Norwegian University of Life Sciences
Faculty of Biosciences
Department of Plant Sciences

Philosophiae Doctor (PhD)
Thesis 2024:9

# Oat breeding in the era of genomics - resistance to Fusarium

## Havreforedling i den genomiske tidsalder - resistens mot Fusarium

Espen Sannes Sørensen

# Oat breeding in the era of genomics-resistance to Fusarium

Havreforedling i den genomiske tidsalder - resistens mot Fusarium

Philosophiae Doctor (PhD) Thesis

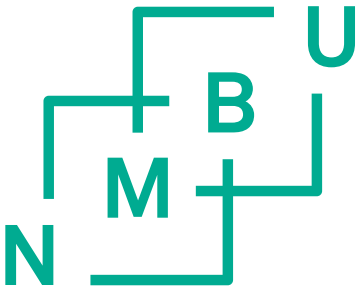
Espen Sannes Sørensen

Norwegian University of Life Sciences

Faculty of Biosciences

Department of Plant Sciences

Ås (2023)




Thesis number 2024:9

ISSN 1894-6402

ISBN 978-82-575-2126-4

# Supervisors and Evaluation Committee

**PhD supervisors:**

Prof. Morten Lillemo
morten.lillemo@nmbu.no
Faculty of Biosciences
Department of Plant Sciences
Norwegian University of Life Sciences (NMBU)
P. O. Box 5003 NMBU, NO-1432 Ås
Norway

Dr. Muath Alsheikh
muath.alsheikh@graminor.no
Graminor AS
Hommelstadvegen 60, NO-2322 Ridabu
Faculty of Biosciences
Department of Plant Sciences
Norwegian University of Life Sciences (NMBU)
P. O. Box 5003 NMBU, NO-1432 Ås
Norway

Dr. Anna Kristina Sonesson
anna.sonesson@nofima.no
Nofima AS
Osloveien 1, 1433 Ås
Norway

Dr. Constantin Jansen
constantin.jansen@graminor.no
Graminor AS
Hommelstadvegen 60, NO-2322 Ridabu
Norway

Dr. Susanne Windju
Susanne.windju@graminor.no
Graminor AS
Hommelstadvegen 60, NO-2322 Ridabu
Norway

**Evaluation committee:**

Dr. Alf Ceplitis
alf.ceplitis@lantmannen.com
Lantmännen seed
SE-205 03 Malmö
Sweden

Marja Johanna Jalli
marja.jalli@luke.fi
Natural Resources Institute Finland (Luke)
Planta, Tietotie 4, FI-31 600 Jokioinen
Finland

# Contents

# Acknowledgements

I would like to express my gratitude to my dedicated team of supervisors: Muath Alsheikh, Morten Lillemo, Anna Kristina Sonesson, Constantin Jansen, and Susanne Windju. Their invaluable guidance and contributions throughout the project have been instrumental in its achievement. I am deeply appreciative of their unwavering support, which has been a source of strength during both the highs and lows of this journey, particularly in the writing phase.

I would also like to thank Graminor AS for selecting me as their PhD candidate. For trusting me and supporting me through many years of extensive field trials. The scope and ambition for this project would not have been realized without their experience, funding and technical support. Further I extend my thanks the dedicated technical staff at Graminor, who managed the disease trials at Staur and conducted the post-harvest analysis. A countless talented individuals have participated in this project in one way or another, from ploughing and tillage, weighing and sowing, fertilizing and spraying against pests, inoculation and irrigation, field notes and scaring away crows, harvesting and cleaning samples, milling and preparing for analysis, ELISA analysis and counting healthy seedlings, for at least 10 000 samples during the PhD period.

I would also like to thank the efforts of the SafeGrains and RESIFUS project members who laid the groundworks for my PhD research in terms of field testing, analysis and material. A special thanks goes to Prof. Åsmund Bjørnstad who's knowledge and experience in oat research gave valuable insight. His advice has been most welcome from start to finish and beyond. I also want to thank Lars Reitan who's work in the RESIFUS project helped implement Fusarium screening at Staur research station under management of Graminor.

A very special thanks and appreciation to Trond Buraas and Halina Ruud who steered a very successful oat breeding program. Their hard work through many decades gave Norway new varieties of high productivity, quality, and disease resistance. Their experience and knowledge in oat breeding has been an inspiration to me and has shaped me into the oat breeder I am today. Some might say that plant breeding is a numbers game, but I would argue that it is also a craftmanship and an artform as Trond and Halina has shown me.

Thank you to the Research Council of Norway for funding and supporting my PhD research.

A big thank you to the department of Plant Sciences, and especially the Plant breeding section. It has been a good place to work and discuss science and challenges through the PhD period. Prof. Morten Lillemo has been a great supervisor and mentor through my education from Master to PhD. His

advice regarding plant breeding is always welcome, and I especially appreciate his newfound enthusiasm for oat breeding.

I also want to thank the International Maize and Wheat Improvement Centre (CIMMYT), and especially Dr. Jose Crossa who took care of me, my family, and colleagues during a 3 month stay in the wonderful country of Mexico. It was a truly inspiring and valuable experience that contributed a lot to my work.

Lastly, I want to thank my wife who has supported me through this period of our lives, who moved with me to a new city, started a family and built a home with me. I look forward to the post-doc experience with you and our family in our home at Ilseng.

Ridabu November 2023, Espen

# Abstract

In Norway, oat, along with wheat and barley, constitute one of the three major crops, valued as a break crop due to its minimal disease overlaps with the others. However, a significant threat to the cereal crops, particularly oats, is Fusarium head blight (FHB) disease, caused by the fungal pathogen *Fusarium graminearum.* FHB negatively impacts yield, quality, and germination ability, and it also generates a mycotoxin called deoxynivalenol (DON), which poses health risks to both humans and animals. The occurrence of Fusarium head blight is promoted by environmental conditions such as continuous precipitation, warm summer, and prolonged harvest periods. Previous outbreaks of this disease have had severe consequences for the oat industry, resulting in reduced grain prices for farmers, diminished high-quality grains for millers and the feed industry, and lower-quality seeds with reduced germination capacity for seed producers.

Developing resistant oat varieties through traditional breeding methods is time-consuming and costly, involving extensive field trials and post-harvest analysis of DON content and germination potential. Recent initiatives aimed at enhancing resistance in Norwegian oat varieties have identified potential sources of resistance and implemented field trials and analysis through the breeding company Graminor, even at a significant cost. Genetic studies have revealed that resistance to Fusarium in oats is a highly quantitative trait influenced by numerous small-effect Quantitative Trait Loci (QTL). This makes it a promising candidate for genomic selection (GS), a DNA-based breeding technique. GS enables breeders to perform selection at earlier stages based on predicted breeding values calculated from statistical models and DNA markers. To effectively implement GS in the Norwegian breeding program, a set of training individuals must be genotyped and accurately phenotyped. This project aimed to introduce genomic selection into the Norwegian breeding program through three tasks. The first task involved evaluating strategies to optimize training populations and resulted in an article published in autumn 2022 which concluded that population size and genetic similarity between training and testing populations was the most important criteria to optimize. Genetic and phenotypic diversity was less important, but still played a vital role as long as the other criteria were optimized. The Prediction core strategy worked best in balancing these optimization criteria and resulted in significantly higher prediction ability than random selection. An optimized training population was sown in mist irrigated and inoculated disease trials in two locations (Staur and Vollebekk) in three years (2020, 2021 and 2023) and analysed for DON content and germination ability. Three validation populations were also analysed in separate years in 2020, 2021 and 2022. Data were used in task 2 and 3. The second task involved conducting genome-wide association mapping to pinpoint crucial regions for FHB resistance in Norwegian germplasm. This research successfully identified 15 different QTL-regions, with five of them consistently validated in breeding material as having a substantial

impact. The third task employed phenotypes, whole-genome markers, and the significantly associated QTL-regions from the association mapping to predict the phenotypes of three distinct breeding populations. Using these QTL-regions as fixed effects in GS, the project achieved moderately high accuracy in predicting DON and germination percentages, ranging between 0.44 and 0.47 for DON in the breeding populations. Although the added QTL-regions enhanced accuracy individually and in a cross-validation setup within the training population, their impact varied across different breeding populations.

The results of this thesis will facilitate Graminor's adoption of GS for FHB resistance, reducing breeding costs and enabling screening in earlier generations. Furthermore, it will serve as a foundation for the implementation of GS for other traits, thereby enhancing the overall efficiency of oat breeding.

# Norsk sammendrag

Havre er blant de tre viktigst kornslagene i Norge, og er ansett som god art for vekstskifte med bygg og hvete fordi de har få sykdommer til felles. Men, en sykdom som har potensielt store konsekvenser i alle kornslagene, og spesielt havre, er aksfusariose. Det er en sykdom forårsaket av sekksporesoppen *Fusarium graminearum*. Den reduserer avling, kvalitet, spireprosent og produserer et giftstoff som heter deoxynivalenol (DON) som er skadelig for både mennesker og dyr. Giftstoffet er strengt regulert i EU og Norge, og i år med store smitteutbrudd kan bøndene få redusert pris på levert korn. I utbruddene mellom 2008-2012 måtte 30% av havrepartiene forkastes på grunn av lav spireevne, og 40% av leverte partier fikk redusert pris på grunn av forhøyede DON verdier. Sammen med jordbearbeiding og god agronomisk praksis er bruk av resistente sorter en av de beste måtene å redusere risikoen for høyt smittepress.

Utviklingen av resistente sorter tar lang tid og koster mye i form av feltforsøk og analyser for DON og spireprosent. Nye forskningsprosjekter satte som mål å finne kilder til resistens og implementere ny kunnskap for å forbedre resistensforedlingen. Disse studiene viste at resistensen i havre er kvantitativ med mange gener som hver for seg gir liten effekt. En foredlingsteknikk som kan tas i bruk for å effektivisere resistensforedlingen er genomisk seleksjon (GS). Dette er en metode som baserer seg på å bruke genetiske og fenotypiske data for en populasjon til å trene en statistisk modell som skal kunne forutsi resistensen til nytt foredlingsmateriale basert på kun genetiske data. Dette vil gjøre det mulig å gjøre seleksjon av materiale i tidlige generasjoner og dermed øke seleksjonsintensiteten for Fusariumresistens. Dette kan også potensielt spare kostnader i feltforsøk og analyser. Målet med dette prosjektet var å implementere genomisk seleksjon i det norske havreforedlingsprogrammet gjennom tre arbeidspakker. I arbeidspakke 1 ble det benyttet en større samling foredlingslinjer og historiske sorter til å evaluere forskjellige strategier for optimalisering av treningspopulasjoner. Dette resulterte i en artikkel som ble publisert i 2022 som konkluderte at de viktigste parameterne å optimalisere var størrelse på populasjonen og den genetiske likheten mellom trenings- og testingspopulasjonen. Genetisk og fenotypisk diversitet spilte mindre viktig rolle, men hadde fortsatt betydning gitt at de andre kriteriene var optimalisert. Det var strategien «Prediction core» som klarte å balansere optimaliseringskriteriene best og strategien resulterte i signifikant høyere treffsikkerhet enn et tilfeldig utvalg. Dette arbeidet resulterte også i en optimalisert treningspopulasjon på 541 genotyper som ble sådd ut i inokulerte Fusarium forsøk med dusjvanning for å gi gode forhold for infeksjon. Forsøkene ble gjennomført med to gjentak, på to steder (Vollebekk og Staur) i tre år (2020, 2021 og 2022). Prøvene ble analysert for DON innhold og spireprosent. Data fra forsøkene ble brukt i arbeidspakke 2 and 3. Arbeidspakke 2 var å gjøre en assosiasjonskartlegging for Fusarium resistens og

resulterte i 48 signifikante markører fordelt over 15 områder på genomet. Resistente og mottakelige alleler av disse områdene ble identifisert, hvorav fem ble validert med konsistent effekt over miljø og viste resistens i tre ulike valideringspopulasjoner. Disse resultatene ble brukt i arbeidspakke 3 hvor genomiske prediksjonsmodeller ble brukt som inkluderte effekten av de fem validerte gen områdene, samt resten av resultatene fra assosiasjonskartleggingen for å forutsi DON og spireprosent i de tre valideringspopulasjonene. Resultatet var modeller som gav moderat høy treffsikkerhet mellom 0.44 og 0.47 for DON. Effekten av å bruke informasjonen fra assosiasjonskartleggingen varierte med egenskap og populasjoner, med i de fleste tilfellene ble treffsikkerheten høyere eller uforandret. Selv om inkludering av markører fra alle signifikante QTL resulterte i høyere treffsikkerhet i treningspopulasjonen, førte det til lavere treffsikkerhet i valideringspopulasjonene. Resultatene i denne avhandlingen vil gi Graminor muligheten til å implementere genomisk seleksjon for Fusariumresistens i havreforedlingsprogrammet. Arbeidet gir også et fundament for videre utvikling av genomisk seleksjon som seleksjonsmetode i planteforedling.

# Abbreviations

| | |
|---|---|
| GS | Genomic selection |
| DTH | Days to heading |
| PH | Plant height |
| DON | Deoxynivalenol |
| GP | Germination percentage |
| PA | Prediction ability |
| SNP | Single nucleotide polymorphism |
| HT | Haplotype |
| VP | Validation population |
| TP | Training population |
| GWAS | Genome wide association study |
| FHB | Fusarium head blight |
| LD | Linkage disequilibrium |
| MAS | Marker assisted selection |
| QTL | Quantitative trait locus/loci |
| DArT | Diversity array technology |
| G x E | Genotype x Environment |
| GBS | Genotyping by sequencing |
| SSR | Short sequence repeat |
| AFLP | Amplified fragment length polymorphism |
| MAF | Minor allele frequency |

# List of papers

1. **Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa* L.).** (Sørensen, E. S., Jansen, C., Windju, S., Crossa, J., Sonesson, A. K., Lillemo, M. & Alsheikh, M. Plant Breeding (2023), 142(1), 41–53)

2. **Identification of haplotypes associated with resistance to *Fusarium graminearum* in spring oat (*Avena sativa* L.).** (Sørensen, E. S., Windju, S., Jansen, C., Sonesson, A. K., Lillemo, M. & Alsheikh, M. Manuscript. (In press 2023)

3. **Evaluation of genomic prediction for fusarium resistance in the Norwegian oat breeding program.** (Sørensen, E. S., Windju, S., Jansen, C., Sonesson, A. K., Lillemo, M. & Alsheikh, M. Manuscript.

# 1. Introduction

## 1.1. Oat production and value

Oats hold a significant position among cereal crops in Norway, ranking third after wheat and barley. Traditionally in Norway, oats have primarily been utilized as animal feed, accounting for approximately 86% of their usage per 2021 (Opplysningskontoret for brød og korn, 2022). However, there has been a significant increase in human consumption in Norway from approximately 3 kg per capita in 2001 to 7 kg in 2021. During the same period, oats occupied an average of 24% of the cereal area (SD 1.6; Statistisk sentralbyrå, 2023a), while wheat encompassed 25% (SD 3.5) and barley 49% (SD 3.9). The average annual oat production in the period was 280 thousand tons (SD 49.9; Statistisk sentralbyrå, 2023b) and the average yield per hectare of 3.8 tons (SD 0.58; Statistisk sentralbyrå, 2023c). The production of oats has faced a notable decline during the same timeframe, with an average decrease of 1990 tons per year, and a decrease in area of 1 048 hectares per year while the yield per hectare increased with only 0.03 tons per year. The shift in consumption patterns, characterized by increased human oat consumption and a shrinking cultivation area, underscores the urgency of enhancing oat yields and quality in the years to come. Achieving this objective will likely require more focus on both field management practices and the development of new oat varieties.

Oats are widely recognized as a highly healthy food source with high degree of polyunsaturated fatty acids, essential amino acids (Rafique *et al.,* 2022), and dietary fiber. Among the dietary fibers found in oats, a significant portion consists of the polysaccharide β-glucan, which has been demonstrated to have cholesterol-lowering properties, thereby reducing the risk of coronary heart disease (Mathews *et al.,* 2020). β-glucan has also shown promise in improving glycemic control, which can significantly mitigate health risks associated with diabetes (Pino *et al*., 2021). Oat in Norway is primarily consumed as part of breakfast cereals, porridge, pastry and bread. However, the potential for expanding the range of oat-based products is substantial. Oats can serve as useful ingredient in culinary applications, including oat-based milk (Basinskiene & Cizeikiene, 2020), noodles (Aydin & Gocmen, 2011) and meat substitutes (Ball *et al.,* 2021) exemplifying the diverse possibilities for incorporating this nutritious grain into a wide array of dishes. One of the main quality criteria for both animal and human consumption is low levels of mycotoxin content produced by different fungal pathogens belonging to the genus *Fusarium.*

## 1.2. Fusarium head blight

*Fusarium* is a large genus of fungi that are commonly found in soil and often associated with plants. The taxonomy is complex, and more than 1000 species have been classified into the *Fusarium* genus. Most of them are harmless and live in the soil microbiome. Some, however, are pathogens that infect plant tissues and cause Fusarum head blight (FHB) in cereals. In Norway, one of the most economically impactful species affecting oats is *Fusarium graminearum* (Bernhoft *et al.,* 2013), which also infect both wheat and barley. The impact of the disease comes from reduced yield (Salgado *et al.,* 2015), quality *(*Havrlentová, 2021), germination ability (Tekle *et al.,* 2013a) and production of a mycotoxin called deoxynivalenol (DON; Sobrova *et al.,* 2010), which is toxic to both humans and animals. Another significant species in Norway is *Fusarium avenaceum,* which is the most widespread species but produces mycotoxins with varying levels of toxicity. *Fusarium langsethiae* is another relevant species that primarily infects oats, producing the toxins HT2 and T2, which have been found to be more toxic than DON (European Food Safety Authority, 2017) but tends to cause less symptoms in the plants than *F. graminearum*. Fusarium infections are highly influenced by environmental conditions. Factors such as high temperature and precipitation during summer, and cold and wet conditions during harvest are strongly correlated with high DON concentration (Langseth & Elen, 1997).

### 1.2.1. Infection cycle and symptoms

*F. graminearum* survives the winter as perithecia or sporodochia on plant residues (Trail, 2009). Perithecia are the results of sexual reproduction and produce wind dispersed ascospores, which under favorable conditions can spread for several kilometers (Keller *et al.,* 2014). During the season, the fungus also produces asexual spores called conidia. These conidia form within structures called sporodochia, which develop on plant tissue and soil debris. The conidia spores are primarily rain-dispersed over short distances (Paul *et al.,* 2004).

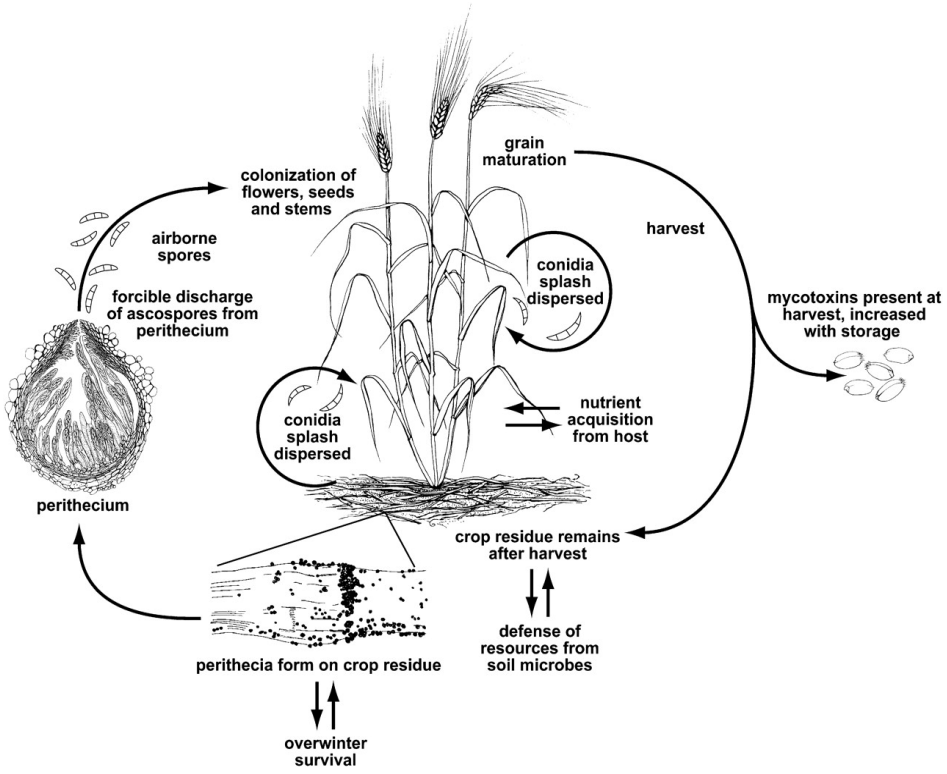



***Figure 1****: Frances Trail. Plant Physiology, Volume 149, Issue 1, January 2009, Pages 103–110, https://doi.org/10.1104/pp.108.129684. © 2009 American Society of Plant Biologists.*

*F. graminearum* is a hemibiotrophic pathogen characterised by both biotrophic and necrotrophic phases. When the spores land on susceptible plant tissue, the fungus produces hyphae that grow in the intercellular space, and subsequently penetrates the cell walls. It is at this stage that DON interacts with the plant's defence mechanisms, which induces cell death (Audenaert *et al.,* 2013). This cell death provides the fungus with resources for continuous growth and propagation. The flowering stage of the plant is particularly vulnerable to infection, as the heads become the primary entry point for the disease. During this time, the kernels are exposed with pollen tissue, providing potential opportunities for the pathogen to establish itself and compromise the crop's overall health and productivity. FHB can be observed in the plants as orange or pink mycelia in the heads (Figure 2). Infected seeds become small and shriveled reducing both yield and test weight. Indigestion of mycotoxin accumulated in the seeds can cause inflammation in the intestines (Kang *et al.,* 2019) which has several symptoms like nausea, diarrhea, headache, and fever (Sobrova *et al.,* 2010). In oats most of the DON is concentrated in the husks (Brodal *et al*., 2020), which are removed during

processing for human consumption, but not necessarily in the case of animal feed production. Contaminated seed gets reduced germination ability, followed by seedling blight after sowing.

