbedriger 2020, autorisation og produktion. Miljø- og Fødevareministeriet, Landbrugsstyrrelsen. Marts 2021* [Online]. Available: https://lbst.dk/fileadmin/user_upload/NaturErhverv/Filer/Tvaergaaende/Oekologi/Statistik/Statistik_over_oekologiske_jordbrugsbedrifter_2020_.pdf [Accessed 8/9 2021].

LBST (2020b). *Vejledning om økologisk jordbrugsproduktion. Miljø- og Fødevareministeriet, Landbrugsstyrrelsen. Februar 2020* [Online]. Available: https://lbst.dk/fileadmin/user_upload/NaturErhverv/Filer/Indsatsomraader/Oekologi/Jordbrugsbedrifter/Vejledning_til_oekologisk_jordbrugsproduktion/OEkologivejledning_februar2020.pdf [Accessed 10/9 2021].

Le May, C., Montarry, J., Morris, C.E., Frenkel, O., and Ravigné, V. (2020). Plant pathogen life-history traits and adaptation to environmental constraints. *Frontiers in Plant Science* 10.

Lee, H.-A., Lee, H.-Y., Seo, E., Lee, J., Kim, S.-B., Oh, S., Choi, E., Choi, E., Lee, S.E., and Choi, D. (2017). Current understandings of plant nonhost resistance. *Molecular Plant-Microbe Interactions* 30**,** 5-15.

Lee, S., Whitaker, V.M., and Hutton, S.F. (2016). Mini review: potential applications of non-host resistance for crop improvement. *Frontiers in Plant Science* 7.

Li, G., Wang, L., Yang, J., He, H., Jin, H., Li, X., Ren, T., Ren, Z., Li, F., Han, X., Zhao, X., Dong, L., Li, Y., Song, Z., Yan, Z., Zheng, N., Shi, C., Wang, Z., Yang, S., Xiong, Z., Zhang, M., Sun, G., Zheng, X., Gou, M., Ji, C., Du, J., Zheng, H., Dolezel, J., Deng, X.W., Stein, N., Yang, Q., Zhang, K., and Wang, D. (2021). A high-quality genome assembly highlights rye genomic characteristics and agronomically important genes. *Nature Genetics* 53**,** 574-584.

Li, M., Dong, L., Li, B., Wang, Z., Xie, J., Qiu, D., Li, Y., Shi, W., Yang, L., and Wu, Q. (2020). A CNL protein in wild emmer wheat confers powdery mildew resistance. *New Phytologist* 228**,** 1027-1037.

Li, Y.-X., Chen, L., Li, C., Bradbury, P.J., Shi, Y.-S., Song, Y., Zhang, D., Zhang, Z., Buckler, E.S., and Li, Y. (2018). Increased experimental conditions and marker densities identified more genetic loci associated with southern and northern leaf blight resistance in maize. *Scientific Reports* 8**,** 1-12.

Li, Z., Lan, C., He, Z., Singh, R.P., Rosewarne, G.M., Chen, X., and Xia, X. (2014). Overview and application of QTL for adult plant resistance to leaf rust and powdery mildew in wheat. *Crop Science* 54**,** 1907-1925.

Limpert, E., Godet, F., and Müller, K. (1999). Dispersal of cereal mildews across europe. *Agricultural and Forest Meterology* 97**,** 293-308.

Lindner, S., Keller, B., Singh, S.P., Hasenkamp, Z., Jung, E., Müller, M.C., Bourras, S., and Keller, B. (2020). Single residues in the LRR domain of the wheat PM3A immune receptor can control the strength and the spectrum of the immune response. *The Plant Journal* 104**,** 200-214.

Liu, H.J., and Yan, J. (2019). Crop genome-wide association study: a harvest of biological relevance. *The Plant Journal* 97**,** 8-18.

Liu, M., Szabo, L.J., Hambleton, S., Anikster, Y., and Kolmer, J.A. (2013). Molecular phylogenetic relationship of the brown leaf rust fungi on wheat, rye, and other grasses. *Plant disease* 97**,** 1408-1417.

Liu, R., Lu, J., Zhou, M., Zheng, S., Liu, Z., Zhang, C., Du, M., Wang, M., Li, Y., and Wu, Y. (2020). Developing stripe rust resistant wheat (*Triticum aestivum* L.) lines with gene pyramiding strategy and marker-assisted selection. *Genetic Resources and Crop Evolution* 67**,** 381-391.

Lu, P., Guo, L., Wang, Z., Li, B., Li, J., Li, Y., Qiu, D., Shi, W., Yang, L., and Wang, N. (2020). A rare gain of function mutation in a wheat tandem kinase confers resistance to powdery mildew. *Nature Communications* 11**,** 1-11.

Lu, Y., and Tsuda, K. (2021). Intimate association of PRR-and NLR-mediated signaling in plant immunity. *Molecular Plant-Microbe Interactions* 34**,** 3-14.

Lundqvist, A. (1956). Self-incompatability in rye. I. Genetic control in the diploid. *Hereditas* 42**,** 293-348.

Ma, P., Han, G., Zheng, Q., Liu, S., Han, F., Wang, J., Luo, Q., and An, D. (2020). Development of novel wheat-rye chromosome 4R translocations and assignment of their powdery mildew resistance. *Plant Disease* 104**,** 260-268.

Ma, P., Xu, H., Xu, Y., Song, L., Liang, S., Sheng, Y., Han, G., Zhang, X., and An, D. (2018). Characterization of a powdery mildew resistance gene in wheat breeding line 10V-2 and its application in marker-assisted selection. *Plant Disease* 102**,** 925-931.

Mackelprang, R., and Lemaux, P.G. (2020). Genetic engineering and editing of plants: an analysis of new and persisting questions. *Annual Review of Plant Biology* 71**,** 659-687.

Madej, L. (1975). Research on male sterility in rye. *Hodowla Rosl Aklim Nasienn*.

Mago, R., Miah, H., Lawrence, G.J., Wellings, C.R., Spielmeyer, W., Bariana, H.S., Mcintosh, R.A., Pryor, A.J., and Ellis, J.G. (2005). High-resolution mapping and mutation analysis separate the rust resistance genes Sr31, Lr26 and Yr9 on the short arm of rye chromosome 1. *Theoretical and Applied Genetics* 112**,** 41-50.

Mainka, S., Dänicke, S., Böhme, H., Ueberschär, K.H., and Liebert, F. (2007). On the composition of ergot and the effects of feeding two different ergot sources on piglets. *Animal Feed Science and Technology* 139**,** 52-68.