***Figure 2**: Infected oat kernel with orange mycelia. Photo: Espen Sannes Sørensen. August 2023.*

### 1.2.2. Resistance and management

The resistance to FHB is commonly characterized into five types (Hautsalo *et al.,* 2018); (1) resistance to initial infection (Schroeder & Christensen, 1963), (2) resistance to spread of infection, (3) resistance to kernel infection, (4), tolerance (Mesterházy, 1995), and (5) resistance to DON accumulation (Miller *et al.,* 1985, Mesterházy *et al.,* 1999), and as such Types 3-5 are confounded in Types 1 and 2. Types 1 and 2 are usually measured visually as number of infected spikelets or with quantitative PCR to quantify the amount of *F. graminearum* DNA. The differentiation between Type 1 and 2 is done in controlled environments where the plants are either spray inoculated (Type 1) or point inoculated (Type 2) (Miedaner *et al.,* 2003, Geddes *et al.,* 2008). Type 3 is usually measured as either number of shriveled seeds or reduction in germination ability. Type 4 is assessed as the reduction of grain yield and degree of kernel damage compared to amount of *F. graminearum* fungus or DON accumulation. Type 5 focuses on the amount of DON in samples, typically as micrograms per kilogram (µg/kg) or parts per million (ppm). In addition to the above-mentioned resistance Types there are several passive avoidance mechanisms that play a role in reducing infection. Some of these mechanisms include the (i) formation of the panicle morphology with increased spacing between the spikelets,

which increases Type 2 resistance, (ii) time of flowering which could ensure that the flowering finishes in periods with lower infection pressure, (iii) plant height which increases the distance from panicles to the conidia spores in the soil, and (iv) anther extrusion which is the ability to throw away the anthers after pollination since remaining anthers can serve as access points for the pathogen (Tekle *et al.,* 2020).

Effective control of *F. graminearum* can be achieved through various field management and agronomic practices. One approach is to minimize available plant material for the pathogen to reproduce. The plant debris can be removed physically, with increased tilling and enhanced decomposition (Leplat *et al.,* 2013). Frequent crop rotation with non-host crops like vegetables or protein crops has been shown to reduce the propagation of the pathogen (Dong *et al.,* 2023). Some weed has also shown potential for propagating *F. graminearum* (Matelionienė *et al.,* 2022), therefore, good weed control would also contribute. A different approach is to make the plants as viable as possible to better protect themselves from infection and tolerate the symptoms. This can be achieved by using certified and healthy seeds, early sowing and sufficient fertilization which makes the plants more robust before infection (Blandino *et al.,* 2009, Arata *et al.,* 2022, Mielniczuk & Skwaryło-Bednarz, 2020). A more direct approach is to use fungicides to reduce the infection, which could reduce the mycotoxin levels by 50% if applied at the optimal time during flowering (Felleskjøpet agri SA., 2023). A final strategy is to treat the seeds post-harvest. The seeds that are to be utilized for sowing can be treated chemically or non-chemically (Piñeros-Guerrero *et al.,* 2019) to avoid reduced germination. The level of DON has been shown to increase during post-harvest storage (Portell *et al.,* 2020) and applying biological control like bacteria or yeast has shown to either convert the DON to less toxic products (Ji *et al*., 2016) or reduce the growth of the fungi (Podgórska-Kryszczuk *et al.,* 2022).

## 1.3. Fusarium research in Norway

### 1.3.1. Fusarium epidemics in oat farming

The issue of Fusarium infection and mycotoxins started to become a significant problem in Norway during the period from 1980 to 2000. During this period, there was noticeable increase in the levels of DON content in analyzed grains, particularly in oats, with mean values exceeding 1000 µg/kg, and some samples exceeded the EU limits of 1750 µg/kg for unprocessed oat (Langseth & Elen, 1997, European Commission, 2006). The main DON producing species at the time were *F. culmorum* (Kosiak *et al.,* 1997 and 2003) while *F. graminearum* was detected in smaller amounts in the main oat growing areas. This changed in the early 2000 as *F. graminearum* became more frequent in Norway (Bernhoft *et al*., 2010) and resulted in elevated DON content. *F. culmorum* does not have a known

sexual stage in Norway (Kant *et al*., 2017), but *F. graminearum* does (Jørstad, 1945) enabling it to adapt and change in virulence and fitness (Becher *et al*., 2010). Recombination of isolates can also lead to resistance against fungicides (de Chaves *et al*., 2022), which emphasizes the need to develop resistant varieties. A study on inoculated and naturally infected fields did not reveal any significant resistance in the varieties at the time (Elen *et al*., 2003), which initiated further research into resistance breeding.

### 1.3.2. SafeGrains project

The elevated DON content and increased frequency of *F. graminearum* led to the initiation of the "SafeGrains" project (The research council of Norway, project # 178273, 2013-2016), which successfully established protocols for spawn inoculation of *F. graminearum* (Tekle *et al.,* 2018). These protocols were implemented at the Norwegian University of Life Sciences (NMBU) and Graminor plant breeding company. During the course of the project, a valuable partnership was established with the University of Minnesota (USA), providing NMBU with affordable access to mycotoxin analysis via gas chromatography (GS/MS). This partnership has provided the project partners with the opportunity to analyze a larger number of samples. One of the main objectives regarding oats was to assess the existing varieties and identify new sources for resistance (Bjørnstad & Skinnes, 2008). This objective was successfully achieved, leading to the removal of susceptible varieties from the market by seed companies. This project also led to Graminor initiating crossing and field testing for Fusarium resistance. Two studies on QTL mapping in oats were conducted (He *et al.,* 2013, Bjørnstad *et al.,* 2017) which identified several loci associated with FHB resistance and concluded that the resistance is caused by many small effect genes, and that there is a significant interaction between the genes and environment.

### 1.3.3. RESIFUS project

A major FHB outbreak took place in the period between 2008 and 2012 (Bernhoft *et al.,* 2013, Sundheim *et al.,* 2013). A substantial 30% of seed lots were discarded due to low germination number in 2009 (Hofgaard *et al.,* 2016), and 40% of the shipments of grain between 2010 and 2012 exceeding regulated limits of DON content (Felleskjøpet agri SA, 2016). This led to the initiation of the "RESIFUS" project in 2014 (The research council of Norway, project # 233908, 2018-2021). The primary focus of this project was the development of a core collection called "Havrebasis" designed to represent the diversity and historical evolution of northern European oat breeding, incorporating potential resistance sources form Europe and North America. Throughout the project's duration, this population was systematically evaluated in field trials with inoculation, concluding that the best sources of FHB resistance already existed in the Norwegian breeding material. Furthermore, the

project aimed to assess cost-effective and high-throughput phenotyping methods, specifically the use of enzyme-linked immunosorbent assay (ELISA) and near-infrared spectroscopy (NIR) calibrations (Tekle *et al.,* 2013b). These results were compared to the GS/MS data from University of Minnesota. The results from the NIR calibrations were successful, particularly during years with high disease pressure (Tekle *et al.,* 2015). However, ELISA showed good correlations with GS/MS even with low disease pressure and was established as the main screening method for FHB resistance at Graminor. This prompted an increased testing of breeding material.

### 1.3.4. FHB resistance breeding at Graminor

Graminor was heavily involved in both SafeGrains and RESIFUS projects resulting in thorough screening of the breeding material for mycotoxin content and germination ability and identification of resistant varieties and breeding lines used for future crossings. Oat is an inbreeding species; therefore, several generations of self-pollination must occur after crossing before homozygous lines can be tested and phenotyped. The traditional breeding strategy at Graminor has been to do shuttle breeding with off-season nurseries in New Zealand allowing two generations per year. Every second generation the most promising single panicles from each population are selected and pooled together into the next generation until the population reaches $F_7$. Single panicles from $F_7$ are selected for phenotypic analysis on single rows in $F_8$, and the most promising lines for simple agronomic traits like height, earliness and overall impression are selected for small yield plots in $F_9$. The small plots are then analyzed for yield, lodging and quality traits like test weight and thousand kernel weight. Approximately 300 of these ($F_{10}$) lines are selected for replicated yield trials over several locations and included in inoculated Fusarium trials at the research stations, Vollebekk and Staur, with subsequent retesting in the next generations. This results in a thorough screening and understanding of the FHB resistance of new varieties with at least three years of field testing in two locations before they are considered for official variety testing. This approach has been successful in providing Norway with highly resistant varieties which dominate the Norwegian market per 2023. The content of DON and *F. graminearum* in analyzed samples from the Food safety department's surveillance program for mycotoxins has been reduced over recent years. In 2022 DON values in oat were the lowest since the surveillance started in 2002 (Bernhoft *et al.,* 2023). This reduction can be due to a combination of favorable weather, introduction of more resistant oat varieties, and improved field management practices. However, with higher temperature and increased rainfall, the future climate predictions are anticipated to favor the pathogen (The Norwegian Centre for Climate Services, 2017), making ongoing breeding efforts to develop more resistant varieties a crucial factor in combating the next potential outbreak.

## 1.4. DNA-based breeding techniques in oat

### 1.4.1. Oat genetics

Oat (*Avena sativa* L.) is an allohexaploid species with a total of 21 different chromosome pairs divided equally into three diploid sub genomes called A, C and D. This genetic makeup is thought to be the results of a fusion of a diploid wild oat named *Avena longiglumis* (AA; Kamal *et al.,* 2022) and the tetraploid species *Avena insularis* (CCDD; Ladizinsky, 1998). The Avena genus contains approximately 30 species. The total size of the genome of cultivated hexaploid oat is estimated to be between 11 and 12.5 Gigabase pairs (Gbp; Yan *et al.,* 2016, Peng *et al.,* 2022). The oat genome is considered to be unstable with large chromosomal translocations within chromosomes and between sub genomes (Jellen *et al.,* 1994, Jellen *et al.,* 1997, Chaffin *et al.,* 2016). The complexity of the oat genome, characterized by its size, instability, and abundance of repeated sequences (Liu *et al.,* 2019) has made the genome assembly and mapping challenging. The first physically anchored consensus map of oat was published in 2013 (Oliver *et al.,* 2013) based on 985 genetic markers. Subsequently, this map was updated by Chaffin *et al*. (2016) with a more extensive set of 7,202 SNP-markers providing a more comprehensive and detailed view of the oat genome. Recently three whole genome reference sequences were published in oats. The first reference genome map was published in 2021 of the American accession "OT3098" (Pepsico, 2021), followed by the Swedish variety "Sang" (Kamal *et al.,* 2022) and the Chinese hulless landrace "Sanfensan" (Peng *et al.,* 2022). These reference genomes hold immense promise in enhancing gene annotations and facilitating more precise genome mapping efforts in the future.

### 1.4.2. DNA-based breeding methods

DNA-based breeding, also known as molecular breeding, has revolutionized the field of plant breeding. This cutting-edge approach utilizes genetic information and DNA markers to accelerate the development of new plant varieties with desirable traits. By identifying specific DNA sequences associated with traits such as disease resistance, higher yields, or improved nutritional content, plant breeders can make more informed and precise breeding choices. This not only expedites the breeding process but also minimizes the need for extensive field trials and reduces the risk of undesirable traits being inadvertently introduced. Two important types of DNA-based methods used in plant breeding are marker-assisted selection (MAS) and genomic selection (GS). Both are established selection methods in plant breeding, relying on marker-trait associations. The choice between them is heavily influenced by the complexity of the underlying genetics. The causal genes can be either qualitative, meaning that the observed trait is caused by a single gene, or quantitative caused by several genes. Many of the breeding traits like yield and quality are often caused by quantitative trait loci (QTL)

which can have large or small effects. In addition, the genotype by environment (G x E) interaction can heavily influence the observed effect. MAS is typically better suited for traits with a limited number of underlying key genes as they are easier to combine into a new variety, and less genetic information is needed to use MAS, making genotyping cheaper. This makes it more applicable in larger populations early in the breeding cycle. GS is a powerful method for complex traits influenced by numerous genetics factors and environmental interactions and has often been found to outperform MAS for quantitative traits (Cerrudo *et al.,* 2018). When traits are controlled by QTL of varying effect MAS and GS could be combined by adding marker effects into GS models (Merrick *et al.,* 2021) or used in different generation of the breeding cycle. The goal of both methods is to make selections as early as possible to screen material before phenotyping. It is therefore important to combine them with speed breeding strategies where more generations per year are achieved in order to make the most of the increased selection intensity. Graminor is currently implementing single seed decent (SSD) using greenhouse facilities to reduce the time from planting to harvest. With SSD it is possible to increase the number of generations per year to four or five (Kigoni *et al.,* 2023).

### 1.4.3. Marker assisted selection

Marker assisted selection is based on the establishment of strong linkage between specific molecular markers and the chromosomal location of the gene(s) responsible for the trait in question. These markers are found through two main methods: linkage mapping (LM) and genome-wide association studies (GWAS). LM often uses biparental populations with limited genetic segregation to detect rare alleles with large effect, which limits the number of markers detected. (Nordborg & Weigel, 2008, Talukder *et al.,* 2019). In contrast, GWAS involves a panel or collection of lines, capitalizing on the extensive recombination events that have occurred over time, resulting in a larger number of markers detected with higher resolution (Nordborg & Weigel, 2008, Myles *et al.,* 2009). In both LM and GWAS, genetic signals that exceed a certain threshold are considered statistically significant, while other marker-trait associations are typically excluded from further analysis. Consequently, in traditional MAS, only a small number of markers are used for each trait because of the strict selection criteria used to identify significant association. There are several challenges with MAS. One is that markers detected with GWAS and LM tend to show larger effects than what is observed in breeding populations. The second is a risk for linkage drag from unwanted QTL and phenotypes. Thirdly, multiple testing requires very high significance thresholds in GWAS, resulting in that only the largest QTL are detected. Fourthly, the detected QTL may track only a limited fraction of the total genetic variance. If the QTL effects are too small as with FHB resistance in oat, there is a high risk that the effect of the remaining unselected resistance QTL will overshadow the effect selected for a single QTL.

### 1.4.4. Genomic selection

Genomic Selection (GS) on the other hand, is a methodology that uses genotypic information from the entire genome and combines it with phenotypic data to make predictions for these traits (Meuwissen *et al.,* 2001). GS holds great potential for improved genetic gain in cereal breeding (Desta & Ortiz, 2014, Crossa *et al.,* 2017). This prediction relies on a statistical model trained using data from a group of individuals with both phenotypic and genotypic information, known as the "training population". This group is used to estimate random effect of individual markers based on a prior distribution, which are then used to calculate the genomic estimated breeding values (GEBV) for selection candidates (i.e., breeding lines) with only genotypic data. There are different kinds of models that have different assumed distributions of marker effects (Meher *et al.,* 2022). Some models can detect non-additive marker effects (Raffo *et al.,* 2022), such as genotype by environment interactions (Zhang *et al.,* 2021). Covariates like associated marker information, correlated traits, or weather parameters can be added to improve the models (Tolhurst *et al.,* 2022). Accuracy of genomic prediction depends on: (i) the number of markers that are well-distributed across the genome, (ii) precision of phenotypic data, (iii) heritability of the traits, (iv) number of individuals in the training population that is genetically and phenotypically diverse and closely related to the target population, (v) genome size (bigger genomes imply more effects to estimate), and, (vi) size of LD blocks along the genome (smaller LD blocks imply more independent effects to estimate). A reduction in price of genotyping in recent years has made GS more available to breeding companies (Huang *et al.,* 2014). Continuous efforts in both cost-effective genotyping and high throughput phenotyping are essential for widespread adoption of this powerful breeding approach. A previous study indicated that GS was favourable over MAS for FHB resistance in wheat (Arruda *et al.,* 2016), which is also assumed in oat given similar genetic complexity with multiple small effect QTL and high G x E (He *et al.,* 2013).

### 1.4.5. Genomic selection in oat

In the early 2000s, genetic research on oats utilized various kind of marker systems, including Diversity array technology (DArT), Amplified Fragment Length Polymorphism (AFLP) and Short Sequence Repeat (SSR) markers (He *et al.,* 2012). These marker systems, while effective for genomic prediction in oats (Asoro *et al.,* 2011), were often costly to develop or did not detect sufficient number of polymorphisms across the genome. In contrast, other crops like wheat developed and used marker assays for single nucleotide polymorphisms (SNP) that were cost-effective and exhibited large number of polymorphisms that were abundant within the genome (Wang *et al.,* 2014). The first commercially available SNP-chip in oat was developed in 2014, containing 4,975 polymorphic markers (Tinker *et al.,* 2014). This SNP-chip was primarily used for gene mapping and, to a limited extent, for

genomic selection (Rio *et al.,* 2021, Brzozowski *et al.,* 2023). The limited adoption of the SNP-chip in oats was likely due to insufficient marker counts for the large and complex oat genome, as well as its relatively higher cost compared to crops like wheat and barley. As a response, breeding companies in the Nordic countries developed an improved customized SNP-chip featuring 18,598 markers (20K SNP Chip), which proved valuable for genomic prediction (Haikka *et al.,* 2020a, Haikka *et al.,* 2020b, Sørensen *et al.,* 2023).

### 1.4.6. Genomic selection to enhance Fusarium resistance

Genomic prediction for resistance to *F. graminearum* have been used successfully in wheat (Rutotski *et al.,* 2012) and barley (Lorenz *et al.,* 2012), where prediction accuracies above 0.6 were shown. Different publications have demonstrated improved prediction accuracy through various strategies, including increasing training population size (Lorenz *et al.,* 2012), inclusion of correlated traits like days to heading and plant height (Schulthess *et al.,* 2018), use of significant marker information from association studies (Rutotski *et al.,* 2012, Alemu *et al.,* 2023), and using models that include environmental effects (Zhang *et al.,* 2021). In case of oats, there is one notable publication on GS for resistance to *F. graminearum* in oat (Haikka *et al.,* 2020b). This study reported mean prediction accuracies of 0.31 for DON, 0.47 for Fusarium infected kernels and 0.59 for germination capacity. The study employed a standard Genomic Best Linear Unbiased Prediction (GBLUP) model. These results suggest that achieving moderate accuracy in FHB resistance prediction in oats is possible, and the choice of phenotype for prediction may play a significant role in the process. The quantitative nature of FHB resistance, availability of accurate field testing and analysis, and reduced costs of genotyping makes GS a suitable option for FHB resistance breeding.

# 2. Thesis

## 2.1. Background and main objective

FHB has significant implications across all sectors of the cereal industry and especially for oat production in Norway. Over the past two decades, oat crops have faced major epidemics. The reduction in both yield and quality directly affects the income for Norwegian farmers. Elevated mycotoxin levels make oat seed unsuitable for both human and animal consumption, which increases the reliance on import. The reduced germination rates in infected seed also pose substantial challenges for the seed companies and the overall oat production sector, as imported varieties may not be optimally adapted to Norwegian conditions. In a global context with scarce food resources and unpredictable geopolitical dynamics affecting imports, it becomes crucial for Norway to maintain its capacity to produce its own food. Oats are well-suited to the Norwegian climate and play a pivotal role in achieving this goal.

The Norwegian oat community has taken a comprehensive approach to tackle FHB in many ways, with active participation from government to farmers alike. This approach has entailed an array of measures including increased surveillance on mycotoxins from the food and safety department to ensure a safer food supply. Funding for research related to FHB detection, treatments and genetic research has also seen a significant increase, advancing our understanding of the disease. The Norwegian Plant Variety Board has increased focus on FHB resistance when evaluating new oat varieties for the domestic market. The plant breeding company Graminor has increased its efforts in testing and research on FHB, resulting in some of the most FHB resistant varieties in Europe. Despite these advancements, there is still room for increased FHB resistance. Furthermore, as the markets are shifting towards new specialized oat products such as oat milk, meat substitutes and even beer it will be increasingly more important to maintain high FHB resistance as a breeding goal, as these products are vulnerable to mycotoxin contamination (Miró-Abella *et al.,* 2017, Pascari *et al.,* 2018).

Genomic selection serves as a cost-efficient tool in this task, offering the potential to predict FHB resistance at earlier breeding stages, to eliminate the most susceptible breeding lines. However, predicting FHB resistance presents a challenging task due to its quantitative nature that is characterised by small effect loci and large genotype by environment interactions. The overall objective of this PhD thesis is to develop predictive models for successful FHB resistance assessment in Graminor's oat breeding material. This objective will be achieved through a structured approach involving three work packages, each resulting in subsequent research papers:

**Paper 1:** The first paper aimed to evaluate different known optimization strategies used for optimization of training populations for genomic prediction, resulting in an optimized training population for further analysis in Paper 2 and 3. A large and diverse germplasm called oat breeding panel was used to ensure the representation of various genetic backgrounds, and phenotypic data were collected from Graminor's database. The evaluation was done on a validation population of breeding lines, and was based on prediction ability, genetic diversity, genetic similarity with the validation population and phenotypic variance of the selected training populations.

**Paper 2:** The second paper aimed to identify loci significantly associated with FHB resistance. This was done by assessing the training population developed based on findings in research Paper 1 in spawn inoculated and mist irrigated disease trials over two locations and three years, analyse samples for DON content and germination percentage and undertake a genome-wide association mapping.

**Paper 3:** The third paper aimed to apply data from the training population and the most important FHB resistance loci identified in Paper 2 and evaluate the performance of genomic prediction that include haplotype information as fixed effects. The models were evaluated as cross-validation in the training population and on three validation populations. In addition, a selection was done with each model of the predicted elite material to validate the expected effect of the models on observed phenotypes and allele frequencies of the resistant haplotype alleles.

## 2.2. Germplasm and phenotyping

### 2.2.1. Oat breeding panel and Validation population (Paper 1)

A comprehensive oat breeding panel was developed, consisting of 1,124 varieties and breeding lines that were tested as part of the RESIFUS project during the period between 2014 to 2018. Within this panel, the majority of lines (86%) were breeding lines and varieties from Graminor while the rest of the material were from Sweden (8%), Netherlands (2%), Germany (2%) and Finland (1%). The remaining lines from America, Australia, Austria, Canada, Denmark, Poland and Slovakia contributed less than 1% each. The primary objective of this larger population was to serve as pool for selecting different sets of optimized training populations, which were subsequently used in genomic prediction. Three different strategies were evaluated for their predictability on breeding material. Phenotypic data for the oat breeding panel were collected from all available yield trials and fusarium trials from 2014 to 2018 for each line accordingly. Phenotypes gathered for each trial were number of days from sowing to heading (DTH), and plant height (PH). The adjusted mean values derived from each trial were used as input data in a generalized linear model, enabling the calculation of overall adjusted means accounting for the random effects of year and location. A testing population of 257

$F_{10}$ breeding lines was evaluated in five environmental conditions in 2019. This was used to gather prediction ability data in order to evaluate the strategies. All lines were genotyped with a Nordic customized SNP-chip containing 18,598 markers (20K SNP-chip). The genotypic data obtained were filtered for missing values (<10%) and minor allele frequency (MAF >5%) in the oat breeding panel using a customized script in R statistical software (R core Team, 2022) . The filtering process resulted in 3,022 polymorphic markers for subsequent analysis.

### 2.2.2. Association mapping-, Training- and Validation- populations (Paper 2 and 3)

A training population of 541 genotypes was selected from the oat breeding panel to be used both for GWAS (Paper 2) and genomic predictions (Paper 3). This population was chosen based on a combination of different optimization criteria like genetic diversity, phenotypic variance and genetic similarity with the breeding material. The effect of the significant markers from the GWAS study, and the prediction ability of genomic selection were evaluated within the training population and on three validation populations called V1, V2 and V3. V1 and V2 consisted of 242 $F_{10}$ breeding lines from Graminor from the years 2020 and 2021, respectively. $F_{10}$ lines were used as they had not yet been selected for fusarium resistance. V3 consisted of 230 lines where 112 were new $F_{10}$ lines from 2022 and the rest overlapped with the other validation populations. Of the 118 remaining lines 88 were from V2, 22 were from V1 and 8 were from the testing population used in Paper 1.

All material was sown in spawn inoculated and mist irrigated disease trials (Tekle *et al.,* 2018) at Vollebekk (59.66°N, 10.75°E) and Staur (60.73°N, 11.10°E) research stations between the years 2020 and 2022. The training population was tested in all six environments, while V1, V2 and V3 were tested in two environments each in 2020, 2021 and 2022 respectively. Phenotypes gathered for each trial were DTH, PH, DON accumulation (ppm) quantified with ELISA, and germination ability quantified as the percentage of healthy germinations following the ISTA protocols (International rules for seed testing, 2021). The marker data from Paper 1 were filtered for missing values (<10%) and minor allele frequency based (MAF >5%) on the new training population which resulted in 3071 polymorphic markers. Of these 2928 were the same as filtered from the Oat breeding panel. The marker data for V1-3 were obtained using a different customized 7K-SNP chip (Polley *et al.,* 2023) containing 6,642 markers which are the 6587 most polymorphic markers from the Nordic 20K SNP-chip and 55 new SNP-markers.

## 2.3. Methodology

### 2.3.1. Optimization strategies

In Paper 1, the strategies for optimizing training populations for Paper 2 (GWAS) and Paper 3 (genomic prediction) were evaluated based on the selected training populations' ability to predict the phenotypes of the validation population described in 2.2.1. These training populations were also subjected to an analysis of their genetic diversity, phenotypic variance and genetic similarity in comparison to the testing population. The first strategy was called the *Diversity Core* and was based on the principle outlined by Franco *et al.* (2005). It is designed to maintain the same level of diversity and population structure in smaller subsamples as found in the larger candidate population. This strategy involves three key steps. In the first step a structure analysis is conducted on the candidate population to identify the most likely number of clusters that represent distinct genetic grouping. In the second step genetic distances between each line within these identified clusters are calculated using marker information. This step helps to quantify the genetic diversity within and between clusters. The final step involves performing 1000 random selections within each cluster. The sample with the highest average distance to the other genotypes in the cluster is selected for the training population. The number of genotypes selected from each cluster is proportional to the size and diversity of the respective cluster.

The second optimization strategy, called *Prediction Core*, was based on the methodologies outlined by Rincent *et al.* (2012) and Akdemir & Isidro-Sanchez (2019). This strategy is designed to minimize the prediction error variance (PEV) for the testing population while simultaneously maximising the genetic diversity of the training population for a selection criterion known as coefficient of determination (CD). The Prediction Core strategy involves two main steps. In step one, a principal component analysis is carried out on the marker information from both candidate and testing populations. The second step involves the use of the first 100 principal components as input in a selection algorithm. This algorithm calculates the PEV and CD values (Akdemir, 2018) which are the optimization criteria. The optimization process begins with the selection of a random sample, and the optimization criteria for the selections being calculated. Subsequently, one of the genotypes is replaced with another genotype that either increases the optimization criteria or does not decrease it. This replacement process is repeated and continues until no further optimization can be achieved.

The third optimization strategy called *Phenotypic selection* involves a straightforward approach: selecting an equal number of genotypes with the highest and lowest breeding values for each trait. Each of the strategy was replicated 20 times per population size of 100, 200, 300, 400 and 500, except for Phenotypic selection that was repeated once per trait and population. Predictions were

conducted with the BGLR package by Perez & Campos (2014) with the R statistical software (R Core Team, 2022). Prediction ability (PA) for each trait (PH and DTH) was calculated as the Pearson correlation between predicted and observed phenotypes. Simultaneous pairwise comparison tests (Minitab, 2021) were performed to determine whether these strategies yielded outcomes significantly different from random selection. The genetic diversity was measured as the expected heterozygosity for each locus, averaged across all loci. The genetic similarity between training and testing population was calculated as the proportion of shared alleles between the training and testing population.

### 2.3.2. Genome wide association mapping

In Paper 2, a GWAS was conducted for the traits germination percentage (GP) and DON accumulation (DON) for individual trials and overall values of the training population. To account for the effect of DTH and PH, a regression analysis was performed, treating GP and DON as separate response variables, and DTH and PH as explanatory variables. The GWAS was carried out using the "farmCPU" method developed by Liu *et al.* (2016) and implemented through the GAPIT3 package by Wang & Zhang (2021) with the R statistical software (R Core Team, 2022).

The farmCPU model was chosen because of its statistical power and its ability to adjust for the potentially confounded effect of kinship. Research, such as the study by Kaler *et al.* (2020), has shown that this model is less prone to false positives and false negatives compared to alternative models. In addition, it has proven to be the most efficient model for detecting resistance QTL for *F. graminearum* in wheat (Nannuru *et al.,* 2022). The analysis was not adjusted for population structure, as model selection analysis with GAPIT determined that zero principal components were the optimal choice. To identify significant markers, a false discovery rate of 0.05 was used as threshold (Benjamini & Hochberg, 1995).

Significant markers were grouped together into QTL-regions based on their linkage disequilibrium (LD) with each other. These regions are from this point called haplotypes and are represented as the combination of significant markers from the GWAS. The haplotypes were analysed using a pairwise t-test for each region to assess if the alleles exhibited significant differences from each other in terms of DON and GP. This was also extended to the three validation populations using their overall phenotypic data. Resistant and susceptible alleles were identified as those that most frequently exhibited significant lower or higher DON and GP than others within the GWAS population. The validated haplotypes were determined as those who showed consistent and significant effect on DON and GP between the resistant and susceptible alleles across environments and populations. The validated resistant alleles were further analysed through a stacking approach by grouping together

genotypes based on number of resistant haplotypes. This was done to investigate whether their combined presence led to an additive reduction in DON content and increase in GP.

### 2.3.3. Genomic prediction

In Paper 3 genomic prediction was evaluated by incorporating significant associated haplotypes identified in the GWAS (Paper 2) as fixed effects to enhance prediction accuracy with genomic prediction. Bayesian ridge regression (BRR) served as the basic model for estimating marker effects. The overall DON and GP values of the training population from the six environments tested were used to train the prediction model. Predictions were conducted with the BGLR package by Perez & Campos (2014) with the R statistical software (R Core Team, 2022). PA was acquired through 5-fold cross-validation for the training population, repeated 30 times, and once for each validation population (V1-3). PA was calculated as the Pearson correlation between predicted and observed phenotypes. Haplotype information was added to the basic model as a fixed effect, with a code 1 for resistance allele, -1 for susceptible alleles, and 0 for neither. In addition, two models were examined: one that evaluated the combined effects of the validated haplotypes, and another that considered all haplotypes from the GWAS. Within the training population, the 30 replications were subjected to pairwise t-test to determine whether they significantly differed from one another in terms of PA.

A breeding scenario was designed, based on the real breeding situation of Graminor, in which the top 30% of lines were chosen based on predicted DON values for further evaluation. These selected genotypes underwent analysis to assess their observed phenotypes, including DON, GP, DTH and PH. In addition, their resistant allele frequencies for all haplotypes identified in the GWAS were examined. This analysis served the dual purpose of confirming that the models performed as intended and identifying any unintended effects on the allele frequencies of different haplotypes. A selection based on observed DON values was done and analysed for allele frequency to see whether the resistant alleles increased as expected.

## 2.4. Main results and discussions

### 2.4.1. Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa* L.) (Paper 1)

The objective of this study was to evaluate various strategies for optimizing training populations for genomic prediction, including Prediction core, Diversity core, Phenotypic selection and Random selection. Among these strategies, Prediction core yielded the most favourable results, showing higher prediction accuracy than the others, particularly in larger populations for both traits (Figure 3). It was the only strategy to show a significantly higher PA compared to Random selection. It also

demonstrated the highest genetic diversity within populations ranging from 200 to 500 individuals (Figure 4A) and displayed a higher degree of genetic similarity to the testing population compared to most other strategies (Figure 4B). Furthermore, the Prediction core showed higher phenotypic variance compared to Diversity core and Random selection.

Diversity core, on the other hand, showed similar prediction accuracy to Random selection. However, selected training populations showed significantly higher genetic diversity especially in smaller populations. It is worth noting that despite the enhanced genetic diversity, the genetic similarity of the training and testing populations remained relatively low in all population sizes compared to the other strategies (Figure 4B). This suggests that Diversity core might not be an optimal strategy for the specific testing population used in the study, as it showed low genetic diversity and striving for high genetic diversity in the training population appeared to reduce genetic similarity between the training and testing populations.

Phenotypic selection also performed reasonably well and close to Prediction core for both traits. It yielded exceptionally high phenotypic variance, but, interestingly, the levels of genetic diversity and similarity varied between the two traits. In summary, two main criteria emerged as essential for optimizing our training populations in our dataset: population size and genetic similarity, both showed significant positive correlations with PA. Population size showed robust correlations, with r-values of 0.77 for DTH and 0.36 for PH. Genetic similarity showed modest and similar correlation, with r-values of 0.4 for DTH and 0.48 for PH. While genetic and phenotypic diversity had non-significant correlation to PA for DTH, they displayed significant correlation with r-values of 0.25 for both criteria for PH.

The Random selection showed the lowest PA of all strategies for both traits for population size 300-500. It also showed the lowest genetic diversity and phenotypic variance in all population sizes. It did however show genetic similarity slightly lower than the Prediction core strategy in all population sizes.

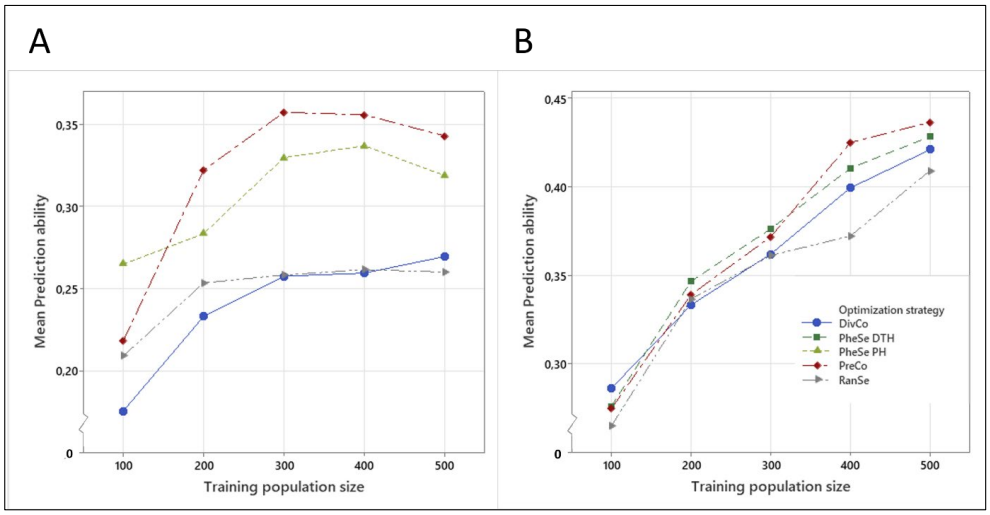



***Figure 3**: Average prediction abilities for (A) plant height (PH) and (B) days to heading (DTH) in oats for the different optimization strategies Prediction core (PreCo), Diversity core (DivCo), Phenotypic selection (PheSe) and Random selection (RanSe) across different training population sizes.*

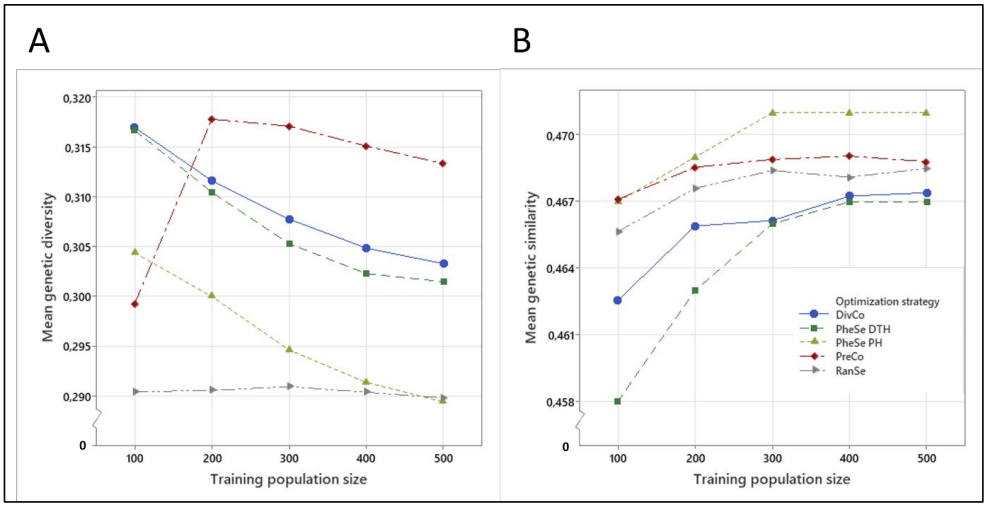



***Figure 4**: (A) Mean genetic diversity and (B) Mean genetic similarity for the optimization strategies Diversity core (DivCo), Prediction core (PreCo), Phenotypic selection (PheSe) for plant height (PH) and days to heading (DTH) and Random selection (RanSe) in population size 100-500.*

This research confirms, as previous studies have shown, that the genetic similarity between training and testing population is one of the most important factors to optimize for achieving high prediction ability (Lorenz & Nice, 2017, Zhang *et al.,* 2017, Xiaogang *et al.*,2018). Higher PA were found with larger population size for DTH (Figure 3B) which matches previous reports on effect of population size

(Asoro *et al.,* 2011, Zhang *et al.,* 2017). But for PH smaller population size gave higher PA (Figure 3A), which supports the idea that smaller training populations suffice for highly heritable traits (Zhang *et al.,* 2017, Kaler *et al.,* 2022) like PH and DTH. It has been shown in previous studies that smaller optimized training population could lead to higher PA compared to larger populations (Adeyemo *et al.,* 2020) which is supported by this study. Several studies have highlighted the importance of having high genetic diversity in the training population (Crossa *et al.,* 2016, Norman *et al.,* 2018, Fernández-González *et al.,* 2023). However, these training populations have been evaluated on diverse material, while this study evaluates optimization on breeding material known to be low in genetic diversity (He *et al.,* 2012). Thus, genetic diversity seems to be more important to increase in combination with increased genetic similarity. A combination of both Diversity Core and Prediction core could be advisable, as accounting for population structure and high genetic diversity would make the training population more robust over time (Berro *et al.,* 2019).

This study could have been improved further to better target some of the research questions. To increase the accuracy of phenotyping and make the study more reliable all candidates for the training population could have been sown together in the same trial, to get a balanced dataset. In addition, at least two more strategies should have been tested. One that maximises the genetic diversity alone, more than the Diversity core does. The second one would use the Prediction core algorithm to select only based on genetic similarity, without increasing the genetic diversity. Both of these strategies could have shed light on the true effect of maximising only genetic diversity and genetic similarity. The last improvement would be to use additional validation populations, to further investigate the stability of the training population over several generations of testing populations, an important aspect when applying genomic prediction (Neyhart *et al.,* 2017).

### 2.4.2. Identification of haplotypes associated with resistance to *Fusarium graminearum* in spring oat (*Avena sativa* L.) (Paper 2)

The GWAS analysis successfully identified a total of 48 significant markers associated with FHB resistance, which include 24 for DON, 22 for GP and 2 for both. 36 of these were clustered into a total of 15 QTL-regions where the markers were in significant LD with each other. The markers in each region were further analysed together as one haplotype. Five of the 15 haplotypes were validated through a pairwise t-test between the most frequently resistance and susceptibility alleles (Table 1). Validated haplotypes were located on linkage groups 1C, 7C-17A, 9D, 12D and 18D. The resistant alleles of these haplotypes showed a reduction of DON between 12 and 23% compared to the susceptible alleles for the overall values in the GWAS population, and increased GP between 1.6 and 4.2% (Table 1). The validation populations mostly showed similar effects, with some inconsistencies for region 1C in V2 and 18D-1 in V1. The QTL-region 7C-17A-1 showed the most pronounced effects in

both DON and GP, and these effects remained relatively consistent across different environmental conditions and populations. Similarly, the QTL region on 9D displayed consistent and highly significant reductions in DON across all tested environments and populations.

***Table* 1**: List of the five validated QTL-regions and the difference between the resistant and susceptible haplotypes in all years (2020, 2021 and 2022) and locations (V = Vollebekk, S = Staur) of the GWAS panel including the overall values (Ov), and the overall values of the validation populations (V1-V3). DON is shown as percentage difference between groups calculated as (difference in mean/average between groups) × 100 while GP is given as the difference in mean phenotypes. Positive values mean that the susceptible haplotype has higher DON and lower GP than the resistant. A < 0.05 = *, < 0.01 = ** and < 0.001 ***

| QTL | 20S | 20V | 21S | 21V | 22S | 22V | Ov | V1 | V2 | V3 |
|---|---|---|---|---|---|---|---|---|---|---|
| **DON** | | | | | | | | | | |
| **1C** | 8 | 5 | 8 | 6 | 24*** | 11* | 12* | 43 | -31 | 14 |
| **7C-17A-1** | 17 | 31* | 18 | 26 | 29* | 19* | 23** | 32 | 6 | 33*** |
| **9D** | 22*** | 19*** | 14** | 14** | 19*** | 12*** | 16*** | 16 | 17 | 15* |
| **12D** | 15 | 26** | 19* | 8 | 22** | 15** | 17*** | NA | 15 | 14 |
| **18D-1** | 9 | 9* | 11** | 15** | 14** | 13*** | 12*** | 2 | 16* | 14** |
| **GP** | | | | | | | | | | |
| **1C** | 1.8 | 0.1 | 1.1 | 1.1 | 6.9*** | 7.9*** | 3.1*** | 1.5 | -1.8 | 5.6 |
| **7C-17A-1** | 2.2 | 2.9 | 0.1 | 3.8* | 4.9 | 9.4*** | 4.2*** | 4.0 | 0.6 | 12.2 |
| **9D** | 2.9*** | 0.9 | 0.5 | -0.2 | 3.1* | 2.7** | 1.6*** | 0.5 | 2.0 | 4.6* |
| **12D** | 5.3*** | 0.6 | -0.1 | -0.7 | 3.7* | 4.2** | 2.1*** | NA | 2.7 | 5.2 |
| **18D-1** | 2.4** | 0.7 | 0.2 | 0.1 | 2.9** | 3.2*** | 1.6*** | -0.2 | 2.2* | 4.9** |

The cumulative impact of the five validated resistant alleles was found to have an additive effect on enhancing resistance (Figure 5a and b). When adding just one resistant haplotype allele, a significant reduction in DON and an increase in GP were observed. These improvements were evident up to a total of three alleles. Interestingly, there was no significant difference between the effects of having three alleles and having four, but a significant improvement was observed when moving from three to five alleles. This indicates that when a certain level of resistance is achieved it requires considerably more resistance sources to improve further, but that this level of resistance is not unattainable.