Maraci, O., Ozkan, H., and Bilgin, R. (2018). Phylogeny and genetic structure in the genus Secale. *PLoS One* 13**,** 1-21.

Marais, G., and Charlesworth, B. (2003). Genome evolution: recombination speeds up adaptive evolution. *Current Biology* 13**,** R68-R70.

Marcec, M.J., Gilroy, S., Poovaiah, B., and Tanaka, K. (2019). Mutual interplay of Ca2+ and ROS signaling in plant immune response. *Plant Science* 283**,** 343-354.

Martis, M.M., Zhou, R., Haseneyer, G., Schmutzer, T., Vrana, J., Kubalakova, M., Konig, S., Kugler, K.G., Scholz, U., Hackauf, B., Korzun, V., Schon, C.C., Dolezel, J., Bauer, E., Mayer, K.F., and Stein, N. (2013). Reticulate evolution of the rye genome. *Plant Cell* 25**,** 3685-3698.

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T., Radchuk, V., Dockter, C., Hedley, P.E., and Russell, J. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544**,** 427-433.

Matzen, N., Heick, T.M., and Jørgensen, L.N. (2019). Control of powdery mildew (*Blumeria graminis* spp.) in cereals by Serenade® ASO (Bacillus amyloliquefaciens (former subtilis) strain QST 713). *Biological Control* 139**,** 104067.

Mcelroy, J.S. (2014). Vavilovian Mimicry: Nikolai Vavilov and his little-known impact on weed Science. *Weed Science* 62**,** 207-216.

Mcintosh, R.A., Frieb, B., Jiang, J., and Gill, B.S. (1995). Cytogenetical studies in wheat XVI. Chromosomal location of a new gene for resistance to leaf rust in a Japanese wheat-rye translocation line. *Euphytica* 82**,** 141-147.

Mcnew, G.L. (1960). "The nature, origin, and evolution of parasitism," in *Plant Pathology: An Advanced Treatise,* eds. J.G. Horsfall & A.E. Diamond. (New York: Academic Press), 19-69.

Meena, R.K., Reddy, K.S., Gautam, R., Maddela, S., Reddy, A.R., and Gudipalli, P. (2021). Improved photosynthetic characteristics correlated with enhanced biomass in a heterotic F 1 hybrid of maize (Zea mays L.). *Photosynthesis Research* 147**,** 253-267.

Melonek, J., Duarte, J., Martin, J., Beuf, L., Murigneux, A., Varenne, P., Comadran, J., Specel, S., Levadoux, S., Bernath-Levin, K., Torney, F., Pichon, J.P., Perez, P., and Small, I. (2021). The genetic basis of cytoplasmic male sterility and fertility restoration in wheat. *Nature Communications* 12**,** 1036.

Melz, G., and Adolf, K. (1991). Genetic analysis of rye (Secale cereale L.) genetics of male sterility of the G-type. *Theoretical and Applied Genetics* 82**,** 761-764.

Melz, G., Kaczmarek, J., and Szigat, G. (1990). Genetical analysis of rye (Secale cereale L.). Location of self-fertility genes in different inbred lines. *Genetica Polonica* 31**,** 1-7.

Melz, G., Melz, G., and Hartman, F. (2003). Genetics of a male-sterile rye of 'G-type' with results of the first F1-hybrids. *Plant Breeding and Seed Science* 47**,** 47-55.

Melz, G., Melz, G., and Winkel, A. (1987). Genetical analysis of rye (Secale cereale L.) III. Self-fertility of the rye mutant vd - inheritance and gene location. *Genetica Polonica* 28.

Menardo, F., Wicker, T., and Keller, B. (2017). Reconstructing the evolutionary history of powdery mildew lineages (*Blumeria graminis*) at different evolutionary time scales with NGS data. *Genome Biology and Evolution* 9**,** 446-456.

Miedaner, T., and Geiger, H.H. (2015). Biology, genetics, and management of ergot (Claviceps spp.) in rye, sorghum, and pearl millet. *Toxins* 7**,** 659-678.

Miedaner, T., Gey, A.-K.M., Sperling, U., and Geiger, H.H. (2002). Quantitative-genetic analysis of leaf-rust resistance in seedling and adult-plant stages of inbred lines and their testcrosses in winter rye. *Plant Breeding* 121**,** 475-479.

Miedaner, T., Glass, C., Dreyer, F., Wilde, P., Wortmann, H., and Geiger, H.H. (2000). Mapping of genes for male-fertility restoration in 'Pampa' CMS winter rye (Secale cereale L.). *Theoretical and Applied Genetics* 101**,** 1226-1233.

Miedaner, T., Haffke, S., Siekmann, D., Fromme, F.J., Roux, S.R., and Hackauf, B. (2018). Dynamic quantitative trait loci (QTL) for plant height predict biomass yield in hybrid rye (Secale cereale L.). *Biomass and Bioenergy* 115**,** 10-18.

Miedaner, T., Herter, C.P., Goßlau, H., Wilde, P., Hackauf, B., and Pillen, K. (2017). Correlated effects of exotic pollen-fertility restorer genes on agronomic and quality traits of hybrid rye. *Plant Breeding* 136**,** 224-229.

Miedaner, T., Klocke, B., Flath, K., Geiger, H.H., and Weber, W.E. (2011). Diversity, spatial variation, and temporal dynamics of virulences in the German leaf rust (*Puccinia recondita* f. sp. *secalis*) population in winter rye. *European Journal of Plant Pathology* 132**,** 23-35.

Miedaner, T., Kodisch, A., Raditschnig, A., and Eifler, J. (2021). Ergot alkaloid contents in hybrid rye are reduced by breeding. *Agriculture* 11.

Miedaner, T., and Korzun, V. (2012). Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* 102**,** 560-566.

Miedaner, T., and Laidig, F. (2019). "Hybrid Breeding in Rye (Secale cereale L.)," in *Advances in Plant Breeding Strategies: Cereals*. Springer, Cham), 343-372.

Miedaner, T., Mirdita, V., Rodemann, B., Drobeck, T., and Rentel, D. (2010). Genetic variation of winter rye cultivars for their ergot (Claviceps purpurea) reaction tested in a field design with minimized interplot interference. *Plant Breeding* 129**,** 58-62.

Miedaner, T., Schmidt, H., and Geiger, H. (1993). Components of variation for quantitative adult-plant resistance to powdery mildew in winter rye. *Phytopathology* 83**,** 1071-1075.

Miedaner, T., and Sperling, U. (1995). Effect of leaf rust and yield components of winter rye hybrids and assessment of quantitative resistance. *Journal of Phytopathology* 143**,** 725-730.

Mielke, H. (1993). Investigations on the control of ergot. *Nachrichtenblatt des Deutschen Pflanzenschutzdienstes* 45**,** 97-102.

Mirdita, V., Dhillon, B., Geiger, H., and Miedaner, T. (2008). Genetic variation for resistance to ergot (Claviceps purpurea [Fr.] Tul.) among full-sib families of five populations of winter rye (Secale cereale L.). *Theoretical and Applied Genetics* 118**,** 85-90.

Mirdita, V., and Miedaner, T. (2009). Resistance to ergot in self-incompatible germplasm resources of winter rye. *Journal of Phytopathology* 157**,** 350-355.

Mohler, V., and Stadlmeier, M. (2019). Dynamic QTL for adult plant resistance to powdery mildew in common wheat (*Triticum aestivum* L.). *Journal of Applied Genetics* 60**,** 291-300.

Monino-Lopez, D., Nijenhuis, M., Kodde, L., Kamoun, S., Salehian, H., Schentsnyi, K., Stam, R., Lokossou, A., Abd-El-Haliem, A., and Visser, R.G. (2021). 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. *The Plant Journal*.

Moose, S.P., and Mumm, R.H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology* 147**,** 969-977.

Morgounov, A., Akin, B., Demir, L., Keser, M., Kokhmetova, A., Martynov, S., Orhan, Ş., Özdemir, F., Özseven, İ., and Sapakhova, Z. (2015). Yield gain due to fungicide application in varieties of winter wheat (*Triticum aestivum*) resistant and susceptible to leaf rust. *Crop and Pasture Science* 66**,** 649-659.

Mourad, A.M., Sallam, A., Belamkar, V., Wegulo, S., Bowden, R., Jin, Y., Mahdy, E., Bakheit, B., El-Wafaa, A.A., and Poland, J. (2018). Genome-wide association study for identification and validation of novel SNP markers for Sr6 stem rust resistance gene in bread wheat. *Frontiers in Plant Science* 9.

Musa, G.L.C., Dyck, P.L., and Samborski, D.J. (1984). The inheritance of resistance in rye to puccinia recondita f.sp. secalis and f.sp. tritici. *Canadian Journal of Plant Science* 64**,** 511-519.

MVFM (2017). *Danish national actionplan on pesticides 2017-2021; Facts, caution and consideration* [Online]. Ministry of Environment and Food of Denmark. Available: https://ec.europa.eu/food/system/files/2019-03/pesticides_sup_nap_dan-rev_en.pdf [Accessed 14/9 2021].

MVFM (2021). *Miljø- og Fødevareministeriet: De officielle kostråd (the Danish food based dietary guidelines)* [Online]. Available: https://altomkost.dk/deofficielleanbefalingertilensundlivsstil/de-officielle-kostraad/ [Accessed 9/9 2021].

Märkle, T. (2005). Nutrition, aspects of land use and environment in medieval times in southern Germany: plant macro-remain analysis from latrines (late 11th–13th century ad.) at the town of Überlingen, Lake Constance. *Vegetation History and Archaeobotany* 14**,** 427-441.

Narang, D., Kaur, S., Steuernagel, B., Ghosh, S., Bansal, U., Li, J., Zhang, P., Bhardwaj, S., Uauy, C., and Wulff, B.B. (2020). Discovery and characterisation of a new leaf rust resistance gene introgressed in wheat from wild wheat Aegilops peregrina. *Scientific reports* 10**,** 1-9.

Newman, T.E., Lee, J., Williams, S.J., Choi, S., Halane, M.K., Zhou, J., Solomon, P., Kobe, B., Jones, J.D., and Segonzac, C. (2019). Autoimmunity and effector recognition in Arabidopsis thaliana can be uncoupled by mutations in the RRS1-R immune receptor. *New Phytologist* 222**,** 954-965.

Niedziela, A., Brukwinski, W., and Bednarek, P.T. (2021). Genetic mapping of pollen fertility restoration QTLs in rye (Secale cereale L.) with CMS Pampa. *Journal of Applied Genetics* 62**,** 1-14.

Nielsen, G.C. (2021). *SEGES Landbrugsinfo. Planter, vejledning til bekæmpelse af svampesygdomme i korn* [Online]. Available: file:///C:/Users/NDSNIVE/Downloads/pm_21_7879_Vejledning_til_bekaempelse_af_svampesygdomme_i_korn%20(2).pdf [Accessed 9/9 2021].

Nielsen, K.A., Nicholson, R.L., Carver, T.L.W., Kunoh, H., and Oliver, R.P. (2000). First touch: an immediate response to surface recognition in conidia of *Blumeria graminis*. *Physiological and Molecular Plant Pathology* 56**,** 63-70.

Nieri, D., Di Donato, A., and Ercolano, M.R. (2017). Analysis of tomato meiotic recombination profile reveals preferential chromosome positions for NB-LRR genes. *Euphytica* 213.

Novakazi, F., Krusell, L., Jensen, J.D., Orabi, J., Jahoor, A., Bengtsson, T., and Consortium, P.B. (2020). You had me at "MAGIC"!: four barley MAGIC populations reveal novel resistance QTL for powdery mildew. *Genes* 11**,** 1512.

Nsabiyera, V., Bariana, H.S., Qureshi, N., Wong, D., Hayden, M.J., and Bansal, U.K. (2018). Characterisation and mapping of adult plant stripe rust resistance in wheat accession Aus27284. *Theoretical and Applied Genetics* 131**,** 1459-1467.

Olesen, M. (2020). *Fremtidens klima. Danish Meteorological Institute* [Online]. Available: https://www.klimatilpasning.dk/viden-om/fremtidens-klima/ [Accessed 5/9 2021].

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

Ounnas, F., Prive, F., Salen, P., Gaci, N., Tottey, W., Calani, L., Bresciani, L., Lopez-Gutierrez, N., Hazane-Puch, F., Laporte, F., Brugere, J.F., Del Rio, D., Demeilliers, C., and De Lorgeril, M. (2016). Whole rye consumption improves blood and liver n-3 fatty acid profile and gut microbiota composition in rats. *PLoS One* 11**,** 1-18.

Park, R.F., Golegaonkar, P.G., Derevnina, L., Sandhu, K.S., Karaoglu, H., Elmansour, H.M., Dracatos, P.M., and Singh, D. (2015). Leaf rust of cultivated barley: pathology and control. *Annual Review of Phytopathology* 53**,** 565-589.