A BLAST search in the GrainGenes database (Yao *et al.,* 2022) against the OT3098 reference genome (PepsiCo, 2021) was conducted within the haplotypes using the marker flanking sequences which revealed several candidate genes (Paper 2, Supplementary Table S2). A total of 15 different genes were annotated with association to disease resistance. Many of these genes have been previously reported to be associated with resistance to *F. graminearum* in both bread wheat (Kugler *et al.,* 2013)

and durum wheat (Sari *et al.,* 2019). Furthermore, these genes are often recognized as components of resistance gene clusters. The QTL-regions on 1C, 7C-17A-1 and 12D supported the concept of gene clusters, as evidenced by the high number of resistance gene copies found in these regions: 20, 16 and 11 copies, receptively. An alternative explanation for the resistance observed in the 1C region relates to the potential presence of DON detoxification genes, which were detected and described by Khairullina *et al.* (2022) as they are situated approximately 4 Mbp from the closest marker in the region.

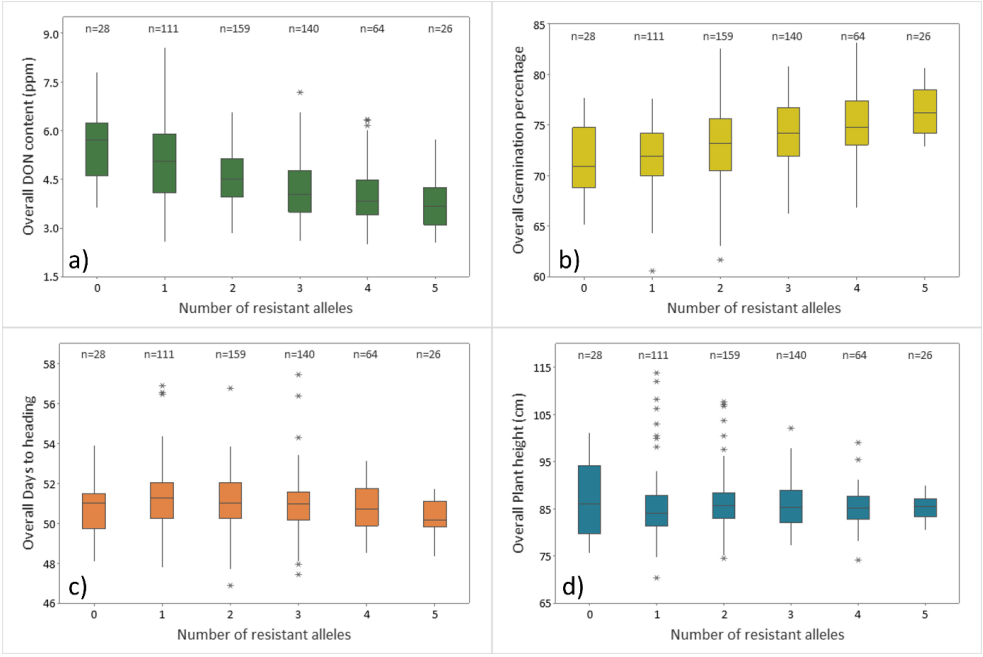



***Figure 5****: Boxplots of overall values from the GWAS population (Y-axis) for the traits DON accumulation in ppm (a), germination percentage (b), days to heading (c) and plant height (d) for genotypes grouped based on the number of validated resistant alleles they carry (X-axis).*

Three regions identified in this study (5C, 9D and 7C-1) have also been detected in previous studies (He *et al.,* 2013, Bjørnstad *et al.,* 2017) while the remaining are novel QTL. However, the BLAST search revealed that the significant markers from their studies are possibly 100-300 Mbp away from the closest significant marker in this study. The region on 7C-1 did not even match the same chromosome on the reference genome. This could be due to the frequent re-arrangement of this specific LG (Jellen *et al.,* 1994). It is worth noting that the genetic maps and populations used in those studies were based on mostly American accessions (Chaffin *et al.,* 2016). However, in our study the populations were a combination of Nordic breeding material and older Nordic varieties (Sørensen *et al.,* 2023),

and the map was based on the one developed by Chaffin *et al.* (2016) and updated with six biparental populations from the Nordic breeding companies. This could explain why the referenced positions do not match with our study. But the mapping done in this study is likely more relevant for further enhancing FHB resistance in the Nordic breeding programs. As of now, there is a Pan genome project in oat (Mascher, 2022) that can help in accurate physical mapping of the markers and shed light on the true dispersion of markers in the genome. It could also result in the detection of more markers, which is highly needed for both mapping and genomic prediction. Fine mapping and functional studies of the QTL-regions discovered in this study would be a major breakthrough for FHB resistance research in oat and contribute to incorporation of durable FHB resistance in breeding material.

There are several things that could have made this GWAS study better. Firstly, ensuring high disease pressure in all experiments would have made the phenotypes more reliable, especially for GP in V1 and V2. Secondly, using other models like BLINK for association mapping could have resulted in better marker detection as it has been described as more statistically powerful than farmCPU (Zhiwu Zhang Laboratory, 2023). Thirdly, add more FHB traits like (1) counting number of infected kernels, (2) Near-Infrared Hyperspectral Imaging, (3) degree of anther extrusion and (4) qPCR for quantification of Fusarium DNA. This could have revealed more regions and more of the genetic characterization of the underlying genes. Finally, a more genetically diverse population should have been used with higher number of polymorphic markers, as the training population was primarily optimized for genomic prediction. This could have mapped markers closer to the causal QTL as LD would have been smaller.

### 2.4.3. Evaluation of genomic prediction models for FHB resistance in oat (*Avena sativa* L.) with models that include significant haplotype information (Paper 3)

The five validated QTL-regions from the GWAS in Paper 2 were added as fixed effects in the genomic prediction model. These markers were included both individually (model M1C, M7C-1, M9D, M12D and M18D-1), as well as collectively (model M5HT) and in combination with the other ten resistance alleles from the GWAS (model M15HT). The goal was to assess whether these markers-based models could improve PA compared to the Bayesian Ridge Regression (BRR) model in both the training population and validation populations. BRR was chosen as the basic model because it assumes that most markers have a small or close to zero effect on the phenotype, and that using a model that assumes some markers to have larger effect would undermine the purpose of adding significant marker information as fixed effect. For the training population, the results demonstrated higher PA for all single haplotype models for both traits, with the exception of M18D-1 (Figure 6). The model with the with the five GWAS-validated haplotypes (M5HT) showed significantly higher PA than any single haplotype model for both traits, while including all fifteen haplotypes (M15HT) showed

significantly higher PA than M5HT for both traits in the training population. However, the results were more variable in the three validation populations (V1, V2 and V3) (Figure 7), and consistent effect on PA across different populations and traits was not always observed. Model M1C showed small differences with BRR with the highest increase in V1. M7C-1 showed consistent increase in PA in V1 and V3 for both DON and GP. M9D showed increased PA in all population for DON (Figure 7a), but in none for GP (Figure 7b). M12D showed reduced PA in all populations and traits. M18D-1 showed almost no difference in PA with BRR for DON (Figure 7a) but increased it for GP in all populations (Figure 7b). M5HT showed the highest PA in V1 for DON and V3 for GP, but reduced PA in V2 and V3 for DON and V2 for GP. Furthermore, M15HT consistently showed notably lower PA in all populations for both traits. These results show that the performance of these marker-based models varies between the training and validation populations, and the choice of which markers to include in the model may depend on the specific trait and population of interest.

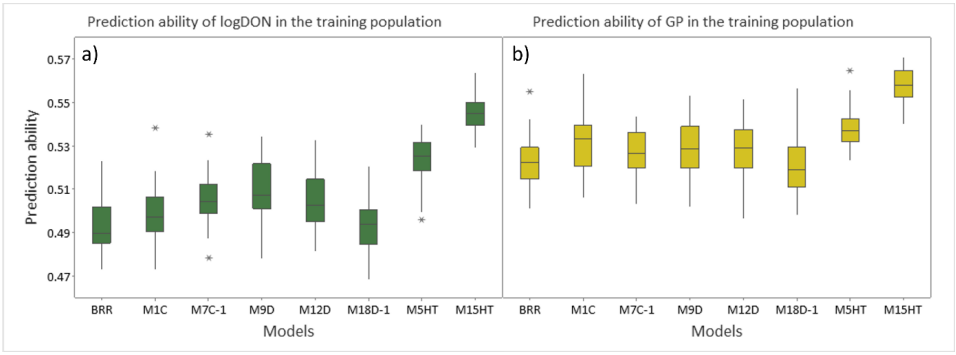



***Figure 6:*** *Prediction ability of log-transformed DON values (logDON; a) and Germination percentage (GP; b) in the training population using 5-fold cross-validation. Each model was replicated 30 times.*

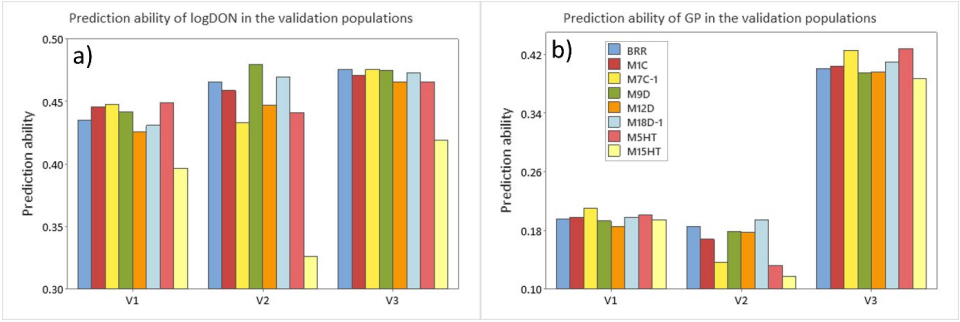



***Figure 7****: Prediction ability of logDON (a) and GP (b) in the three validation populations for the different models with prediction ability on the y axis and populations on the x-axis (V1, V2 and V3). The different models are indicated by the colours specified in the figure.*

In the selection process, the models consistently reduced DON and increased GP in the training population, with increased effect with higher PA (Figure 8). There was no significant difference in observed DTH in selection. But for PH there were significantly higher plants in selection with M12D, M5HT and M15HT than with BRR with the tallest plant being observed with M15HT. However, the opposite was observed in the validation populations, as M15HT consistently resulted in the shortest plants (Paper 3, supplementary Figure S1). This could indicate a non-random linkage between the QTL-regions and PH that could differ between the training and validation populations, which in case would partly explain the difference in PA between the training and validation population for M15HT. When assessing observed DON content in the validation populations, a consistent trend emerged (Paper 3, Figure 4). Selections made with the M5HT showed higher observed DON content, and this effect was even more profound in selections with the M15HT model. This pattern contradicts the findings in the training population (Figure 8a and b), as the prediction that increased PA with M5HT also showed higher DON values in selections. The M12D model was the only single haplotype model that increased DON content in the validation population V1 which matches the lower PA (Figure 7a). The M7C-1 and M18D-1 models consistently reduced DON content in selections for validation populations V2 and V3. However, this reduction in DON content was not reflected in an increase in PA for these models. In some instances, their PA remained similar to that of the BRR model, and in the case of M7C-1 in V2, the PA was lower. These observations underscore the complex interplay between model selection, traits, and populations, and shows that increased PA does not always result in better observed phenotypes in elite material.

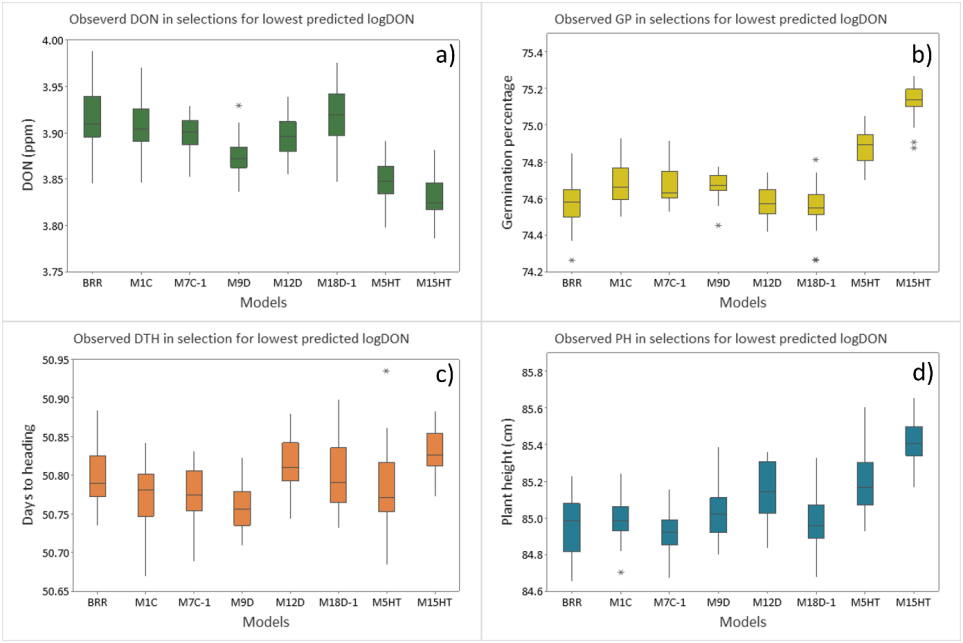



***Figure 8**: Average observed phenotypes of DON (a), Germination percentage (GP; b), Days to heading (DTH; c) and Plant height (PH; d) in populations selected based on 1/3 of genotypes with the lowest predicted logDON with each model in the 30 cross-validation replicates.*

A plausible explanation for the varying effects between the models and between populations is that when amplifying an assumed allele effect in the model it reduces the effect of other important QTL. This could lead to the unintended reduction of resistance alleles from other regions in selected elite material as highlighted in Figure 9. There it is shown that along with the significant increase of resistant alleles of targeted QTL-regions, there is also a significant reduction in other resistance alleles in different regions as with M1C, M9D, M12D and M18D-1. The inclusion of these effects is further amplified in the M5HT model. When using the M15HT model, the frequencies of most alleles increase, but not all. Notably, alleles in regions 5C-1 and 20D still show a significant reduction compared to BRR, despite their effects being included in the model. It is important to note that the changes in allele frequencies in the validation populations were inconsistent (Paper 3, Table 5). But the resistant alleles of the target regions always increased which showed that the models worked as intended in the validation populations as well. A striking difference between the training and validation populations is the change in allele frequencies on region 21D in selections. Initially, there is a significant increase with M5HT and M15HT in the training population, but in the validation populations, a consistent decrease is evident. This finding suggests that the QTL-region on 21D might play a more significant role than previously recorded in the GWAS. Another notable result from

analysing the validation population is that specifically in V3 the frequency of the resistant alleles of 12D and 18D-1 were reduced in alle single haplotype model, except for the ones where they were specifically targeted.

| HT | M1C | M7C-1 | M9D | M12D | M18D-1 | M5HT | M15HT |
|---|---|---|---|---|---|---|---|
| 1C | *** | | | ** | | *** | *** |
| 2C | | | | | | | *** |
| 5C-1 | | | | | | *** | *** |
| 5C-2 | | | | | | | *** |
| 6C | | | | | | ** | |
| 7C-1 | | *** | | | | *** | *** |
| 7C-2 | | | | | ** | *** | *** |
| 9D | | | *** | | | *** | *** |
| 12D | ** | | | *** | | *** | * |
| 15A | | | | | | | *** |
| 18D-1 | | | | | *** | *** | |
| 18D-2 | | | | | | ** | |
| 19A | * | | | | | *** | *** |
| 20D | | | ** | | | | *** |
| 21D | | | *** | | | *** | *** |


***Figure 9****: Difference between the fixed effect models and the base model BRR in frequencies of resistance haplotypes (HT) within populations selected based on the 1/3 lowest predicted DON values. The number are rounded up to the closest second decimal, and empty cells means that the difference was less than 0.005. p-values < 0.05 = *, 0.01 = ** and 0.001 = ***.*

Previous studies that include fixed effect of significant QTL for FHB resistance report increased prediction accuracy (Rutkoski *et al.,* 2012, Zhang *et al.,* 2021, Alemu *et al.,* 2023). In most cases several QTL are added simultaneously, which matches the results found in this study within the training population as more QTL-regions significantly increased PA. This shows the added benefit of adding fixed effects of QTL with large effect as well as the small effect. However, the results from the validation population reveal that this effect can be inconsistent and potentially have negative impact on PA. Some studies have pointed out that there are potential risks of adding fixed effects to models, especially when the effects are relatively small (Poland & Rutkoski, 2016, Herter *et al.,* 2019), as with Fusarium resistance in oat (He *et al.,* 2013, Bjørnstad *et al.,* 2017).

There are several potential explanations to the difference between the training and validation populations. The most important ones to consider are the relationship between the training and validation populations, difference in LD between population, gene interactions, the environmental effects and disease pressure. When there is high LD between two important QTL-regions, it means that their alleles appear together more often than expected based on their individual frequencies.

And if the alleles that segregate are resistant on one locus, and non-resistant on the other it would probably not increase PA as expected. When LD is high across chromosomes it is likely caused by family structure (Nei & Li, 1973) or epistasis (Lewontin & Kojima, 1960). Ideally the QTL-regions used as fixed effects in GS models should not be in high LD with each other to avoid adding the effects of non-resistant allele through LD with targeted QTL-regions. LD between known QTL-regions should therefore be analysed in the breeding population before performing genomic prediction with fixed effects of QTL-regions. Another way that LD affects the results is that the LD between the markers and causal QTL detected in the GWAS could change due to recombination in the breeding material, making the assumed effect wrong and therefore ineffective in increasing PA. The effect of gene interaction and epistasis could have substantial impact on PA (Wientjes *et al.,* 2022). This has been described for FHB resistance in wheat (Kage *et al.,* 2017), where transcription factors of some genes affect the expression of others. The expression of certain genes could also have been affected by environmental effects as described in disease resistance in *Arabidopsis thaliana* (Macqueen & Bergelson, 2016) where different validation populations were tested in separate years. A large variation in disease pressure between the years, has shown to affect gene expression of FHB resistance in wheat (Manghwar *et al.,* 2021), which could also be the case in our study. When phenotypic selection was performed, most of the validated alleles increased in frequency (Paper 3, Figure 5), but not for all validation populations. This indicate that there might be a difference in QTL effect between the populations and years.

The PA found in this study were moderately high for DON compared to another study on FHB resistance in oat (Haikka *et al.,* 2020b). Similar PA were also shown across all three validation populations. The general complexity of the oat genome, and especially FHB resistance makes it a trait difficult to predict. Four ways that could improve PA based on the current approach of added effects of significant markers are proposed. First is to differentiate the effects of the alleles as the GWAS showed that each region and allele contributes different effects on resistance. This could be done by not using the presumed effects of resistant and susceptible haplotype allele that we have from the GWAS, and just use the marker information. This would make the model able to calculate the effects of other marker combinations as well as the ones included as haplotypes in this study. A second is to use a model including epistatic effects, which could capture the potential gene interaction for the trait (Raffo *et al.,* 2022). Third is to analyse the LD between the potential QTL to include and use only the ones that have the lowest LD with other known QTL-regions. This could reduce the risk of co-segregation between resistance and susceptible alleles of different loci. A fourth is to do more precise mapping of the causal QTL, and locating markers very close them. This would reduce the risk of loss of LD between markers and QTL due to segregation in new breeding populations.

Other aspects of improving PA include the following: (1) using models that include GxE interaction like "RKHS" (Hu *et al.,* 2023) as G x E is shown to have a significant effect on DON (He *et al.,* 2013) and GP (Table 1), (2) use multi-trait models that include effects of DTH and PH to improve PA (Gaire *et al.,* 2022, Zhang *et al.,* 2022) as they have shown to correlate with both DON and GP (Figure 9), and (3) to increase the number of SNP markers used, as the approximately 3000 markers used in this study do not necessarily capture all small effect-QTL in the large oat genome. All improvements mentioned above could be used together in a multi-trait-multi-environment model (Gill *et al.,* 2021) with fixed effect of significant markers.