Pedersen, Å.P. (1996). *Oversigt over landsforsøgene. Forsøg og undersøgelser i de landøkonomiske foreninger.* Århus, Denmark: Scanprint A/S.

Pedras, M.S.C., and Abdoli, A. (2017). Pathogen inactivation of cruciferous phytoalexins: Detoxification reactions, enzymes and inhibitors. *RSC Advances* 7**,** 23633-23646.

Pel, M.J., and Pieterse, C.M. (2013). Microbial recognition and evasion of host immunity. *Journal of Experimental Botany* 64**,** 1237-1248.

Peng, Y., Van Wersch, R., and Zhang, Y. (2018). Convergent and divergent signaling in PAMP-triggered immunity and effector-triggered immunity. *Molecular Plant-Microbe Interactions* 31**,** 403-409.

Pereira, D., Mcdonald, B.A., and Croll, D. (2020). The genetic architecture of emerging fungicide resistance in populations of a global wheat pathogen. *Genome Biology and Evolution* 12**,** 2231-2244.

Persson, K., and Von Bothmer, R. (2002). Genetic diversity amongst landraces of rye (Secale cereale L.) from northern Europe. *Hereditas* 136**,** 29-38.

Piasecka, A., Jedrzejczak-Rey, N., and Bednarek, P. (2015). Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. *New Phytologist* 206**,** 948-964.

Pilet-Nayel, M.L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.C., Fournet, S., Durel, C.E., and Delourme, R. (2017). Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Frontiers in Plant Science* 8.

Platten, J.D., Cobb, J.N., and Zantua, R.E. (2019). Criteria for evaluating molecular markers: comprehensive quality metrics to improve marker-assisted selection. *PLoS One* 14**,** e0210529.

Qi, D., Deyoung, B.J., and Innes, R.W. (2012). Structure-function analysis of the coiled-coil and leucine-rich repeat domains of the RPS5 disease resistance protein. *Plant Physiology* 158**,** 1819-1832.

Rabanus-Wallace, M.T., Hackauf, B., Mascher, M., Lux, T., Wicker, T., Gundlach, H., Báez, M., Houben, A., Mayer, K.F.X., Guo, L., Poland, J., Pozniak, C.J., Walkowiak, S., Melonek, J., Praz, C., Schreiber, M., Budak, H., Heuberger, M., Steuernagel, B., Wulff, B., Börner, A., Byrns, B., Čížková, J., Fowler, D.B., Fritz, A., Himmelbach, A., Kaithakottil, G., Keilwagen, J., Keller, B., Konkin, D., Larsen, J., Li, Q., Myśków, B., Padmarasu, S., Rawat, N., Sesiz, U., Sezgi, B., Sharpe, A., Šimková, H., Small, I., Swarbreck, D., Toegelová, H., Tsvetkova, N., Voylokov, A.V., Vrána, J., Bauer, E., Bolibok-Bragoszewska, H., Doležel, J., Hall, A., Jia, J., Korzun, V., Laroche, A., Ma, X.-F., Ordon, F., Özkan, H., Rakoczy-Trojanowska, M., Scholz, U., Schulman, A.H., Siekmann, D., Stojałowski, S., Tiwari, V., Spannagl, M., and Stein, N. (2021). Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. *Nature Genetics* 53**,** 564-573.

Racca, P., Kakau, J., Kleinhenz, B., and Kuhn, C. (2015). Impact of climate change on the phenological development of winter wheat, sugar beet and winter oilseed rape in Lower Saxony, Germany. *Journal of Plant Diseases and Protection* 122**,** 16-27.

Rajczak, J., and Schär, C. (2017). Projections of future precipitation extremes over Europe: A multimodel assessment of climate simulations. *Journal of Geophysical Research: Atmospheres* 122**,** 10,773-710,800.

Reichel, A. (1981). Untersuchungen zur horizontalen resistenz von roggen und weizen gegenüber braunrost. PhD Thesis. University of Halle, Germany.

Reif, J.C., Melchinger, A.E., Xia, X.C., Warburton, M.L., Holsington, D.A., Vasal, S.K., and M., F. (2003). Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Science* 43**,** 1275-1282.

Ren, T., Sun, Z., Ren, Z., Tan, F., Luo, P., Tang, Z., Fu, S., and Li, Z. (2020). Molecular and cytogenetic characterization of a wheat-rye 7BS.7RL translocation line with resistance to stripe rust, powdery mildew, and fusarium head blight. *Phytopathology* 110**,** 1713-1720.

Ren, Y., Hou, W., Lan, C., Basnet, B.R., Singh, R.P., Zhu, W., Cheng, X., Cui, D., and Chen, F. (2017). QTL analysis and nested association mapping for adult plant resistance to powdery mildew in two bread wheat populations. *Frontiers in Plant Science* 8.

Riache, M., Revilla, P., Maafi, O., Malvar, R.A., and Djemel, A. (2021). Combining Ability and Heterosis of Algerian Saharan Maize Populations (Zea mays L.) for Tolerance to No-Nitrogen Fertilization and Drought. *Agronomy* 11**,** 492.

Roelfs, A. (1989). Epidemiology of the cereal rusts in North America. *Canadian Journal of Plant Pathology* 11**,** 86-90.

Roux, S.R., Hackauf, B., Linz, A., Ruge, B., Klocke, B., and Wehling, P. (2004). Leaf-rust resistance in rye (Secale cereale L.). 2. Genetic analysis and mapping of resistance genes Pr3, Pr4, and Pr5. *Theoretical and Applied Genetics* 110**,** 192-201.

Roux, S.R., Ruge, B., Linz, A., and Wehling, P. (2000). Leaf rust resistance in rye - evaluaton, genetic analysis and molecular mapping. *Acta Phytopathologica et Entomologica Hungarica* 35**,** 65-73.

Roux, S.R., and Wehling, P. (2010). Nature of mixed infection type 2(5) observed in rye (secale cereale L.) plants carrying the pr1 leaf-rust resistance gene. *Journal Für Kulturpflanzen* 62**,** 29-34.

Rubek, F. (2019). *Årssammendrag 2019. Danish Meteorological Institute* [Online]. Available: https://www.dmi.dk/fileadmin/user_upload/Afrapportering/Aarssammendrag/Sammendrag_2019.pdf [Accessed 5/9 2021].

Rubek, F., Scharling, M., and Cappelen, J. (2020). *Danmarks klima 2019 - with English summary. DMI rapport 20-01. Danish Meteorological Institute* [Online]. Available: https://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2020/DMIRap20-01.pdf [Accessed 10/9 2021].

Rösch, M. (2000). Long-term human impact as registered in an upland pollen profile from the southern Black Forest, south-western Germany. *Vegetation History and Archaeobotany* 9**,** 205-218.

Raaymakers, T.M., and Van Den Ackerveken, G. (2016). Extracellular recognition of oomycetes during biotrophic infection of plants. *Frontiers in Plant Science* 7.

Saintenac, C., Lee, W.-S., Cambon, F., Rudd, J.J., King, R.C., Marande, W., Powers, S.J., Bergès, H., Phillips, A.L., and Uauy, C. (2018). Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to fungal pathogen Zymoseptoria tritici. *Nature Genetics* 50**,** 368-374.

Sakuma, S., Salomon, B., and Komatsuda, T. (2011). The domestication syndrome genes responsible for the major changes in plant form in the Triticeae crops. *Plant and Cell Physiology* 52**,** 738-749.

Salamini, F., Ozkan, H., Brandolini, A., Schafer-Pregl, R., and Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. *Nature Reviews Genetics* 3**,** 429-441.

Saleh, G., Abdullah, D., and Anuar, A. (2002). Performance, heterosis and heritability in selected tropical maize single, double and three-way cross hybrids. *The Journal of Agricultural Science* 138**,** 21-28.

Santos, E., Pinto-Carnide, O., Figueiras, A., Benito, C., and Matos, M. (2019). Biochemical, physiological and genetic analysis of aluminum tolerance of different rye species. *Environmental and Experimental Botany* 162**,** 87-94.

Schenke, D., and Cai, D. (2020). Applications of CRISPR/Cas to improve crop disease resistance–beyond inactivation of susceptibility factors. *iScience* 23.

Schils, R., Olesen, J.E., Kersebaum, K.-C., Rijk, B., Oberforster, M., Kalyada, V., Khitrykau, M., Gobin, A., Kirchev, H., and Manolova, V. (2018). Cereal yield gaps across Europe. *European Journal of Agronomy* 101**,** 109-120.

Schittenhelm, S., Kraft, M., and Wittich, K.P. (2014). Performance of winter cereals grown on field-stored soil moisture only. *European Journal of Agronomy* 52.

Schlegel, R., and Korzun, V. (1997). About the origin of 1RS. 1BL wheat-rye chromosome translocations from Germany. *Plant Breeding* 116**,** 537-540.
Schlegel, R.H.J. (2014). *Rye genetics, breeding and cultivation.* CRC Press Taylor & Francis Group, LLC.
Schlegel, R.H.J. (2016). Hybrid breeding boosted molecular genetics in rye. *Russian Journal of Genetics: Applied Research* 6**,** 569-583.
Schnable, P.S., and Springer, N.M. (2013). Progress toward understanding heterosis in crop plants. *Annual Review of Plant Biology* 64**,** 71-88.
Schreiber, M., Himmelbach, A., Börner, A., and Mascher, M. (2019). Genetic diversity and relationship between domesticated rye and its wild relatives as revealed through genotyping-by-sequencing. *Evolutionary Applications* 12**,** 66-77.
Schwarz, T., Kuleta, W., Turek, A., Tuz, R., Nowicki, J., Rudzki, B., and Bartlewski, P.M. (2014). Assessing the efficiency of using a modern hybrid rye cultivar for pig fattening, with emphasis on production costs and carcass quality. *Animal Production Science* 55**,** 467-473.
Schwarz, T., Turek, A., Nowicki, J., Tuz, R., Rudzki, B., and Bartlewski, P.M. (2016). Production value and cost-effectiveness of pig fattening using liquid feeding or enzyme-supplemented dry mixes containing rye grain. *Czech Journal of Animal Science* 61**,** 341-350.
Scott, M.F., Ladejobi, O., Amer, S., Bentley, A.R., Biernaskie, J., Boden, S.A., Clark, M., Dell'acqua, M., Dixon, L.E., and Filippi, C.V. (2020). Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity* 125**,** 396-416.
Seibel, W., and Weipert, D. (1994). The quality of German rye crop. *Die Muehle + Mischfuttertechnik* 131**,** 575-581.
Senior, B.J. (1947). Pig-production methods in Denmark, Sweden, Holland and Ireland. *Journal of the Statistical and Social Inquiry Society of Ireland* 18**,** 217.
Seong, K., Seo, E., Li, M., and Staskawicz, B. (2019). Evolution of NLR resistance genes with non-canonical N-terminal domains in wild tomato species. *BioRxiv* 786194.
Septiani, P., Lanubile, A., Stagnati, L., Busconi, M., Nelissen, H., Pè, M.E., Dell'acqua, M., and Marocco, A. (2019). Unravelling the genetic basis of Fusarium seedling rot resistance in the MAGIC maize population: novel targets for breeding. *Scientific Reports* 9**,** 1-10.
Shao, Z.-Q., Xue, J.-Y., Wu, P., Zhang, Y.-M., Wu, Y., Hang, Y.-Y., Wang, B., and Chen, J.-Q. (2016). Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiology* 170**,** 2095-2109.
Sharma, A., Srivastava, P., Mavi, G., Kaur, S., Kaur, J., Bala, R., Sohu, V., Chhuneja, P., and Bains, N.S. (2021). Resurrection of wheat cultivar 'PBW343'using marker assisted gene pyramiding for rust resistance. *Frontiers in Plant Science* 12.
Shull, G.H. (1948). What is" heterosis"? *Genetics* 33**,** 439.
Silva, V., Yang, X., Fleskens, L., Ritsema, C., and Geissen, V. (2021). Environmental and human health at risk–scenarios to achieve the EU's 50% pesticide reduction goals. *Research Square***,** 1-25.
Singh, R., Huerta-Espino, J., Bhavani, S., Herrera-Foessel, S., Singh, D., Singh, P., Velu, G., Mason, R., Jin, Y., and Njau, P. (2011a). Race non-specific resistance to rust diseases in CIMMYT spring wheats. *Euphytica* 179**,** 175-186.
Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., Herrera-Foessel, S., Singh, P.K., Singh, S., and Govindan, V. (2011b). The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. *Annual Review of Phytopathology* 49**,** 465-481.
Singh, S.P., Hurni, S., Ruinelli, M., Brunner, S., Sanchez-Martin, J., Krukowski, P., Peditto, D., Buchmann, G., Zbinden, H., and Keller, B. (2018). Evolutionary divergence of the rye Pm17 and Pm8 resistance genes reveals ancient diversity. *Plant Molecular Biology* 98**,** 249-260.
Skowrońska, R., Mariańska, M., Ulaszewski, W., Tomkowiak, A., Nawracała, J., and Kwiatek, M.T. (2020). Development of triticale× wheat prebreeding germplasm with loci for slow-rusting resistance. *Frontiers in Plant Science* 11.
Skriver, K. (1987). *Oversigt over landsforsøgene. Forsøg og undersøgelser i de landøkonomiske foreninger.* Århus, Denmark: Aarhuus Stiftbogtrykkerie.
Słomińska-Durdasiak, K.M., Kollers, S., Korzun, V., Nowara, D., Schweizer, P., Djamei, A., and Reif, J.C. (2020). Association mapping of wheat Fusarium head blight resistance-related regions using a candidate-gene approach and their verification in a biparental population. *Theoretical and Applied Genetics* 133**,** 341-351.
Solodukhina, O.V. (2002). Genetic characterization of rye accessions with regard to leaf rust resistance. *Russian Journal of Genetics* 38**,** 497-506.