### 2.4.4. Fixed effects of QTL-regions and evaluation of GS

The effect of adding haplotype information as fixed effect were expected to increase PA in general and the frequency of only the targeted resistant allele in selections of elite material. But the significant reduction of other QTL alleles as a side effect as with 12D with M1C and 1C with M12D (Figure 9) poses the question of how the QTL-regions are linked. If there is a linkage between the resistant allele of 1C and non-resistant alleles of 12D and vice versa, then that could be evident in the entire population and thus possible to account for when selecting markers for model improvement. Another explanation could be that the effects of some of the resistant alleles are overestimated in the BRR model, thus their effects are reduced when fixed effect of other alleles are added. This is indicated by the fact that there is a bigger increase of all five validated resistant alleles in selection with BRR than with phenotypic selection in the training population (Paper 3, Figure 5). As PA is calculated as the correlation between the predicted and observed phenotypes it could be argued that to achieve 100% accuracy of the models, the allele frequency in selections should be the same with BRR as with phenotypic selection (Figure 5), thus the reduction of effects of non-targeted QTL with the single haplotype models could also increase PA. Another aspect of this is that the higher frequency of resistant alleles in selection with BRR over phenotypic selection indicates an increased selection intensity for these regions with GS, thus increasing genetic gain faster. A drawback of this is that alleles can get fixated faster in the breeding material, or that rare alleles are lost due to genetic drift, both of which reduce the genetic diversity of the breeding material, potentially hampering future genetic gains with the same material (Wientjes *et al.,* 2023).

### 2.4.5. Implementation of genomic prediction for FHB resistance

The consistent PA found in this study makes it possible to apply the models in the Norwegian breeding program. However, per now the PA is not high enough to be given as much weight as phenotypic data. Thus, we recommend confirming resistances in phenotypic trials before official variety trials. A second benefit of this is that data from field trials could be used to improve the

prediction models by adding new lines to the training population. Four options for implementation at Graminor based on the current breeding program as described in 1.3.4. can be suggested. First is to make genomic selection in the $F_{10}$ generation as has been done previously with phenotypic selection. This would be cheap as there are few lines to genotype, and costs normally spent on phenotyping can be saved. Second is to make selection based on GS in $F_9$. The number of lines in this generation is almost 10 times as high as $F_{10}$ which increases the price for genotyping but would increase selection intensity for FHB resistance. The PA does not merit selecting the top 10% with high confidence, but could be used to discard the most susceptible, and then do phenotypic selection for other agronomic traits on the rest. A third application would be to make selection in $F_8$ where we have lines in single rows. Here are 10 times more genotypes than in $F_9$ making genotyping even more expensive, and there are less traits to phenotype as they are sown in single rows. Here the model could be used to discard the most susceptible lines as there are many potentially favourable traits that is not observed, or to use models for other agronomical traits as well as FHB resistance. Finally, if the models improve it could be used to genotype single panicles in earlier generations and use them as parents for new crossings. After a certain number of crossing cycles, the progeny is advanced to pure lines (Gaynor *et al.,* 2017). This is a method of rapid improvements of genetic gains but could be considered unreliable as prediction are made on still segregating lines. This approach would be more applicable with more accurate models, and for more traits, which would reduce the risk of unfavourable genes and phenotypes accumulating in the germplasm. The use of fixed effects of QTL could be useful even if they do not improve PA in that it does increase the frequency of desirable alleles in the selections. It can be viewed as a form of MAS within the GS framework and would be especially useful for alleles that are in low frequency in the population as these could be lost in future generation due to genetic drift. But it would be advisable to only use markers with consistent large effect across populations as it is expected based on this study that added QTL would be selected in disfavour of others.

# Conclusion

This study has provided valuable insights into the optimization of training populations in oats, as well as the application of genomic prediction models and the identification of significant QTL-regions for FHB resistance:

**Training Population Optimization (Paper 1):** The study successfully evaluated different strategies for training population optimization and identified two main criteria for optimizing training populations: population size and genetic similarity between training and testing population. Genetic and phenotypic diversity were less important as criteria by themselves. But the results indicate that no single criterion improved PA by itself, but in combinations. The Prediction core strategy successfully balanced high genetic similarity between training and testing population with relatively high genetic and phenotypic diversity, resulting significantly higher PA than random selection with 0.09 higher average PA for PH in population size 300 and 400. The findings from this study resulted in an optimized training population for future genomic prediction and association studies in the context of FHB resistance.

**Genome-Wide Association Study (GWAS) of Fusarium resistance (Paper 2):** A total of 15 significant QTL-regions associated with Fusarium resistance in oats were detected. Out of these, five QTL regions exhibited consistent effects across different environments and populations. A haplotype analysis was conducted, enabling the identification of both resistant and susceptible alleles within each region. Each of the five resistant alleles showed between 12 and 23% reduction in DON content compared to the susceptible allele. The cumulative effects of these five validated QTL-regions were found to collectively reduce the DON content by 38%. In addition, the study identified several disease resistance genes within these regions that had previously been associated with resistance to *F. graminearum* in other crops. This research has affirmed that Fusarium resistance is a complex trait influenced by multiple QTL-regions across the genome, each with varying effects. The knowledge gained from this study is expected to assist breeders in making informed decisions when selecting parent plants for future crossings, thereby facilitating the combination of multiple resistant QTL regions. While the individual effects of the markers detected in the GWAS may be relatively small, this information can be leveraged to enhance genomic selection models for FHB resistance.

**Genomic Prediction Models for Fusarium resistance (Paper 3):** The integration of fixed effects using the resistant alleles identified in the GWAS improved the prediction accuracy for both DON and GP in the training population with average PA of 0.49 and 0.52 with the base model. Adding QTL information increased PA with the highest when all GWAS results were included resulting in average PA of 0.55 and 0.56 for DON and GP respectively. The PA in the validation populations were

moderately high with the base model when heritability of the trait was high, with PA for DON between 0.44 and 0.48. However, the individual effects of these resistant alleles in the validation populations were inconsistent with different QTL outperforming BRR in different traits and populations. The lowest PA was achieved when all QTL information was added, emphasizing that results from GWAS are not necessarily the same in breeding populations. It is therefore advisable to validate the effect of marker information in breeding material before application of markers in models. The reason for the inconsistency could possibly be due to differences in populations or environmental factors, given that the three populations were tested in different years. Moreover, the frequency of these resistant haplotype alleles revealed that enhancing the effect of specific QTL can downregulate the effect of other resistance alleles. This downregulating between populations is likely influenced by non-random linkage and family structure, and subsequent co-segregation of resistant and non-resistant alleles at separate loci. Additionally, it is quite possible that the LD between markers detected in the GWAS, and the causal QTL has been partially broken due to recombination in subsequent breeding populations. Therefore, it is advisable to thoroughly study the linkage between known resistance QTL-regions before integrating them into GS models to avoid unintentional negative effects on predictions.

High PA is achievable for FHB resistance in oat as long as there is a good relationship between the training and breeding material, the training population is large and diverse enough, the traits are correctly assessed with low error rate and high heritability, there is enough marker information to capture the small effects of resistance genes throughout the genome, and that the model used manages to capture the complex interplay of epistasis, G x E and population structure within breeding programs. Adding fixed effects of significant markers can improve PA, but the markers effects from a GWAS are not necessarily the same in the breeding population. More research into precise mapping and function of resistance genes would improve the reliability of this approach.

# References


Abrahamsen, U., Brodal, G. & Waalen, W. (2016) Virkning av ulike forgrøder på neste års avling av hvete. NIBIO BOK 2 (1) Jord- og Plantekultur 2016, 106-111. URL: https://nibio.brage.unit.no/nibio-xmlui/handle/11250/2577326.

Adeyemo, E., Bajgain, P., Conley, E. J., Sallam, A. H., & Anderson, J. A. (2020) Optimizing training population size and content to improve prediction accuracy of FHB-related traits in wheat. *Agronomy*, *10*(4), 543. https://doi.org/10.3390/agronomy10040543.

Akdermir, D. (2018) STPGA: Selection of Training Populations by Genetic Algorithm. R package version 5.2.1. https://CRAN.R-project.org/package=STPGA.

Akdemir, D. & Isidro-Sánchez, J. (2019). Design of training populations for selective phenotyping in genomic prediction. *Scientific reports* 9(1446). https://doi.org/10.1038/s41598-018-38081-6.

Alemu, A., Batista, L., Singh, P. K., Ceplitis, A., & Chawade, A. (2023) Haplotype-tagged SNPs improve genomic prediction accuracy for Fusarium head blight resistance and yield-related traits in wheat. *Theoretical and applied genetics*, *136*(4), 92. https://doi.org/10.1007/s00122-023-04352-8.

Arata, G, J. Martínez, M., Elguezábal, C., Rojas, D., Cristos, D., Dinolfo, M. I. & Arata, A. F. (2022) Effects of sowing date, nitrogen fertilization, and *Fusarium graminearum* in an Argentinean bread wheat: Integrated analysis of disease parameters, mycotoxin contamination, grain quality, and seed deterioration. *Journal of Food Composition and Analysis, 107*, 104364. https://doi.org/10.1016/j.jfca.2021.104364.

Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C. Brown-Guedira, G. Dong, Y., Foresman, B. J & Kolb, F. L. (2016) Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding, 36*, 84. https://doi.org/10.1007/s11032-016-0508-5.

Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P. & Jannink, J-L. (2011) Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. *The Plant Genome, 4*(2). https://doi.org/10.3835/plantgenome2011.02.0007.

Audenaert, K., Vanheule, A., Höfte, M., & Haesaert, G. (2013) Deoxynivalenol: a major player in the multifaceted response of Fusarium to its environment. *Toxins*, *6*(1), 1–19. https://doi.org/10.3390/toxins6010001.

Aydin, E. & Cocmen, D. (2011) Cooking quality and sensorial properties of noodle supplemented with oat flour. *Food Science and biotechnology, 20***,** 507-511. https://doi.org/10.1007/s10068-011-0070-1.

Ball, J. J., Wyatt, R. P., Coursen, M. M., Lambert, B. D., & Sawyer, J. T. (2021) Meat Substitution with Oat Protein Can Improve Ground Beef Patty Characteristics. *Foods, 10*(12), 3071. https://doi.org/10.3390/foods10123071.

Basinskiene, L. & Cizeikiene, D. (2020) Cereal-Based Nonalcoholic Beverages. In C. M. Galanakis, Ed. *Trends in Non-alcoholic Beverages*, 63-99. Academic Press. https://doi.org/10.1016/B978-0-12-816938-4.00003-3.

Becher, R., Hettwer, U., Karlovsky, P., Deising, H. B., & Wirsel, S. G. (2010) Adaptation of Fusarium graminearum to tebuconazole yielded descendants diverging for levels of fitness, fungicide resistance, virulence, and mycotoxin production. *Phytopathology*, *100*(5), 444–453. https://doi.org/10.1094/PHYTO-100-5-0444.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bernhoft, A., Clasen, P. E., Kristoffersen, A. B., & Torp, M. (2010) Less Fusarium infestation and mycotoxin contamination in organic than in conventional cereals. *Food additives & contaminants. Part A, Chemistry, analysis, control, exposure & risk assessment*, *27*(6), 842–852. https://doi.org/10.1080/19440041003645761.

Bernhoft, A. Eriksen, G. S., Sundheim, L., Berntssen, M., Brantsæter, A. L., Brodal, G., Fæste, C. K., Hofgaard, I. S., Rafoss, T., Sivertsen, T. & Tronsmo, A. M. (2013) Risk Assessment of Mycotoxins. in Cereal Grain in Norway. *Opinion of the Scientific Steering Committee of the Norwegian Scientific Committee for Food Safety; VKM (Norwegian Scientific Committee for Food Safety):* Oslo, Norway, 2013. URL: https://vkm.no/english/riskassessments/allpublications/riskassessmentofmycotoxinsincerealgraininnorway.4.27ef9ca915e07938c3b2bad8.html.

Bernhoft, A., Christensen, E., Tukun, F-L., Gilbert, D., & Er, J. C. & Eriksen, G. (2023) The surveillance programme for feed and feed materials for terrestrial animals in Norway 2022, *Norwegian Veterinary Institute.* URL: https://www.mattilsynet.no/dyr_og_dyrehold/for/foranalyser_2022__mykotoksiner_og_sopp.50042/binary/F%C3%B4ranalyser%202022%20-%20mykotoksiner%20og%20sopp.

Berro, I., Lado, B., Nalin, R. S., Quincke, M., & Gutiérrez, L. (2019) Training Population Optimization for Genomic Selection. *The plant genome*, *12*(3), 1–14. https://doi.org/10.3835/plantgenome2019.04.0028.

Bjørnstad, Å. & Skinnes, H. (2008) Resistance to *Fusarium* infection in oats (*Avena sativa* L.). *Cereal research communications 36 (Suppl 6),* 57–62. https://doi.org/10.1556/CRC.36.2008.Suppl.B.9.

Bjørnstad, Å., He, X., Tekle, S., Klos, K., Huang, Y-F., Tinker, N. A., Dong, Y. & Skinnes, H. (2017) Genetic variation and associations involving Fusarium head blight and deoxynivalenol accumulation in cultivated oat (*Avena sativa* L.). *Plant Breeding*, *136*, 620– 636. https://doi.org/10.1111/pbr.12502.

Blandino, M., Reyneri, A., Colombari, G. & Pietri, A. (2009) Comparison of integrated field programmes for the reduction of fumonisin contamination in maize kernels. *Field Crops Research*, *111*, 284–289. https://doi.org/10.1016/j.fcr.2009.01.004.

Brodal, G., Aamot, H. U., Almvik, M., & Hofgaard, I. S. (2020) Removal of Small Kernels Reduces the Content of *Fusarium* Mycotoxins in Oat Grain. *Toxins*, *12*(5), 346. https://doi.org/10.3390/toxins12050346.

Brzozowski, L. J., Campbell, M. T., Hu, H., Yao, L., Caffe, M., Gutiérrez, L., Smith, K. P., Sorrells, M. E., Gore, M. A., & Jannink, J-L. (2023) Genomic prediction of seed nutritional traits in biparental families of oat (*Avena sativa*). *The Plant Genome*, 00, e20370. https://doi.org/10.1002/tpg2.20370.

Cerrudo, D., Cao, S., Yuan, Y., Martinez, C., Suarez, E. A., Babu, R., Zhang, X., & Trachsel, S. (2018) Genomic Selection Outperforms Marker Assisted Selection for Grain Yield and Physiological Traits in a Maize Doubled Haploid Population Across Water Treatments. *Frontiers in plant science*, *9*, 366. https://doi.org/10.3389/fpls.2018.00366.

Chaffin, A. S., Huang, Y.-F., Smith, S., Bekele, W. A., Babiker, E., Gnanesh, B. N., Foresman, B. J., Blanchard, S. G., Jay, J. J., Reid, R. W., Wight, C. P., Chao, S., Oliver, R., Islamovic, E., Kolb, F. L., McCartney, C., Mitchell Fetch, J. W., Beattie, A. D., Bjørnstad, Å., Bonman, J. M., Langdon, T., Howarth,

C. J., Brouwer, C. R., Jellen, E. N., Klos, K. E., Poland, J. A., Hsieh, T.-F., Brown, R., Jackson, E., Schlueter, J. A. & Tinker, N. A. (2016) A Consensus Map in Cultivated Hexaploid Oat Reveals Conserved Grass Synteny with Substantial Subgenome Rearrangement. *The Plant Genome, 9*: plantgenome2015.10.0102. https://doi.org/10.3835/plantgenome2015.10.0102.

de Chaves, M.A., Reginatto, P., da Costa, B.S., de Paschoal, R. I., Teixeira, M. L. & Fuentefria, A. M. (2022) Fungicide Resistance in *Fusarium graminearum* Species Complex. *Current Microbiology* 79, 62. https://doi.org/10.1007/s00284-021-02759-4.

Crossa, J., Jarquín, D., Franco, J., Pérez-Rodríguez, P., Burgueño, J., Saint-Pierre, C., Vikram, P., Sansaloni, C., Petroli, C., Akdemir, D., Sneller, C., Reynolds, M., Tattaris, M., Payne, T., Guzman, C., Peña, R. J., Wenzl, P. & Singh, S. (2016) Genomic Prediction of Gene Bank Wheat Landraces. *G3: Genes, Genomes, Genetics, 6*(7), 1819-1834. https://doi.org/10.1534/g3.116.029637.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de Los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017) Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in plant science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011.

Desta, Z. A. & Ortiz, R. (2014) Genomic selection: genome-wide prediction of plant improvement. Trends in Plant Science, 19(9). https://doi.org/10.1016/j.tplants.2014.05.006.

Dong, F., Chen, X., Lei, X., Wu, D., Zhang, Y., Lee, Y. W., Mokoena, M. P., Olaniran, A. O., Li, Y., Shen, G., Liu, X., Xu, J. H., & Shi, J. R. (2023) Effect of Crop Rotation on *Fusarium* Mycotoxins and *Fusarium* Species in Cereals in Sichuan Province (China). *Plant disease*, *107*(4), 1060–1066. https://doi.org/10.1094/PDIS-01-22-0024-RE.

Elen, O., Liu, W., Langseth, W., H Skinnes, H., Gullord, M. & Sungheim, L. (2003), Deoxynivalenol Content of Cereal Grain from Naturally Infected and Artificially Inoculated Plants in Field Trials in Norway, *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science, 53*:4, 183-189. https://doi.org/10.1080/09064710310015229.

European commission. (2006) Commission regulation (EC) No 1881/2006 of 19 December 2006. Setting maximum levels for certain contaminants in foodstuffs. Annex section 2: Mycotoxins, 2.4. Deoxynivalenol, 2.4.2. Unprocessed durum wheat and oats. URL: http://data.europa.eu/eli/reg/2006/1881/2023-01-01.

European Food Safety Authority (EFSA), Arcella, D., Gergelova, P., Innocenti, M. L. & Steinkellner, H. (2017) Human and animal dietary exposure to T-2 and HT-2 toxin. *EFSA Journal*, *15*, 57. https://doi.org/10.2903/j.efsa.2017.4972.

Felleskjøpet Agri SA. (2016) Prognose for tilgang og forbruk av norsk korn for sesongen 2016/2017. URL: https://www.fk.no/Media/Files/prognose-for-tilgang-og-forbruk-av-norsk-korn-for-sesongen-2016-2017.

Felleskjøpet Agri SA. (2023) Plantevernkatalogen 2023, Page 26. URL: https://ipaper.ipapercms.dk/TidRom/felleskjopet/plantevern/.

Fernández-González, J., Akdemir, D. & Isidro y Sánchez, J. (2023) A comparison of methods for training population optimization in genomic selection. *Theoretical and Applied Genetics, 136*, 30. https://doi.org/10.1007/s00122-023-04265-6.

Franco, J., Crossa, J., Taba, S. & Shands, H. A. (2005) Sampling strategy for conserving genetic diversity when forming core subsets. *Crop Science. 45*(3), 1035–1044. https://doi.org/10.2135/cropsci2004.0292.

Gaire, R., de Arruda, M. P., Mohammadi, M., Brown-Guedira, G., Kolb, F. L., & Rutkoski, J. (2022) Multi-trait genomic selection can increase selection accuracy for deoxynivalenol accumulation resulting from fusarium head blight in wheat. *The plant genome*, *15*(1), e20188. https://doi.org/10.1002/tpg2.20188.

Gaynor, R. C., Gorjanc, G., Bentley, A. R., Ober, E. S., Howell, P., Jackson, R., Mackay, I. J. & Hickey, J. M. (2017) A Two-Part Strategy for Using Genomic Selection to Develop Inbred Lines. *Crop Science, 57*: 2372-2386. https://doi.org/10.2135/cropsci2016.09.0742.

Geddes, J., Eudes, F., Tucker, J. R., Legge, W. G. & Selinger, L. B. (2008) Evaluation of inoculation methods on infection and deoxynivalenol production by *Fusarium graminearum* on barley, *Canadian Journal of Plant Pathology, 30*:1, 66-73. https://doi.org/10.1080/07060660809507497.

Haikka, H., Knürr, T., Manninen, O., Pietilä, L., Isolahti, M., Teperi, E., Mäntysaari, E. A. & Strandèn, I. (2020) Genomic prediction of grain yield in commercial Finnish oat (*Avena sativa*) and barley (*Hordeum vulgare*) breeding programmes. *Plant Breeding*,139, 550-561. https://doi.org/10.1111/pbr.12807.

Haikka, H., Manninen, O., Hautsalo, J., Pietilä, L., Jalli, M. & Veteläinen, M. (2020) Genome-wide Association Study and Genomic Prediction for *Fusarium graminearum* Resistance Traits in Nordic Oat (*Avena sativa* L.). Agronomy*, 10*(2), 174. https://doi.org/10.3390/agronomy10020174.

Hautsalo, J., Jalli, M., Manninen, O. & Veteläinen M. (2018) Evaluation of resistance to *Fusarium graminearum* in oats. *Euphytica, 214,* 139. https://doi.org/10.1007/s10681-018-2222-3.

Havrlentová, M., Šliková, S., Gregusová, V., Kovácsová, B., Lančaričová, A., Nemeček, P., Hendrichová, J., & Hozlár, P. (2021) The Influence of Artificial *Fusarium* Infection on Oat Grain Quality. *Microorganisms*, *9*(10), 2108. https://doi.org/10.3390/microorganisms9102108.

He, X., & Bjørnstad, Å. (2012) Diversity of North European oat analyzed by SSR, AFLP and DArT markers. *Theoretical and applied genetics*, *125*(1), 57–70. https://doi.org/10.1007/s00122-012-1816-8.

He, X. Skinnes, H. Oliver, R. E., Jackson, E. W. & Bjørnstad, Å. (2013) Linkage mapping and identification of QTL affecting deoxynivalenol (DON) content (*Fusarium* resistance) in oats (*Avena sativa L.*). *Theoretical and Applied Genetics, 126*, 2655-2670. https://doi.org/10.1007/s00122-013-2163-0.