Sorkheh, K., Malysheva-Otto, L.V., Wirthensohn, M.G., Tarkesh-Esfahani, S., and Martínez-Gómez, P. (2008). Linkage disequilibrium, genetic association mapping and gene localization in crop plants. *Genetics and Molecular Biology* 31**,** 805-814.

Sortsinfo (2021). *The Danish official trial records* [Online]. Available: https://sortinfo.dk/ [Accessed 20/8 2021].

Spinoni, J., Vogt, J.V., Naumann, G., Barbosa, P., and Dosio, A. (2018). Will drought events become more frequent and severe in Europe? *International Journal of Climatology* 38**,** 1718-1736.

Stam, R., Silva-Arias, G.A., and Tellier, A. (2019). Subsets of NLR genes show differential signatures of adaptation during colonization of new habitats. *New Phytologist* 224**,** 367-379.

Steuernagel, B., Periyannan, S.K., Hernandez-Pinzon, I., Witek, K., Rouse, M.N., Yu, G., Hatta, A., Ayliffe, M., Bariana, H., Jones, J.D., Lagudah, E.S., and Wulff, B.B. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology* 34**,** 652-655.

Stojalowski, S.A., Milczarski, P., Hanek, M., Bolibok-Bragoszewska, H., Myskow, B., Kilian, A., and Rakoczy-Trojanowska, M. (2011). DArT markers tightly linked with the Rfc1 gene controlling restoration of male fertility in the CMS-C system in cultivated rye (Secale cereale L.). *Journal of Applied Genetics* 52**,** 313-318.

Stracke, S., Schilling, A.G., Forster, J., Weiss, C., Glass, C., Miedaner, T., and Geiger, H.H. (2003). Development of PCR-based markers linked to dominant genes for male-fertility restoration in Pampa CMS of rye (Secale cereale L.). *Theoretical and Applied Genetics* 106**,** 1184-1190.

Sucher, J., Boni, R., Yang, P., Rogowsky, P., Buchner, H., Kastner, C., Kumlehn, J., Krattinger, S.G., and Keller, B. (2017). The durable wheat disease resistance gene Lr34 confers common rust and northern corn leaf blight resistance in maize. *Plant Biotechnology Journal* 15**,** 489-496.

Suffert, F., and Thompson, R. (2018). Some reasons why the latent period should not always be considered constant over the course of a plant disease epidemic. *Plant Pathology* 67**,** 1831-1840.

Sul, J.H., Martin, L.S., and Eskin, E. (2018). Population structure in genetic studies: Confounding factors and mixed models. *PLoS genetics* 14**,** e1007309.

Szal, M., Kupryjanowicz, M., Wyczółkowski, M., and Tylmann, W. (2014). The Iron Age in the Mrągowo Lake District, Masuria, NE Poland: the Salęt settlement microregion as an example of long-lasting human impact on vegetation. *Vegetation History and Archaeobotany* 23**,** 419-437.

Sørensen, C.K., Carciofi, M., Hasler-Sheetal, H., Zafari, S., Andrzejczak, O., Hovmøller, M., Møller, I., and Hebelstrup, K. (2019). Overexpression of phytoglobin in barley alters both compatible and incompatible interactions with the mildew pathogen *Blumeria graminis*. *Plant Pathology* 68**,** 152-162.

Tamborski, J., and Krasileva, K.V. (2020). Evolution of plant NLRs: from natural history to precise modifications. *Annual Review of Plant Biology* 71**,** 355-378.

Tang, X., Cao, X., Xu, X., Jiang, Y., Luo, Y., Ma, Z., Fan, J., and Zhou, Y. (2017). Effects of climate change on epidemics of powdery mildew in winter wheat in China. *Plant Disease* 101**,** 1753-1760.

Targonska, M., Bolibok-Bragoszewska, H., and Rakoczy-Trojanowska, M. (2016). Assessment of genetic diversity in Secale cereale based on SSR markers. *Plant Molecular Biology Reporter* 34**,** 37-51.

Thompson, J.N., and Burdon, J.J. (1992). Gene-for-gene coevolution between plants and parasites. *Nature* 360**,** 121-125.

Thor, K., Jiang, S., Michard, E., George, J., Scherzer, S., Huang, S., Dindas, J., Derbyshire, P., Leitão, N., and Defalco, T.A. (2020). The calcium-permeable channel OSCA1. 3 regulates plant stomatal immunity. *Nature* 585**,** 569-573.

Thorling, L., Ditlefsen, C., Ernstsen, V., Hanse, B., Johnsen, A.R., and Troldborg, L. (2021). *Grundvandsovervågning, status og udvikling 1989 - 2019. De nationale geologiske undersøgelser for Danmark og Grøndland, Klima, Energi- og Forsyningsministeriet* [Online]. Available: https://www.geus.dk/Media/2/1/Grundvandsovervågning%201989-2019.pdf [Accessed 9/9 2021].

Torp, J., Jensen, H.P., and Jørgensen, J.H. (1978). Powdery mildew resistance genes in 107 Northwest European spring barley varieties. *Kongelige Veterinær- og Landbohøjskole***,** 75-102.

Tripathi, L., Ntui, V.O., and Tripathi, J.N. (2020). CRISPR/Cas9-based genome editing of banana for disease resistance. *Current Opinion in Plant Biology* 56**,** 118-126.