Herter, C. P., Ebmeyer, E., Kollers, S., Korzun, V., Würschum, T., & Miedaner, T. (2019) Accuracy of within- and among-family genomic prediction for Fusarium head blight and Septoria tritici blotch in winter wheat. *Theoretical and applied genetics*, *132*(4), 1121–1135. https://doi.org/10.1007/s00122-018-3264-6.

Hofgaard, I., Aamot, H., Torp, T., Jestoi, M., Lattanzio, V., Klemsdal, S., Waalwijk, C., Van der Lee, T., & Brodal, G. (2016) Associations between *Fusarium* species and mycotoxins in oats and spring wheat from farmers' fields in Norway over a six-year period. *World Mycotoxin Journal, 9*, 365–378. https://doi.org/10.3920/WMJ2015.2003.

Hu, X., Carver, B. F., El-Kassaby, Y. A., Zhu, L. & Chen, C. (2023) Weighted kernels improve multi-environment genomic prediction. *Heredity,* 130, 82–91. https://doi.org/10.1038/s41437-022-00582-6.

Huang, Y. F., Poland, J. A., Wight, C. P., Jackson, E. W., Tinker, N. A. (2014) Using Genotyping-By-Sequencing (GBS) for Genomic Discovery in Cultivated Oat. *PLOS ONE, 9*(7): e102448. https://doi.org/10.1371/journal.pone.0102448.

International Rules for Seed Testing, Chapter 5, (2021) i–5-56 (64)

Jellen, E. N., Gill, B. S. & Cox. T. S. (1994) Genomic in situ hybridization differentiates between A/D- and C-genome chromatin and detects intergenomic translocations in polyploid oat species (genus *Avena*). *Genome***,** *37*(4): 613-618. https://doi.org/10.1139/g94-087.

Jellen, E. N., Rines, H. W., Fox, S. L., Davis, D. W., Phillips, R. L. & Gill, B. S. (1997) Characterization of 'Sun II' oat monosomics through C-banding and identification of eight new 'Sun II' monosomics. *Theoretical and Applied Genetics*, *95,* 1190–1195. https://doi.org/10.1007/s001220050680.

Ji, C., Fan, Y. & Zhao, L. (2016) Review on biological degradation of mycotoxins. *Animal Nutrition*, *2*, 127–133. https://doi.org/10.1016/j.aninu.2016.07.003.

Jørstad, I. (1945) Parasittsoppene på kultur- og nyttevekster i Norge. I. Sekksporesopper (Ascomycetes) og konidiesopper (Fungi imperfecti). *Tillegg C til Landbruksdirektørens melding for 1943. Melding fra Statens plantepatologiske institutt. Nr. I.*, *50*, 1–142.

Kage, U., Yogendra, K. & Kushalappa, A. (2017) TaWRKY70 transcription factor in wheat *QTL-2DL* regulates downstream metabolite biosynthetic genes to resist *Fusarium graminearum* infection spread within spike. *Scientific Reports, 7*, 42596. https://doi.org/10.1038/srep42596.

Kaler, A. S., Gillman, J.D., Beissinger, T. & Purcell, L.C. (2020) Comparing different statistical models and multiple testing corrections for association mapping in Soybean and Maize. *Frontier of Plant Science, 10*, 1794. https://doi.org/10.3389/fpls.2019.01794.

Kaler, A. S., Purcell, L. C., Beissinger, T. & Gillman, J. D. (2022) Genomic prediction models for traits differing in heritability for soybean, rice, and maize. *BMC Plant Biology, 22***,** 87. https://doi.org/10.1186/s12870-022-03479-y.

Kamal, N., Renhuldt, N. T., Bentzer, J. *et al.* (2022) The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature, 606*, 113–119. https://doi.org/10.1038/s41586-022-04732-y.

Kang, R., Li, R., Dai, P., Li, Z., Li, Y., & Li, C. (2019) Deoxynivalenol induced apoptosis and inflammation of IPEC-J2 cells by promoting ROS production. *Environmental pollution*, *251*, 689–698. https://doi.org/10.1016/j.envpol.2019.05.026.

Kant, P., Reinprecht, Y., Martin, C. J., Islam, R., & Pauls, K. P. (2017) *Comprehensive Biotechnology* (M. Moo-Young, Ed.) (3rd ed., pp. 789-805). Pergamon. https://doi.org/10.1016/B978-0-12-809633-8.09244-X.

Keller, M. D., Bergstrom, G. C., Shields & E. J. (2014) The aerobiology of Fusarium graminearum. *Aerobiologia, 30*, 123–136. https://doi.org/10.1007/s10453-013-9321-3.

Khairullina, A., Tsardakas Renhuldt, N., Wiesenberger, G., Bentzer, J., Collinge, D. B., Adam, G., & Bülow, L. (2022) Identification and Functional Characterisation of Two Oat UDP-Glucosyltransferases Involved in Deoxynivalenol Detoxification. *Toxins*, *14*(7), 446. https://doi.org/10.3390/toxins14070446.

Kigoni, M., Choi, M. & Arbelaez, J.D. (2023) 'Single-Seed-SpeedBulks:' a protocol that combines 'speed breeding' with a cost-efficient modified single-seed descent method for rapid-generation-advancement in oat (*Avena sativa* L.). *Plant Methods, 19*, 92. https://doi.org/10.1186/s13007-023-01067-1.

Kosiak, B., Torp, M. & Thrane, U. (1997) The occurrence of Fusarium SPP. In Norwegian Grain – A survey. *Cereal research communications, 25,* 595-596. https://doi.org/10.1007/BF03543787.

Kosiak, B. Torp, M., Skjerve, E. & Thrane, U. (2003) The Prevalence and Distribution of *Fusarium* species in Norwegian Cereals: a Survey. *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science, 53*:4, 168-176. https://doi.org./10.1080/09064710310018118.

Kugler, K. G., Siegwart, G., Nussbaumer, T. Ametz, C., Spannagl, M., Steiner, B., Lemmens, M., Mayer, K. F. X., Buerstmayr, H. & Schweiger, W. (2013) Quantitative trait loci-dependent analysis of a gene co-expression network associated with Fusarium head blight resistance in bread wheat (*Triticum aestivum*L.). *BMC Genomics, 14*, 728. https://doi.org/10.1186/1471-2164-14-728.

Ladizinsky, G. A. (1998) A new species of Avena from Siciliy, possibly the tetraploid progenitor of hexaploidy oats. *Genetic Resources and Crop Evolution, 45*, 263-269. https://doi.org/10.1023/A:1008657530466.

Langseth, W. & Elen, O. (1997) The occurrence of deoxynivalenol in Norwegian cereals—differences between years and districts, 1988–1996. *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science, 47(3)*, 176-184. https://doi.org/10.1080/09064719709362458.

Leplat, J., Friberg, H., Abid, M & Steinberg, C. (2013) Survival of *Fusarium graminearum*, the causal agent of Fusarium head blight. *Agronomy of Sustainable Development, 33*, 97-111. https://doi.org/10.1007/s13593-012-0098-5.

Lewontin, R. C. & Kojima, K. (1960) The Evolutionary Dynamics of Complex Polymorphisms. *Evolution*, *14*(4), 458–472. https://doi.org/10.2307/2405995.

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016) Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS genetics*, *12*(2), e1005767. https://doi.org/10.1371/journal.pgen.1005767.

Liu, Q., Li, X., Zhou, X., Li, M., Zhang, F., Schwarzacher, T., & Heslop-Harrison, J. S. (2019) The repetitive DNA landscape in Avena (Poaceae): chromosome and genome evolution defined by major repeat classes in whole-genome sequence reads. *BMC plant biology*, *19*(1), 226. https://doi.org/10.1186/s12870-019-1769-z.

Lorenz, A. J., Smith, K. P. and Jannink, J-L. (2012) Potential and Optimization of Genomic Selection for Fusarium Head Blight Resistance in Six-Row Barley. *Crop Science, 52*, 1609-1621. https://doi.org/10.2135/cropsci2011.09.0503.

Lorenz A. & Nice L. (2017) Training Population Design and Resource Allocation for Genomic Selection in Plant Breeding. In: Varshney R., Roorkiwal M., Sorrells M. (eds) *Genomic Selection for Crop Improvement. Springer, Cham.* https://doi.org/10.1007/978-3-319-63170-7_2.

MacQueen, A. & Bergelson, J. (2016) Modulation of R-gene expression across environments. *Journal of experimental botany*, *67*(7), 2093–2105. https://doi.org/10.1093/jxb/erv530.

Manghwar, H., Hussain, A., Ali, Q., Saleem, M. H., Abualreesh, M. H., Alatawi, A., Ali, S., Munis, M. F. H. (2021) Disease Severity, Resistance Analysis, and Expression Profiling of Pathogenesis-Related

Protein Genes after the Inoculation of *Fusarium equiseti* in Wheat. *Agronomy, 11*(11):2124. https://doi.org/10.3390/agronomy11112124.

Mascher, M. (2022) Towards a pan-genome of hexaploid oat (abstract). *International Oat Conference OAT 2022 Program book,* page 57. URL: https://issuu.com/internationaloat/docs/oat2022_program_book.

Matelionienė, N., Supronienė, S., Shamshitov, A., Zavtrikovienė, E., Janavičienė, S. & Kadžienė, G. (2022) Weeds in Cereal Crop Rotations May Host *Fusarium* Species That Cause Fusarium Head Blight and Grain Weight Losses in Wheat. Agronomy*, 12(*11):2741. https://doi.org/10.3390/agronomy12112741.

Mathews, R., Kamil, A., & Chu, Y. (2020) Global review of heart health claims for oat beta-glucan products. *Nutrition reviews*, *78*(Suppl 1), 78–97. https://doi.org/10.1093/nutrit/nuz069.

Meher, P. K., Rustgi, S. & Kumar, A. (2022) Performance of Bayesian and BLUP alphabets for genomic prediction: analysis, comparison and results. *Heredity 128*, 519–530. https://doi.org/10.1038/s41437-022-00539-9.

Merrick, L. F., Burke, A. B., Chen, X., & Carter, A. H. (2021) Breeding With Major and Minor Genes: Genomic Selection for Quantitative Disease Resistance. *Frontiers in plant science*, *12*, 713667. https://doi.org/10.3389/fpls.2021.713667.

Mesterházy, A. (1995) Types and components of resistance to *Fusarium* head blight of wheat. *Plant Breeding, 114*: 377-386. https://doi.org/10.1111/j.1439-0523.1995.tb00816.x.

Mesterházy, Á., Bartók, T., Mirocha, C.G. and Komoróczy, R. (1999) Nature of wheat resistance to *Fusarium* head blight and the role of deoxynivalenol for breeding. *Plant Breeding, 118*: 97-110. https://doi.org/10.1046/j.1439-0523.1999.118002097.x.

Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819.

Miedaner, T., Moldovan, M., & Ittu, M. (2003) Comparison of spray and point inoculation to assess resistance to fusarium head blight in a multienvironment wheat trial. *Phytopathology*, *93*(9), 1068–1072. https://doi.org/10.1094/PHYTO.2003.93.9.1068.

Mielniczuk, E. & Skwaryło-Bednarz, B. (2020) Fusarium Head Blight, Mycotoxins and Strategies for Their Reduction. *Agronomy*, *10*, 509. https://doi.org/10.3390/agronomy10040509.

Miller, J. D., Young, J. C., & Sampson, D. R. (1985) Deoxynivalenol and Fusarium Head Blight Resistance in Spring Cereals. Journal of Phytopathology, 113: 359-367. https://doi.org/10.1111/j.1439-0434.1985.tb04837.x.

Minitab, LLC. (2021). *Minitab*. Retrieved from https://www.minitab.com

Miró-Abella, E., Herrero, P., Canela, N., Arola, L., Borrull, F., Ras, R., & Fontanals, N. (2017) Determination of mycotoxins in plant-based beverages using QuEChERS and liquid chromatography-tandem mass spectrometry. *Food chemistry*, *229*, 366–372. https://doi.org/10.1016/j.foodchem.2017.02.078.

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., & Buckler, E. S. (2009) Association mapping: critical considerations shift from genotyping to experimental design. *The Plant cell*, *21*(8), 2194–2202. https://doi.org/10.1105/tpc.109.068437.

Nannuru, V. K. R., Windju, S. S., Belova, T., Dieseth, J. A., Alsheikh, M., Dong, Y., McCartney, C. A., Henriques, M. A., Buerstmayr, H., Michel, S., Meuwissen, T. H. E., & Lillemo, M. (2022) Genetic architecture of fusarium head blight disease resistance and associated traits in Nordic spring wheat. *Theoretical and applied genetics*, *135*(7), 2247–2263. https://doi.org/10.1007/s00122-022-04109-9.

Nei, M. & Li, W. H. (1973) Linkage disequilibrium in subdivided populations. *Genetics*, *75*(1), 213–219. https://doi.org/10.1093/genetics/75.1.213.

Neyhart, J. L., Tiede, T., Lorenz, A. J., & Smith, K. P. (2017) Evaluating Methods of Updating Training Data in Long-Term Genomewide Selection. *G3*, *7*(5), 1499–1510. https://doi.org/10.1534/g3.117.040550.

Nordborg, M. & Weigel, D. (2008) Next-generation genetics in plants. Nature 456, 720-723. https://doi.org/10.1038/nature07629.

Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018) Optimising Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population Structure on Prediction Accuracy. *G3, 8*(9), 2889–2899. https://doi.org/10.1534/g3.118.200311.

Oliver, R. E., Tinker, N. A., Lazo, G. R., Chao, S., Jellen, E. N., *et al.* (2013) SNP Discovery and Chromosome Anchoring Provide the First Physically-Anchored Hexaploid Oat Map and Reveal Synteny with Model Species. *PLOS ONE, 8*(3): e58068. https://doi.org/10.1371/journal.pone.0058068.

Opplysningskontoret for brød og korn. (2022, August 3) *Fakta om havre*. Brød og Korn. Retrieved October 7, 2023. From https://brodogkorn.no/oppskrift_tema/havre/.

Pascari, X., Ramos, A. J., Marín, S. & Sanchis, V. (2018) Mycotoxins and beer. Impact of beer production process on mycotoxin contamination. A review. Food Research International, 103: 121-129. https://doi.org/10.1016/J.FOODRES.2017.07.038.

Paul, P. A., El-Allaf, S. M., Lipps, P. E., & Madden, L. V. (2004) Rain Splash Dispersal of Gibberella zeae Within Wheat Canopies in Ohio. *Phytopathology*, *94*(12), 1342–1349. https://doi.org/10.1094/PHYTO.2004.94.12.1342.

Peng, Y., Yan, H., Guo, L. *et al.* (2022) Reference genome assemblies reveal the origin and evolution of allohexaploid oat. *Nature Genetics, 54*, 1248–1258. https://doi.org/10.1038/s41588-022-01127-7.

PepsiCo. (2021). OT3098 v2 Hexaploid Oat. URL: https://wheat.pw.usda.gov/jb?data=/ggds/oat-ot3098v2-pepsico.

Perez, P., & Campos, G., (2014) Genome-Wide Regression and Prediction with the BGLR Statistical Package. *Genetics, 198*(2), 483-495. https://doi.org/10.1534/genetics.114.164442.

Pino, J. L., Mujica, V. & Arredondo, M. (2021) Effect of dietary supplementation with oat β-glucan for 3 months in subjects with type 2 diabetes: A randomized, double-blind, controlled clinical trial. *Journal of Functional Foods*, 77, 104331. https://doi.org/10.1016/j.jff.2020.104311.

Piñeros-Guerrero, N., Maldonado-Archila, G. and Gómez-Caro, S. (2019) Effect of thermal and in vitro fungicide treatments on pathogens of the genus Fusarium associated with maize seeds. *Agronomía Colombiana*, *37*(3), 228–238. https://doi.org/10.15446/agron.colomb.v37n3.80302.

Podgórska-Kryszczuk, I., Solarska, E., & Kordowska-Wiater, M. (2022) Biological Control of *Fusarium culmorum*, *Fusarium graminearum* and *Fusarium poae* by Antagonistic Yeasts. *Pathogens 11*(1), 86. https://doi.org/10.3390/pathogens11010086.

Poland, J., & Rutkoski, J. (2016) Advances and Challenges in Genomic Selection for Disease Resistance. *Annual review of phytopathology*, *54*, 79–98. https://doi.org/10.1146/annurev-phyto-080615-100056.

Polley, A., Plieske, J., Grafarend-Belau, E., Manninen, O., Vetelainen, M., Ceplitis, A., Vallenback, P., Alsheikh, M., Gnad, H. & Ganal, M. W. (2023) A new 7K SNP genotyping array for oat. *Oat newsletter, 60:4.* URL: https://oatnews.org/oatnews_pdfs/2020/oatnews_2023_Ganal.pdf.

Portell, X., Verheecke-Vaessen, C., Torrelles-Ràfales, R., Medina, A., Otten, W., Magan, N., & García-Cela, E. (2020) Three-Dimensional Study of *F. graminearum* Colonisation of Stored Wheat: Post-Harvest Growth Patterns, Dry Matter Losses and Mycotoxin Contamination. *Microorganisms*, *8*(8), 1170. https://doi.org/10.3390/microorganisms8081170.

R Core Team. (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Raffo, M. A., Sarup, P., Guo, X., Liu, H., Andersen, J. R., Orabi, J., Jahoor, A., & Jensen, J. (2022) Improvement of genomic prediction in advanced wheat breeding lines by including additive-by-additive epistasis. *Theoretical and applied genetics,135*(3), 965–978. https://doi.org/10.1007/s00122-021-04009-4.

Rafique, H., Dong, R., Wang, X., Alim, A., Aadil, R. M., Li, L., Zou, L., & Hu, X. (2022) Dietary-Nutraceutical Properties of Oat Protein and Peptides. *Frontiers in nutrition*, *9,* 950400. https://doi.org/10.3389/fnut.2022.950400.

Rincent, R., Laloë, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., Rodríguez, V. M., Moreno-Gonzalez, J., Melchinger, A., Bauer, E., Schoen, CC., Meyer, N., Giauffret, C., Bauland, C., Jamin, P., Laborde, J., Monod, H., Flament, P., Charcosset, A. & Moreau, L. (2012) Maximizing the Reliability of Genomic Selection by Optimizing the Calibration Set of Reference Individuals: Comparison of Methods in Two Diverse Groups of Maize Inbreds (*Zea mays* L.). *Genetics, 192*(2), 715-728. https://doi.org/10.1534/genetics.112.141473.

Rio, S., Gallego-Sánchez, L., Montilla-Bascón, G., Canales, F. J., Isidro Y Sánchez, J., & Prats, E. (2021) Genomic prediction and training set optimization in a structured Mediterranean oat population. *Theoretical and applied genetics, 134*(11), 3595–3609. https://doi.org/10.1007/s00122-021-03916-w.

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J-L. & Sorrells, M. (2012) Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *The Plant Genome, 5*. https://doi.org/10.3835/plantgenome2012.02.0001.

Salgado, J. D., Madden, L. V., & Paul, P. A. (2015) Quantifying the effects of fusarium head blight on grain yield and test weight in soft red winter wheat. *Phytopathology*, *105*(3), 295–306. https://doi.org/10.1094/PHYTO-08-14-0215-R.

Sari, E., Cabral, A. L., Polley, B., Tan, Y., Hsueh, E., Konkin, D. J., Knox, R. E., Ruan, Y., & Fobert, P. R. (2019) Weighted gene co-expression network analysis unveils gene networks associated with the Fusarium head blight resistance in tetraploid wheat. *BMC genomics*, *20*(1), 925. https://doi.org/10.1186/s12864-019-6161-8.

Schroeder, H. W. & Christensen, J. J. (1963) Factors Affecting Resistance of Wheat to Scab Caused by Gibberella zeae. *Phytopathology, 53*, 831-838.

Schulthess, A. W., Zhao, Y., Longin, C. F. H. & Reif, J. C. (2018) Advantages and limitations of multiple-trait genomic prediction for *Fusarium* head blight severity in hybrid wheat (*Triticum*

*aestivum* L.). *Theoretical and Applied Genetics*, *131*, 685–701. https://doi.org/10.1007/s00122-017-3029-7.

Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., & Kizek, R. (2010) Deoxynivalenol and its toxicity. *Interdisciplinary toxicology*, *3*(3), 94–99. https://doi.org/10.2478/v10102-010-0019-x.

Statistisk sentralbyrå, Statistics Norway (2023, February 21) *04607: Areal av korn- og oljevekster (F) 2001 - 2021*. ssb.no. Retrieved October 7, 2023. From https://www.ssb.no/statbank/table/04607/.

Statistisk sentralbyrå, Statistics Norway (2023, February 21) *07479: Kornavling (1 000 tonn) 1989 - 2022*. ssb.no. Retrieved October 7, 2023. From https://www.ssb.no/statbank/table/07479.

Statistisk sentralbyrå, Statistics Norway (2023, February 21) *07480: Kornavling per dekar (kg) 1989 - 2022*. Ssb.no. Retrieved October 7, 2023. From https://www.ssb.no/statbank/table/07480/.

Sundheim, L., Brodal, G., Hofgaard, I.S. & Rafoss T. (2013) Temporal Variation of Mycotoxin Producing Fungi in Norwegian Cereals. *Microorganisms*, *1*(1):188-198. https://doi.org/10.3390/microorganisms1010188.

Sørensen, E. S., Jansen, C., Windju, S., Crossa, J., Sonesson, A. K., Lillemo, M., & Alsheikh, M. (2023) Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa*). *Plant Breeding, 142*(1), 41–53. https://doi.org/10.1111/pbr.13061.

Talukder, Z. I., Ma, G., Hulke, B. S., Jan, C. C., & Qi, L. (2019) Linkage Mapping and Genome-Wide Association Studies of the *Rf* Gene Cluster in Sunflower (*Helianthus annuus* L.) and Their Distribution in World Sunflower Collections. *Frontiers in genetics*, *10*, 216. https://doi.org/10.3389/fgene.2019.00216.

Tekle, S. Skinnes H. & Bjørnstad, Å. (2013) The germination problem of oat seed lots affected by Fusarium head blight. *European Jorunal of Plant Pathology, 135,* 147-158. https://doi.org/10.1007/s10658-012-0074-6.

Tekle, S., Bjørnstad, Å., Skinnes, H., Dong, Y. & Segtnan, V. H. (2013) Estimating Deoxynivalenol Content of Ground Oats Using VIS-NIR Spectroscopy. *Cereal Chemistry, 90*, 181-185. https://doi.org/10.1094/CCHEM-07-12-0084-R.

Tekle, S., Måge, I., Segtnan, V. H. and Bjørnstad, Å. (2015) Near-Infrared Hyperspectral Imaging of *Fusarium*-Damaged Oats (*Avena sativa* L.). *Cereal Chemistry, 92*, 73-80. https://doi.org/10.1094/CCHEM-04-14-0074-R.

Tekle, S., Lillemo, M., Skinnes, H., Reitan, L., Buraas, T. & Bjørnstad, Å. (2018) Screening of Oat Accessions for Fusarium Head Blight Resistance Using Spawn-Inoculated Field Experiments. *Crop Science, 58*: 143-151. https://doi.org/10.2135/cropsci2017.04.0264.

Tekle, S., Stråbø, S. S., He, X., Dong, Y. & Bjørnstad Å. (2020) Variation in Anther Extrusion and Its Impact on *Fusarium* Head Blight and Deoxynivalenol Content in Oat (*Avena sativa* L.). *Agronomy*, *10*(3), 354. https://doi.org/10.3390/agronomy10030354.

The Norwegian Centre for Climate Services (NCCS). Hanssen-Bauer, I., Førland, E. J., Haddeland, I., Hisdal, H., Mayer, S., Nesje, A., Nilsen, J. E. Ø., Sandven, S., Sandø, A. B., Sorteberg, A. & Ådlandsvik, B. (2017) Climate in Norway 2100 – a knowledge base for climate adaptation. URL: https://www.miljodirektoratet.no/publikasjoner/2017/mai-2017/climate-in-norway-2100--a-knowledge-base-for-climate-adaptation/.

The research council of Norway. (2013) Safe grains: Mycotoxin prevention through resistant wheat and oats. URL: https://prosjektbanken.forskningsradet.no/en/project/FORISS/178273.

The research council of Norway. (2018) Resistenskilder, resistensmekanismer og seleksjonsmetoder mot Fusarium og mykotoksiner i havre. URL: https://prosjektbanken.forskningsradet.no/en/project/FORISS/233908.

Tinker, N. A., Chao, S., Lazo, G.R., Oliver, R. E., Huang, Y-F., Poland, J. A., Jellen, E. N., Maughan, P. J., Kilian, A. & Jackson, E. W. (2014) A SNP Genotyping Array for Hexaploid Oat. *The Plant Genome, 7*: plantgenome2014.03.0010. https://doi.org/10.3835/plantgenome2014.03.0010.

Tolhurst, D. J., Gaynor, R. C., Gardunia, B., Hickey, J. M., & Gorjanc, G. (2022) Genomic selection using random regressions on known and latent environmental covariates. *Theoretical and applied genetics*, *135*(10), 3393–3415. https://doi.org/10.1007/s00122-022-04186-w.

Trail, F. (2009) For Blighted Waves of Grain: *Fusarium graminearum* in the Postgenomics Era. Plant physiology, *149*, 103-10. https://doi.org/10.1104/pp.108.129684.

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., 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. & Akhunov, E. (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnology Journal, 12, 787-796. https://doi.org/10.1111/pbi.12183.

Wang, J. & Zhang, Z. (2021) GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics, Proteomics & Bioinformatics, 19(4)*, 629-640. https://doi.org/10.1016/j.gpb.2021.08.005.

Wientjes, Y. C. J., Bijma, P., Calus, M. P. L., Zwaan, B. J., Vitezica, Z. G. & Heuvel, van den J. (2022) The long-term effects of genomic selection: 1. Response to selection, additive genetic variance, and genetic architecture. *Genetics Selection Evolution 54*, 19. https://doi.org/10.1186/s12711-022-00709-7.

Wientjes, Y. C. J., Bijma, P., van den Heuvel, J., Zwaan, B. J., Vitezica, Z. G., & Calus, M. P. L. (2023) The long-term effects of genomic selection: 2. Changes in allele frequencies of causal loci and new mutations. *Genetics*, *225*(1), iyad141. https://doi.org/10.1093/genetics/iyad141.

Xiaogang, L., Hongwu, W., Hui, W., Zifeng, G., Xiaojie, X., Jiacheng, L., Shanhong, W., Wen-Xue, L., Cheng, Z., Boddupalli, M, P., Olsen, M, S., Changling, H., Yunbi, X. (2018) Factors affecting genomic selection revealed by empirical evidence in maize. *The Crop Journal, 6*(4), 341-352. https://doi.org/10.1016/j.cj.2018.03.005.

Yan, H., Martin, S. L., Bekele, W. A., Latta, R. G., Diederichsen, A., Peng, Y., & Tinker, N. A. (2016) Genome size variation in the genus Avena. *Genome*, *59*(3), 209–220. https://doi.org/10.1139/gen-2015-0132.

Yao, E., Blake, V. C, Cooper, L., Wight, C. P., Michel, S., Cagirici, H. B., Lazo, G. R., Birkett, C. L., Waring, D. J., Jannink, J-L., Holmes, I., Waters, A. J., Eickholt, D. P. & Sen, T. Z. (2022) "GrainGenes: A Data-Rich Repository for Small Grains Genetics and Genomics", Database, Volume 2022, baac034. https://doi.org/10.1093/database/baac034.

Zhang, A., Wang, H., Beyene,Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., Vicente, F. S., Olsen, M., Prasanna B. M., Crossa, J., Yu, H. & Zhang, X. (2017) Effect of Trait Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 bi-parental Tropical Maize Populations" *Frontiers of plant science, 8*. https://doi.org/10.3389/fpls.2017.01916.

Zhang, W., Boyle, K., Brule-Babel, A., Fedak, G., Gao, P., Djama, Z. R., Polley, B., Cuthbert, R., Randhawa, H., Graf, R., Jiang, F., Eudes, F., & Fobert, P. R. (2021) Evaluation of Genomic Prediction for Fusarium Head Blight Resistance with a Multi-Parental Population. *Biology*, *10*(8), 756. https://doi.org/10.3390/biology10080756.

Zhang, J., Gill, H. S., Brar, N. K., Halder, J., Ali, S., Liu, X., Bernardo, A., St. Amand, P., Bai, G., Gill, U. S., Turnipseed, B. & Sehgal, S. K. (2022) Genomic prediction of Fusarium head blight resistance in early stages using advanced breeding lines in hard winter wheat. *The Crop Journal*, 10(6), 1695-1704. https://doi.org/10.1016/j.cj.2022.03.010.

Zhiwu Zhang Laboratory, Washington State University. (2023) User manual for Genomic Association and Prediction Integrated Tool, Version 3. Page 14. URL: https://zzlab.net/GAPIT/gapit_help_document.pdf.

# Paper I

Received: 8 April 2022 | Revised: 23 September 2022 | Accepted: 20 October 2022

DOI: 10.1111/pbr.13061

ORIGINAL ARTICLE

Plant Breeding WILEY

# Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa*)

Espen Sannes Sørensen[1,2] | Constantin Jansen[1] | Susanne Windju[1] | Jose Crossa[4] | Anna Kristina Sonesson[3] | Morten Lillemo[2] | Muath Alsheikh[1,2]

[1]Graminor AS Ridabu, Ridabu, Norway

[2]Department of Plant Science, Norwegian University of Life Sciences, Ås, Norway

[3]Nofima AS, Ås, Norway

[4]International Maize and Wheat Improvement Center, Texcoco, Mexico

**Correspondence**
Espen Sannes Sørensen, Department of Plant Science, Norwegian University of Life Sciences, Ås, Norway.
Email: espen.sorensen@nmbu.no

**Funding information**
Research Council of Norway and Graminor, Grant/Award Number: 286606

## Abstract

Genomic selection is a promising breeding methodology that could increase selection accuracy and intensity and reduce generation interval. As the cost of genotyping decreases, it will be important to optimize training populations for costly phenotypic experiments for many complex traits. The aim of this research was to evaluate different optimization strategies, by using historical data from the Norwegian oat breeding programme at Graminor. In this paper, we focus on the optimization criteria: genetic diversity, phenotypic variance and genetic similarity between the training and testing populations. The four training population strategies—prediction core, diversity core, phenotypic selection and random selection—were applied to an oat candidate population of 1124 lines. An independent testing population was used to calculate the mean prediction abilities for the traits days to heading and plant height. Moreover, the strategies were tested in three independent wheat populations. The results showed that prediction core was the most promising strategy to select training populations with high genetic similarity to the testing set, high genetic diversity, and high phenotypic variance, which resulted in higher prediction ability across population sizes and traits.

**KEYWORDS**
genetic diversity, genetic similarity, optimization criteria, phenotypic variance, prediction ability, training population

## 1 | INTRODUCTION

The genetic gains per year of conventional breeding have been estimated to be 1% (Li et al., 2018). With the introduction of new molecular DNA-based technologies, breeders can increase selection accuracy and intensity and reduce the generation interval. This increased breeding efficiency (Heffner et al., 2010; Bhat et al., 2016; Xu et al., 2020) is key to increasing food production in the future. A promising marker-based breeding technique is genomic selection (GS; Meuwissen et al., 2001, Crossa et al., 2017, Wang et al., 2018), which uses whole-genome DNA markers and phenotypic information of a training population to predict the marker effects of a specific trait using statistical models. The marker effects are used to predict the breeding values of non-phenotyped individuals called testing population. GS has become a more available breeding methodology in recent years. As genotyping costs continue to decrease, cost of phenotyping will become the limiting factor of GS (Bhat et al., 2016).

Although GS in plant breeding was considered challenging (Desta & Ortiz, 2014), it has been successfully implemented in cereal crops, for example, wheat and barley (Ankamah-Yeboah et al., 2020;

Larkin et al., 2019), and has great potential for improved selection for yield and disease resistance in oats (Haikka, Knürr, et al., 2020; Haikka, Manninen, et al., 2020; Mellers et al., 2020). Genomics-resources and marker systems were for a long time limited in oats due to the complexity of the oat genome and reduced research investments compared with other major crops (Latta et al., 2019). However, the development of the 6K-SNP chip (Tinker et al., 2014) made GS more available for implementation in oat breeding and has already been implemented in Nordic breeding programmes (Ceplitis, 2014).

The composition of the training population highly affects the prediction ability, which is crucial for successful implementation of GS (Akdemir & Isidro-Sánchez, 2019; Berro et al., 2019; Crossa et al., 2016). The main criteria for training population optimization are (i) population size, (ii) genetic diversity, (iii) phenotypic diversity (Isidro et al., 2015), (iv) genetic relationship between the training and testing population (Crossa et al., 2014), and (v) degree of population structure (Werner et al., 2020). Some of these criteria could be more important than others for different traits, populations, and species (Crossa et al., 2010). Optimizing the training population is especially useful when phenotyping costs are high in traits with low heritability and in cases of high genotype by environment interaction. High heritability is also related to high prediction ability but is not something we try to optimize in this study.

By using the criteria mentioned above we have evaluated three different strategies for training population optimization. The first strategy preserves the genetic diversity and population structure from a larger population in smaller training populations (Crossa et al., 2016; Franco et al., 2005). The second strategy uses the genetic relationship between the training and testing population to identify individuals that have the lowest mean prediction error variance (PEV; Rincent et al., 2012, Isidro et al., 2015). The third strategy is based on selecting training populations with high phenotypic variation for a specific trait.

The goal of this study was to use historical data and breeding lines from the Norwegian oat breeding programme at Graminor, to develop an optimal training population for further research. The strategies mentioned above were applied to a large candidate population, with an independent breeding population as testing population. The main hypothesis is that an optimization strategy will give higher prediction abilities than a random selection. The optimization criteria genetic diversity, phenotypic diversity, and genetic similarity between training and testing population were analysed in all strategies. A wheat dataset from CIMMYT was used to validate the strategies in a completely independent breeding germplasm. The outcome of this study could contribute to the implementation of GS in commercial plant breeding programmes.

## 2 | MATERIALS AND METHODS

### 2.1 | Germplasm

Oat lines in this study were provided by Graminor plant breeding company and are listed in Table S1. Summary of the number of lines, SNP-markers, environments, and heritability is given in Table 1. Table 2 shows the number of lines tested in each location and year, and Table S8 shows the percentage of overlapping lines between the environments. All lines have been evaluated for the traits days to heading (DTH) and plant height (PH) by Graminor and the Norwegian University of Life Sciences from yield trials that were a randomized complete block design with plot size of 1.5 m × 5 m, and irrigated disease trials that were an alpha lattice design (Patterson & Williams, 1976) with plot size of 1.5 m × 1.25 m. Spatial variation was analysed by using nearest neighbour for yield trials (Cover & Hart, 1967) and alpha lattice for the disease trials. Plant height was collected by measuring the height of the plant from the ground to the top of the head 2–3 weeks after heading. Days to heading were recorded as the number of days from sowing until the date when 50% of the heads have emerged more than 50% from the flag leaf.

**TABLE 1** Summary description of oat germplasm, genotypic data and phenotypic data used in this study

| | Training population | Testing population |
|---|---|---|
| Population size | 1124 | 257 |
| Number of SNP markers | 3022 | 3022 |
| Locations | 4 | 4 |
| Years | 5 | 1 |
| Heritability of plant height | .71 | .82 |
| Heritability of days to heading | .62 | .90 |

*Note*: The number of years and locations refers to the phenotypic trials of the training population candidates, which were used to calculate the heritability of the traits.

The training population candidates consisted of 65% $F_9$ and 16% $F_{10}$–$F_{12}$ breeding lines from Graminor, and 19% are a collection of diverse material from Europe, North America and Australia. The testing population consisted of 257 Graminor $F_9$ breeding lines from 2019. The $F_9$ lines were tested for at least one year at three locations, the $F_{10}$–$F_{13}$ lines were tested for at least two years at four locations, and the diverse materials were tested for at least two locations in 2016 and one in 2017.

### 2.2 | Phenotypic data

The phenotypic data used in the genomic prediction models come from a two-stage analysis. The first stage is the calculated adjusted mean values from field designs to account for the effect of replicate and block. The second stage is to use adjusted mean values in mixed linear models to account for the environmental effects of year, locations and experiment within the same environment. The following models were used in stage two:

$$\mathbf{y} = \mathbf{b_0} + \mathbf{b_1 x_g} + \mathbf{b_2 x_l} + \mathbf{b_3 x_y} + \mathbf{b_4 x_{ly}} + \mathbf{e} \quad (1)$$

**TABLE 2** Number of lines tested in each year and location for the training population candidates

| Year | Bjørke (60.80°N, 11.20°E) | Staur (60.73°N, 11.10°E) | Rød (59.34°N, 10.89°E) | Vollebekk (59.66°N, 10.75°E) |
|---|---|---|---|---|
| 2014 | 304 | 304 | 304 | 34 |
| 2015 | 136 | 153 | 136 | 93 |
| 2016 | 174 | 344 | 174 | 289 |
| 2017 | 407 | 440 | 407 | 260 |
| 2018 | 337 | 356 | 337 | 357 |
| 2019 | 257 | 257 | 257 | 257 |

$$\mathbf{y} = \mathbf{b_0} + \mathbf{b_1 x_g} + \mathbf{b_2 x_l} + \mathbf{b_3 x_t(x_l)} + \mathbf{e} \quad (2)$$

$$\mathbf{y} = \mathbf{b_0} + \mathbf{b_1 x_g} + \mathbf{b_2 x_l} + \mathbf{e} \quad (3)$$

In the equations (Equations 1–3) **y** is the response phenotype, $\mathbf{b_0}$ the intercept, $\mathbf{b_1}$-$\mathbf{b_4}$ are coefficients, $\mathbf{x_g}$ the fixed effect of genotype, $\mathbf{x_l}$ the random effect of location, $\mathbf{x_y}$ the random effect of year, $\mathbf{x_t}$ the random effect of trial, $\mathbf{x_{ly}}$ the interaction between year and location, $\mathbf{x_t\,(x_l)}$ the effect of experiment nested in location, and **e** the error term.

Equation (1) was used on plant height in the training population, Equation (2) on days to heading in both training and testing population, and Equation (3) on plant height in the testing population. Plant height and days to heading were normalized with different models because data collected from the irrigated disease trials differed for days to heading compared with the other trials in the same year and location, so the factor of experiment was added to the model. The factor of experiment also contains the effect of year as the same experiments are only tested for one year. For plant height, it was sufficient to use year and location as factor. Material with phenotypic values two standard deviations from the mean were excluded as the distribution became skewed.

## 2.3 | Heritability

The broad sense heritability ($h^2$) was calculated as

$$h^2 = V_G/V_P \quad (4)$$

where $V_G$ is the variance of genotype and $V_P$ is the variance of phenotype. $V_P$ is equal to the $V_G + V_e$, where $V_e$ is the variance of error. $V_G$ was estimated using the $x_g$ term using the following mixed linear model:

$$\mathbf{y} = \mathbf{b_0} + \mathbf{b_1 x_g} + \mathbf{b_2 x_l} + \mathbf{b_3 x_t(x_l)} + \mathbf{e} \quad (5)$$

where **y** is the response phenotype, $\mathbf{b_0}$ is the intercept, $\mathbf{b_1}$–$\mathbf{b_3}$ are coefficients, $\mathbf{x_g}$ is the random effect of genotype, $\mathbf{x_l}$ is the fixed effect of location, $\mathbf{x_t\,(x_l)}$ is the fixed effect of trial nested in location, and **e** is the error using the Minitab software (Minitab, 2010). This calculation accounts for the fixed effect of environment, leaving only the effect of genotype and error in Equation (4).

## 2.4 | Genotyping

All lines were genotyped with a customized, unpublished 20 k SNP chip. The genetic data were analysed and filtered with a 10% missing values threshold and 5% MAF based on the training population candidates, resulting in 3022 polymorphic markers. The missing marker data were imputed with the 'impute' function and 'means' method with the package 'e1071' in the R statistical software (Meyer et al., 2021).

## 2.5 | Experimental design

Each optimization strategy was repeated 20 times for each population size of 100, 200, 300, 400 and 500. Average prediction ability was calculated as the average correlation between predicted and observed breeding values of the testing population. Bayesian ridge regression (BRR) was used to compute the marker effects, and the 'BGLR' function of the 'BGLR' package in the R software (Pérez & de los Campos, 2014) was used to calculate the genomic estimated breeding values of the testing population. The number of iterations were set to 30,000 and the burnin to 15,000.

## 2.6 | Training population optimization strategies

This study aimed to optimize known training population criteria. Each strategy was compared with a random selection. The correlation between the optimization criteria and the prediction abilities were calculated and tested for significance with ANOVA.

### 2.6.1 | Phenotypic selection

Phenotypic selection aims to maximize phenotypic variation in the training populations and is abbreviated to PheSe for the rest of the paper. Based on the MLM output data (Figure 1), equal proportions of lines with the most extreme highest and lowest adjusted breeding values were selected for the PheSe populations. This was done once for each population size and not replicated 20 times like the other strategies. A similar approach was proposed by Zhao et al. (2012) in a slightly different premise. They argue that a fraction of the training

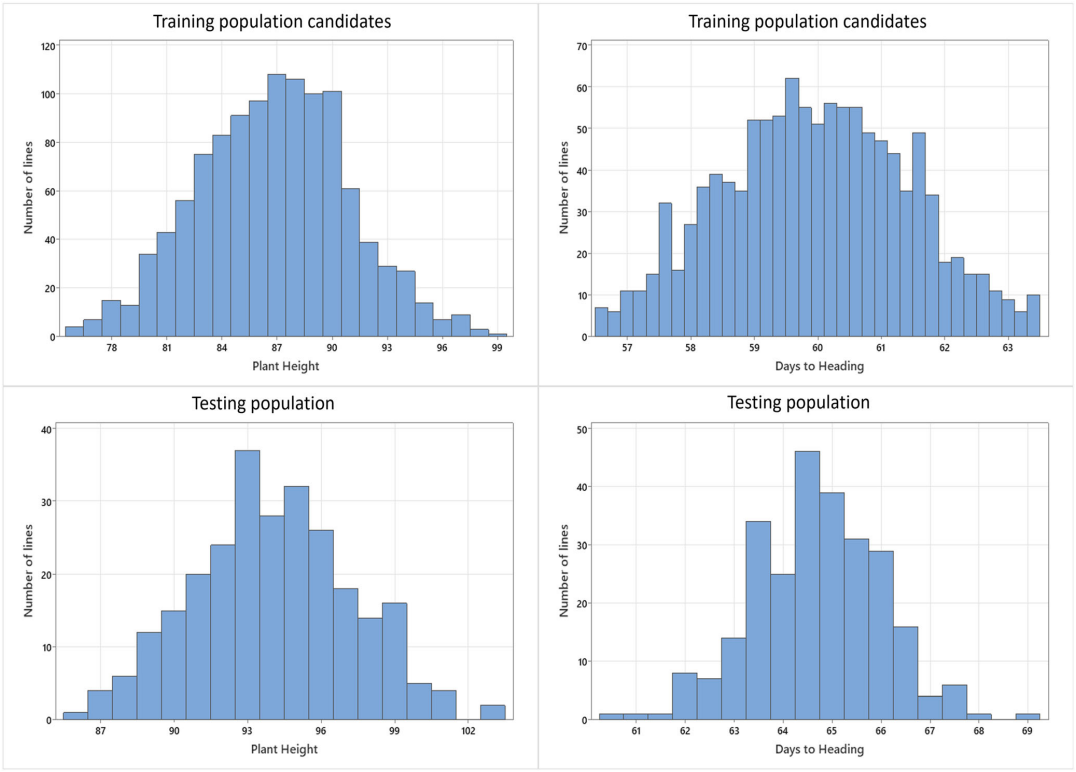



**FIGURE 1** Phenotypic distribution of the training population candidates and testing population for days to heading and plant height after applying the models in Equations (1)–(3) [Color figure can be viewed at wileyonlinelibrary.com]

population should consist of inferior material to increase prediction accuracy (Zhao et al., 2012).