Trivedi, P., Leach, J.E., Tringe, S.G., Sa, T., and Singh, B.K. (2020). Plant–microbiome interactions: from community assembly to plant health. *Nature Reviews Microbiology* 18**,** 607-621.

Tystofte (2021). *TystofteFoundation, CPVO-entrusted examination office* [Online]. Available: https://www.tystofte.dk/en/ [Accessed 13/9 2021].

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., and Bemm, F. (2019). A species-wide inventory of NLR genes and alleles in Arabidopsis thaliana. *Cell* 178**,** 1260-1272. e1214.

Van Den Berg, F., Gaucel, S., Lannou, C., Gilligan, C., and Van Den Bosch, F. (2013). High levels of auto-infection in plant pathogens favour short latent periods: a theoretical approach. *Evolutionary Ecology* 27**,** 409-428.

Van Wersch, S., and Li, X. (2019). Stronger when together: clustering of plant NLR disease resistance genes. *Trends in Plant Science* 24**,** 688-699.

Varden, F.A., De La Concepcion, J.C., Maidment, J.H., and Banfield, M.J. (2017). Taking the stage: effectors in the spotlight. *Current Opinion in Plant Biology* 38**,** 25-33.

Vilhjalmsson, B.J., and Nordborg, M. (2013). The nature of confounding in genome-wide association studies. *Nature Reviews Genetics* 14**,** 1-2.

Voylokov, A., Fuong, F., and Smirnov, V. (1993). Genetic studies of self-fertility in rye (Secale cereale L.). 1. The identification of genotypes of self-fertile lines for the Sf alleles of self-incompatibility genes. *Theoretical and Applied Genetics* 87**,** 616-618.

Voylokov, A., Korzun, V., and Börner, A. (1998). Mapping of three self-fertility mutations in rye (Secale cereale L.) using RFLP, isozyme and morphological markers. *Theoretical and Applied Genetics* 97**,** 147-153.

Wang, J., Wang, J., Hu, M., Wu, S., Qi, J., Wang, G., Han, Z., Qi, Y., Gao, N., and Wang, H.-W. (2019a). Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science* 364.

Wang, J., and Zhang, Z. (2021). GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *Genomics, Proteomics & Bioinformatics*.

Wang, L., Zhao, L., Zhang, X., Zhang, Q., Jia, Y., Wang, G., Li, S., Tian, D., Li, W.-H., and Yang, S. (2019b). Large-scale identification and functional analysis of NLR genes in blast resistance in the Tetep rice genome sequence. *Proceedings of the National Academy of Sciences* 116**,** 18479-18487.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M., Salvi, S., Milner, S.G., Cattivelli, L., Mastrangelo, A.M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., International Wheat Genome Sequencing, C., Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A.R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K.J., Hayden, M., and Akhunov, E. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol Journal* 12**,** 787-796.

Wang, X., Bakkeren, G., and Mccallum, B. (2010). Virulence and molecular polymorphisms of the wheat leaf rust fungus *Puccinia triticina* in Canada from 1997 to 2007. *Botany* 88**,** 575-589.

Wang, Y., Li, X., Fan, B., Zhu, C., and Chen, Z. (2021). Regulation and function of defense-related callose deposition in plants. *International Journal of Molecular Sciences* 22**,** 2393.

Wang, Y., Tan, J., Wu, Z., Vandenlangenberg, K., Wehner, T.C., Wen, C., Zheng, X., Owens, K., Thornton, A., and Bang, H.H. (2019c). Staygreen, stay healthy: a loss-of-susceptibility mutation in the staygreen gene provides durable, broad-spectrum disease resistances for over 50 years of US cucumber production. *New Phytologist* 221**,** 415-430.

Ward, S.V., and Manners, J.G. (1974). Environmental effects on the quantity and viability of conidia produced by Erysiphe graminis. *Transactions of the British Mycological Society* 62**,** 119-128.

Wehling, P., Linz, A., Hackauf, B., Roux, S.R., Ruge, B., and Klocke, B. (2003). Leaf-rust resistance in rye (Secale cereale L.). 1. Genetic analysis and mapping of resistance genes Pr1 and Pr2. *Theoretical and Applied Genetics* 107**,** 432-438.

Weiss, E., and Zohary, D. (2011). The Neolithic Southwest Asian founder crops: their biology and archaeobotany. *Current Anthropology* 52**,** S237-S254.

Weiss, E., Zohary, D., and Hopf, M. (2012). *Domestication of plants in the old world - The origin and spread of domesticated plants in South-west Asia, Europe, and the Mediterranean Basin.* Oxford University Press on Demand.

Wilde, P., Korzun, V., Menzel, J., Ruonan, Z., Stein, N., and Hackauf, B. (2019). *Restorer Plants. KWS SAAT SE. (U.S. Patent App. 16/064304)*.

Williams, S.J., Sornaraj, P., Decourcy-Ireland, E., Menz, R.I., Kobe, B., Ellis, J.G., Dodds, P.N., and Anderson, P.A. (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant-Microbe Interactions* 24**,** 897-906.

Witek, K., Jupe, F., Witek, A.I., Baker, D., Clark, M.D., and Jones, J.D. (2016). Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology* 34**,** 656-660.

Wójtowicz, A., Wójtowicz, M., Sigvald, R., Czernecki, B., Ratajkiewicz, H., Łacka, A., Zacharczuk, M., and Pasternak, M. (2020). Assessment of the impact of climate change on the latency period of leaf rust on triticale in Poland. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science* 70**,** 195-207.

Wójtowicz, A., Wójtowicz, M., Sigvald, R., and Pasternak, M. (2017). Predicting the effects of climate change on the latency period of wheat leaf rust in western Poland. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science* 67**,** 223-234.

Wolfe, M.S., Brändle, U., Koller, B., Limpert, E., and Mcdermott, J. (1992). Barley mildew in Europe: population biology and host resistance. *Euphytica* 63**,** 125-139.

Wu, J., Van Der Burgh, A.M., Bi, G., Zhang, L., Alfano, J.R., Martin, G.B., and Joosten, M.H. (2018). The bacterial effector AvrPto targets the regulatory coreceptor SOBIR1 and suppresses defense signaling mediated by the receptor-like protein Cf-4. *Molecular Plant-Microbe Interactions* 31**,** 75-85.

Wu, X.X., Xu, X.F., Chen, R.Z., Li, T.Y., and Cao, Y.Y. (2019). Virulence structure and its genetic diversity analyses of *Blumeria graminis* f. sp. *tritici* isolates in China. *BMC Eolutionary Biology* 19**,** 1-11.

Wu, Y., San Vicente, F., Huang, K., Dhliwayo, T., Costich, D.E., Semagn, K., Sudha, N., Olsen, M., Prasanna, B.M., Zhang, X., and Babu, R. (2016). Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theoretical and Applied Genetics* 129**,** 753-765.

Xing, L., Hu, P., Liu, J., Witek, K., Zhou, S., Xu, J., Zhou, W., Gao, L., Huang, Z., and Zhang, R. (2018). Pm21 from Haynaldia villosa encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Molecular Plant* 11**,** 874-878.

Xu, G., Greene, G.H., Yoo, H., Liu, L., Marqués, J., Motley, J., and Dong, X. (2017). Global translational reprogramming is a fundamental layer of immune regulation in plants. *Nature* 545**,** 487-490.

Xue, A.G., Lim, S., Chen, Y., Humphreys, G., Cao, W., Menzies, J., Cowger, C., Li, H., and Serajazari, M. (2021). Virulence structure of *Blumeria graminis* f. sp. *tritici*, the causal agent of wheat powdery mildew, in Ontario, Canada, in 2018 and 2019. *Canadian Journal of Plant Pathology***,** 1-9.

Yadav, P.S., Mishra, V.K., Arun, B., Chand, R., Vishwakarma, M., Vasistha, N., Mishra, A., Kalappanavar, I., and Joshi, A.K. (2015). Enhanced resistance in wheat against stem rust achieved by marker assisted backcrossing involving three independent Sr genes. *Current Plant Biology* 2**,** 25-33.

Yang, D., Tang, J., Yang, D., Chen, Y., Ali, J., and Mou, T. (2019). Improving rice blast resistance of Feng39S through molecular marker-assisted backcrossing. *Rice* 12**,** 1-16.

Yoder, C. (2012). Let them eat cake? Status-based differences in diet in medieval Denmark. *Journal of Archaeological Science* 39**,** 1183-1193.

Yoo, H., Greene, G.H., Yuan, M., Xu, G., Burton, D., Liu, L., Marqués, J., and Dong, X. (2020). Translational regulation of metabolic dynamics during effector-triggered immunity. *Molecular Plant* 13**,** 88-98.

Yu, K., Wang, H., Liu, X., Xu, C., Li, Z., Xu, X., Liu, J., Wang, Z., and Xu, Y. (2020). Large-scale analysis of combining ability and heterosis for development of hybrid maize breeding strategies using diverse germplasm resources. *Frontiers in Plant Science* 11.

Zaw, H., Raghavan, C., Pocsedio, A., Swamy, B.M., Jubay, M.L., Singh, R.K., Bonifacio, J., Mauleon, R., Hernandez, J.E., and Mendioro, M.S. (2019). Exploring genetic architecture of grain yield and quality traits in a 16-way indica by japonica rice MAGIC global population. *Scientific Reports* 9**,** 1-11.

Zhang, B., Chi, D., Hiebert, C., Fetch, T., Mccallum, B., Xue, A., Cao, W., Depauw, R., and Fedak, G. (2019). Pyramiding stem rust resistance genes to race TTKSK (Ug99) in wheat. *Canadian Journal of Plant Pathology* 41**,** 443-449.

Zhang, F., and Batley, J. (2020). Exploring the application of wild species for crop improvement in a changing climate. *Current Opinion in Plant Biology* 56**,** 218-222.

Zhang, J., Pu, R., Yuan, L., Huang, W., Nie, C., and Yang, G. (2014). Integrating remotely sensed and meteorological observations to forecast wheat powdery mildew at a regional scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7**,** 4328-4339.

Zhang, L., and Dickinson, M. (2001). Fluorescence from rust fungi: a simple and effective method to monitor the dynamics of fungal growth in planta. *Physiological and Molecular Plant Pathology* 59**,** 137-141.

Zhang, M., Chiang, Y.-H., Toruño, T.Y., Lee, D., Ma, M., Liang, X., Lal, N.K., Lemos, M., Lu, Y.-J., and Ma, S. (2018). The MAP4 kinase SIK1 ensures robust extracellular ROS burst and antibacterial immunity in plants. *Cell Host & Microbe* 24**,** 379-391. e375.

Zhang, Z., Henderson, C., and Gurr, S.J. (2004). *Blumeria graminis* secretes an extracellular catalase during infection of barley: potential role in suppression of host defence. *Molecular Plant Pathology* 5**,** 537-547.

Zhang, Z., Henderson, C., Perfect, E., Carver, T., Thomas, B., Skamnioti, P., and Gurr, S. (2005). Of genes and genomes, needles and haystacks: *Blumeria graminis* and functionality. *Molecular Plant Pathology* 6**,** 561-575.

Zhang, Z., Liu, Y., Huang, H., Gao, M., Wu, D., Kong, Q., and Zhang, Y. (2017). The NLR protein SUMM 2 senses the disruption of an immune signaling MAP kinase cascade via CRCK 3. *EMBO Reports* 18**,** 292-302.

Zhao, J., Xu, C., Xu, J., Huang, L., Zhang, D., and Liang, D. (2018). Forecasting the wheat powdery mildew (*Blumeria graminis* f. sp. *tritici*) using a remote sensing-based decision-tree classification at a provincial scale. *Australasian Plant Pathology* 47**,** 53-61.

Zhong, Z., Marcel, T.C., Hartmann, F.E., Ma, X., Plissonneau, C., Zala, M., Ducasse, A., Confais, J., Compain, J., and Lapalu, N. (2017). A small secreted protein in Zymoseptoria tritici is responsible for avirulence on wheat cultivars carrying the Stb6 resistance gene. *New Phytologist* 214**,** 619-631.

Zhou, Q., Li, L., and Wang, L. (2021). Complex System Analysis of *Blumeria graminis* f. sp. *tritici* Isolates Collected in Central Hebei Province, China. *Discrete dynamics in nature and society* 2021**,** 1-8.

Zhu, H., Li, C., and Gao, C. (2020). Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nature Reviews Molecular Cell Biology* 21**,** 661-677.

Zhu, J., Zhou, Y., Shang, Y., Hua, W., Wang, J., Jia, Q., Liu, M., and Yang, J. (2016). Genetic evidence of local adaption and long distance migration in *Blumeria graminis* f. sp. *hordei* populations from China. *Journal of General Plant Pathology* 82**,** 69-81.

Ziv, C., Zhao, Z., Gao, Y.G., and Xia, Y. (2018). Multifunctional roles of plant cuticle during plant-pathogen interactions. *Frontiers in Plant Science* 9.