### 2.6.2 | Prediction core

The prediction core strategy aims to optimize the genetic relationship between the training and testing population by minimizing the PEV of the testing population, while also maintaining high diversity in the training population, which is done by calculating an optimization criterion called coefficient of determination (CD; Laloë, 1993). In the rest of the paper prediction core is abbreviated to PreCo. The strategy was published in 2012 by Rincent et al. (2012) and made into an R package by Akdemir (2018).

PreCo populations were selected by performing principal component analyses (PCA) on the genetic markers of the training and testing populations (Akdemir et al., 2015). The first 100 principal components (PCs) were used as input for a selection algorithm using the function 'GenAlgForSubsetSelection' in the R package 'STPGA', which starts off with a random sample, calculates the CD values and replaces one genotype at the time until it finds one that increases or gives the same the CD value. This process is repeated until no further increase in CD values is achieved (Akdemir, 2018). CDMEAN2 was used as selection criterion. The arguments of the function were set to npop = 300, nelite = 20, niterations = 5000, and minitbefstop = 1000. 'Npop' refers to the number of crosses in the testing population, and 'nelite' refers to the number of parents used. We chose higher parameters than required in order to give the algorithm more power and better solutions. The 'niterations' argument is the maximum number of iterations the selections algorithm use to find the optimal solution, whereas the 'minitbefstop' argument is the number of equal solution required for the algorithm to stop before the maximum is reached.

### 2.6.3 | Diversity core

The diversity core strategy aim to preserve the genetic diversity and population structure from the total candidate population in smaller populations (Crossa et al., 2016; Franco et al., 2005). Hereafter, diversity core is abbreviated to DivCo.

DivCo populations were selected by performing a structure analysis with the software STRUCTURE (Hubisz et al., 2009), and the structure harvester (Earl & vonHoldt, 2012) to determine the optimal number of clusters. A dendrogram was created with the 'hclust' function which performs a hierarchical clustering of a distance matrix based on the genetic markers. The Ward.D2 method was used in the

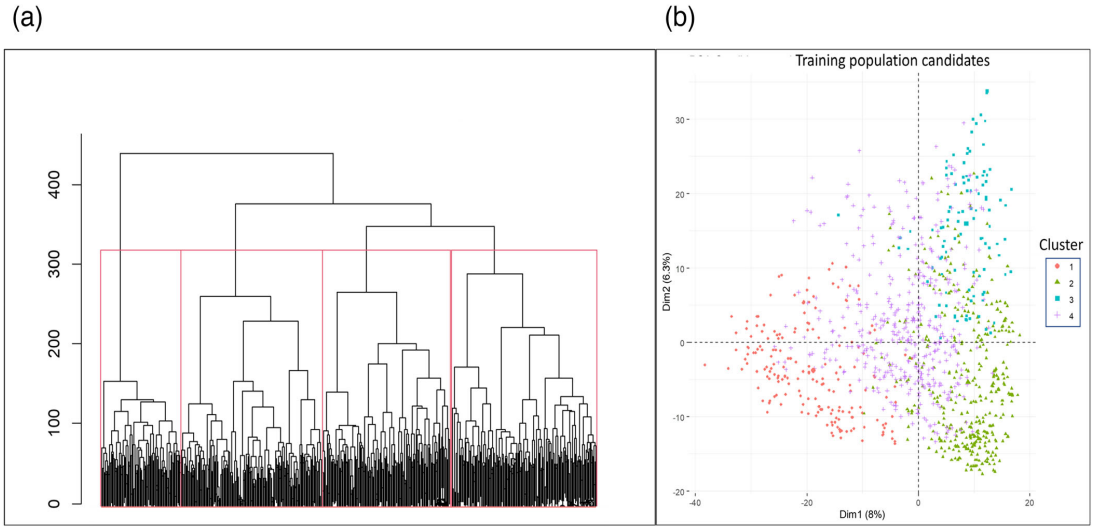



**FIGURE 2** (a) Dendrogram of the training population candidates separated into four clusters. The height of the dendrogram is given as the total sum of squares between individuals and each cluster (b) principal component analysis of the training population candidates separated into four clusters [Color figure can be viewed at wileyonlinelibrary.com]

clustering to ensure that the within-group distance is low, and the between-group distance is high (Ward, 1963). The dendrogram was separated into the optimal number of clusters from the structure analysis with the 'rect.hclust' function in R, which isolates the clusters with the highest genetic distance to each other (Figure 2a). Figure 2b shows the four clusters in the PCA.

The mean distance (MD) of each cluster was summed up, and the number of genotypes selected from each cluster were proportional to sum of MD from all clusters. A stratified random sampling was done 1000 times in each cluster, and the subsamples with the highest average mean distance were selected for the DivCo populations.

### 2.6.4 | Random selection

Random selection was included as a control to represent random unoptimized training populations and is abbreviated to RanSe for the rest of the paper. Populations were selected by using the 'sample_n' function from the 'dplyr' package in the R software (Wickham et al., 2021), which randomly selects a given number of random rows from a dataframe.

## 2.7 | Statistical analysis

ANOVA was used to identify significant effects of optimization strategy on prediction ability, and the equation is stated as:

$$\mathbf{y} = \mathbf{b_0} + \mathbf{b_1 x_1} + \mathbf{b_2 x_2} + \mathbf{b_3 x_{12}} + \mathbf{e} \quad (6)$$

where $\mathbf{y}$ is the response prediction ability, $\mathbf{b_0}$ the intercept, $\mathbf{b_1}$–$\mathbf{b_3}$ are coefficients, $\mathbf{x_1}$ the fixed effect of optimization strategy, $\mathbf{x_2}$ the fixed effect of population size, $\mathbf{x_{12}}$ the interaction between population size and optimization strategy, and $\mathbf{e}$ the error term.

A Tukey pairwise comparison test was used for each pair of strategies to identify if they were significantly different from each other.

Bootstrapping was also used to calculate the significant differences between the strategies within each defined population size. Bootstrapping was used because the PheSe strategy was not replicated. The sample closest to the mean prediction ability in each strategy and population size was compared with each other. The bootstrapping was conducted by removing a random set of lines from the testing population and calculating the prediction ability of the remaining lines in the testing population for both training populations. The procedure was done using the R package 'GRousellet/bootcorci' (Rousselet et al., 2019) and the function 'twocorci.ov' by removing a random set of lines from the testing population, and calculating the prediction ability for the remaining lines. The bootstrapping was done with a significance level of $\alpha = .05$ and 2000 iterations.

## 2.8 | Optimization criteria

The genetic diversity was calculated as the mean expected heterozygosity by using the R package 'diveRsity' and the function 'Divbasic' in the R software which calculates the frequencies of the alleles of each marker using the formula (2)*p*q, where p and q is the frequencies of the different alleles. Then then mean 2pq is calculated for all markers (Keenan et al., 2013).

The genetic similarity between the training and testing populations was calculated as the proportion of shared alleles per pair of populations based on the allele frequencies, summed and averaged across all loci. The calculations were done by using the R package 'PopGenReport' and the function 'pairwise.propShared' in the R software (Adamack & Gruber, 2014).

The Phenotypic variance was calculated by using the 'var' function of the 'base' R software.

## 2.9 | Wheat (*Triticum aestivum*) validation sets

Four different datasets were provided by CIMMYT to validate the results of this study and is described in Table 3. The largest were chosen as the training population candidates, whereas the rest were used as testing populations.

**TABLE 3** Summary description of wheat germplasm, genotypic data and phenotypic data used in this study (Montesinos-López et al., 2019)

| | |
|---|---|
| Training population candidates | 980 lines |
| Testing population 1 | 766 lines |
| Testing population 2 | 775 lines |
| Testing population 3 | 964 lines |
| Number of SNP markers | 9285 |
| Locations | 6 per year for each population |
| Years | 4, 1 year per population |
| Traits | Plant height, days to heading and grain yield |

# 3 | RESULTS

## 3.1 | Prediction ability

For plant height (Figure 3a), RanSe and DivCo showed similar prediction abilities of .26 using population sizes 300–500, whereas the prediction abilities of PheSe and PreCo were higher at .33 and .35, respectively. PreCo performed significantly better than DivCo in size 300 and 400, whereas PheSe performed significantly better than DivCo at size 400 and RanSe at 400 and 500 (Table 4). PreCo performed approximately .025 points better than PheSe in size 200–500, but this difference was not significant. The prediction ability of all

**TABLE 4** Significant *p*-values values from bootstrapping tests for plant height for the optimization strategies prediction core (PreCo), diversity core (DivCo), phenotypic selection (PheSe) and random selection (RanSe)

| Optimization strategies | Population size | $p < .05$ |
|---|---|---|
| DivCo vs. PreCo | 300 | .036 |
| DivCo vs. PreCo | 400 | .018 |
| DivCo vs. PheSe | 400 | .019 |
| RanSe vs. PheSe | 400 | .038 |
| RanSe vs. PheSe | 500 | .046 |

*Note*: Days to heading had no significant differences in bootstrapping.

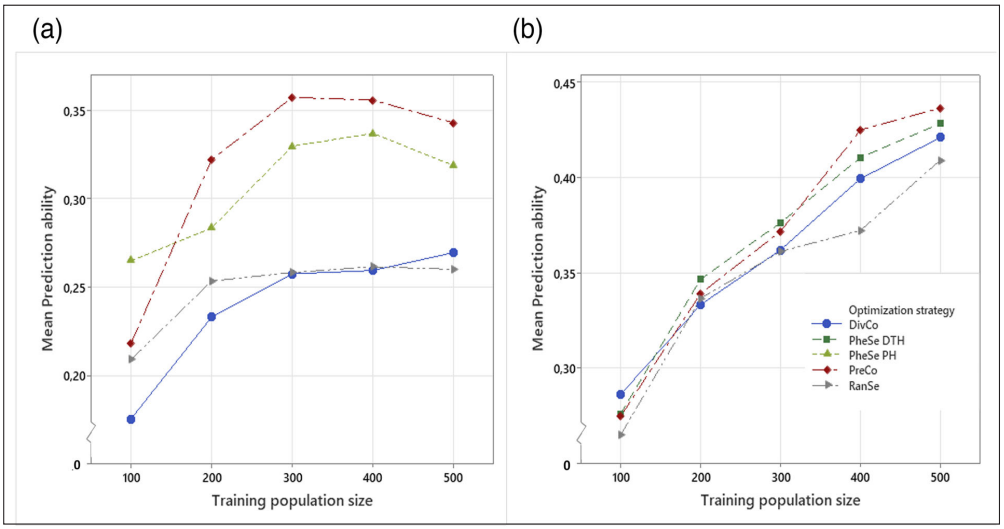



**FIGURE 3** Average prediction abilities for (a) plant height (PH) and (b) days to heading (DTH in oats for the different optimization strategies prediction core (PreCo), diversity core (DivCo), phenotypic selection (PheSe) and random selection (RanSe) across different training population sizes [Color figure can be viewed at wileyonlinelibrary.com]

lines plateaued at size 300, with a slight decrease in larger populations for PreCo and PheCo as the prediction ability of the total candidate population was .33 (data not shown).

For days to heading (Figure 3b), all strategies performed similarly for population sizes 100–300 but differed more for population sizes 400 and 500, where the PreCo showed the highest prediction ability followed by PheSe, DivCo and RanSe. However, none of these differences were significant in a bootstrapping test. Prediction ability increased linearly with population size with a maximum of .44 for PreCo at size 500. Using all lines in the candidate population resulted in a prediction ability of .49 (data not shown).

## 3.2 | ANOVA

The ANOVA results (Table 5) showed that population size contributed to about 62% of the variation in prediction ability for days to heading and 20% for plant height. Optimization strategy contributed to less than 2% of the variation for days to heading, and 21% for plant height. Both factors were significant for both traits, whereas the interaction term was not significant. The PreCo populations yielded significantly higher prediction ability than RanSe for days to heading, and significantly better than DivCo and RanSe for plant height (Table 6). No other significant differences were detected using the Tukey test.

## 3.3 | Genetic diversity, similarity and phenotypic variance

The optimizations criteria (Figures 4 and 5) showed that the RanSe populations had the lowest phenotypic and genetic diversity, and intermediate genetic similarities. DivCo populations had high genetic diversity, low genetic similarity, and intermediate phenotypic diversity. PreCo populations had the highest genetic diversity, high genetic similarity, and high phenotypic diversity. PheSe populations had very high phenotypic diversity for both traits, intermediate genetic diversity for days to heading and low for plant height, and the highest genetic similarity for plant height and the lowest for days to heading.

There was a significant positive correlation between prediction ability and genetic similarity for both traits, with r values of .48 for plant height and .4 for days to heading. There was also a significant positive correlation between prediction ability and genetic and phenotypic diversity for plant height, but with a low r of .25 for both criteria. Population size had a high significant positive correlation with prediction ability with r values of .77 for days to heading and .36 for plant height.

## 3.4 | Wheat validation results

Population size had a significant large effect on the variation in prediction ability for all three testing populations in all traits (Table 7). Selection strategy showed a significant contribution in two out of the three

**TABLE 6** Results from the pairwise comparisons of the prediction abilities of the four optimization strategies prediction core (PreCo), diversity core (DivCo), phenotypic selection (PheSe) and random selection (RanSe) for the traits days to heading and plant height

| Strategies compared | | Difference in means | *p*-values |
|---|---|---|---|
| **Days to heading** | | | |
| PheSe | DivCo | 0.007 | .980 |
| PreCo | DivCo | 0.009 | .384 |
| RanSe | DivCo | −0.012 | .177 |
| PreCo | PheSe | 0.002 | 1.000 |
| RanSe | PheSe | −0.019 | .740 |
| RanSe | PreCo | −0.020** | .002 |
| **Plant height** | | | |
| PheSe | DivCo | 0.068 | .062 |
| PreCo | DivCo | 0.080** | <.001 |
| RanSe | DivCo | 0.010 | .673 |
| PreCo | PheSe | 0.013 | .968 |
| RanSe | PheSe | −0.058 | .140 |
| RanSe | PreCo | −0.071** | <.001 |

*$p < .05$.**$p < .01$.

**TABLE 5** ANOVA for oats with prediction ability as response variable and population size (size), optimization strategy (strategy) and the interaction term Size*Strategy as factors for days to heading and plant height

| Source | *df* | Contribution | Adj SS | Adj MS | *F*-value | *p*-value |
|---|---|---|---|---|---|---|
| **Days to heading** | | | | | | |
| Size | 4 | 61.51% | 0.187 | 0.047 | 28.95 | <.001 |
| Strategy | 3 | 1.65% | 0.022 | 0.007 | 4.45 | .004 |
| Size*Strategy | 12 | 1.56% | 0.020 | 0.002 | 1.05 | .404 |
| Error | 285 | 35.27% | 0.461 | 0.002 | | |
| **Plant height** | | | | | | |
| Size | 4 | 20.33% | 0.074 | 0.019 | 5.25 | <.001 |
| Strategy | 3 | 21.43% | 0.392 | 0.131 | 37 | <.001 |
| Size*strategy | 12 | 3.23% | 0.059 | 0.005 | 1.4 | .167 |
| Error | 285 | 55.01% | 1.007 | 0.004 | | |

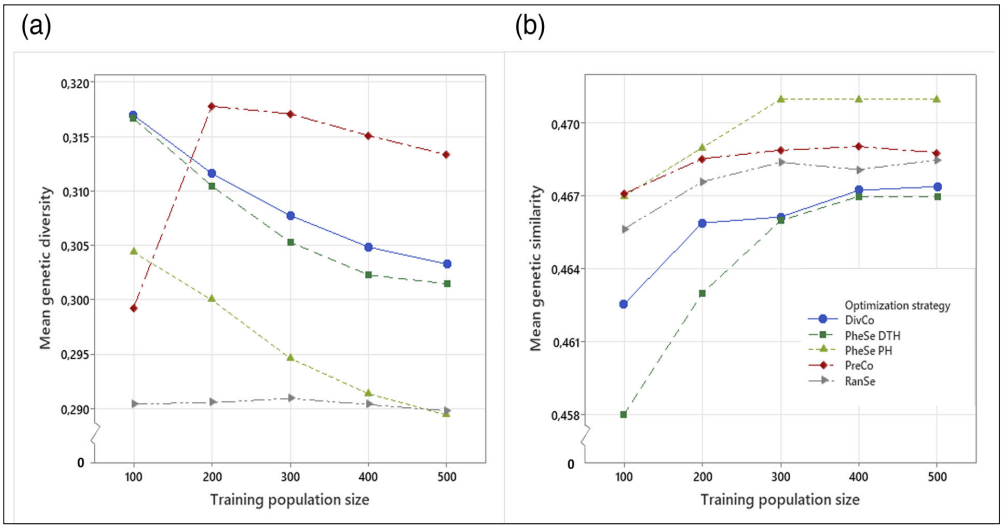



**FIGURE 4** (a) Mean genetic diversity and (b) mean genetic similarity for the optimization strategies diversity core (DivCo), prediction core (PreCo), phenotypic selection (PheSe) for plant height (PH) and days to heading (DTH) and random selection (RanSe) in population size 100–500 [Color figure can be viewed at wileyonlinelibrary.com]

populations for days to heading, and one out of the three populations for plant height and grain yield. In all cases of with non-significant contribution of selection strategy for prediction ability of days to heading and plant height, the Size × Strategy interaction term was significant. For grain yield prediction, the error term was always non-significant.

DivCo were the best strategy in terms of prediction ability in three of four significant pairwise comparisons in testing population 1 (Table S5), PreCo were best in three of four comparisons in population 2 (Table S6), and PheSe were best in three of five comparisons in population 3 (Table S7). Full ANOVA tables are available in Tables S2–S4 along with plots of the mean prediction ability for the optimization strategies (Figure S1).

## 4 | DISCUSSION

Our study compared three training population optimization strategies (prediction core, diversity core and phenotypic selection) to random selection. The training population criteria optimized were genetic diversity, phenotypic variance and genetic similarity. The four strategies were validated for their prediction ability, that is, their ability to predict the phenotypes of a given testing population, and analysed for their genetic diversity, genetic similarity between the training and testing population, and phenotypic diversity.

The broad sense heritability was high for the traits days to heading (.62) and plant height (.71) which is expected for these traits. Studies have shown that smaller training populations are needed for traits with high heritability (Kaler et al., 2022; Zhang et al., 2017), and others have shown that high prediction ability can be achieved for plant height and days to heading with small training populations (Baertschi et al., 2021; Haikka, Knürr, et al., 2020). A study done on unbalanced agronomic traits showed that the standard broad sense heritability calculation overestimates the heritability (Schmidt et al., 2019), which is also likely true for the dataset of this research. This can explain the relatively low prediction abilities in this study. But the overestimation does not likely affect the ranking of the strategies as an adjustment of the heritability as suggested by Schmidt et al. (2019) would shrink the heritability for both traits equally. There are also large G × E effects on unbalanced data, which could also have contributed to the low prediction abilities. The observation that the maximum prediction ability was reached at population size 300 for plant height, whereas it was not yet reached at size 500 for days to heading is likely an effect of the difference in heritability between the two traits.

The main factors effecting the prediction abilities were population size and genetic similarity, which has been highlighted as important training population criteria in several studies (Liu et al., 2018; Lorenz & Nice, 2017; Zhang et al., 2017). Genetic and phenotypic diversity were however less important since increasing these criteria alone would decrease the genetic similarity as the testing population has low genetic diversity. Other studies also found that genetic diversity is more important when population structure is present (Berro et al., 2019; Isidro et al., 2015). One study on diversity core and prediction core found that they gave similar prediction abilities (Crossa et al., 2016), which is different from what this study concludes. But in their study, the authors used diverse landrace populations with a much higher degree of genetic diversity and population structure. These populations are better suited for the diversity core strategy than the more narrow breeding population used in this study.

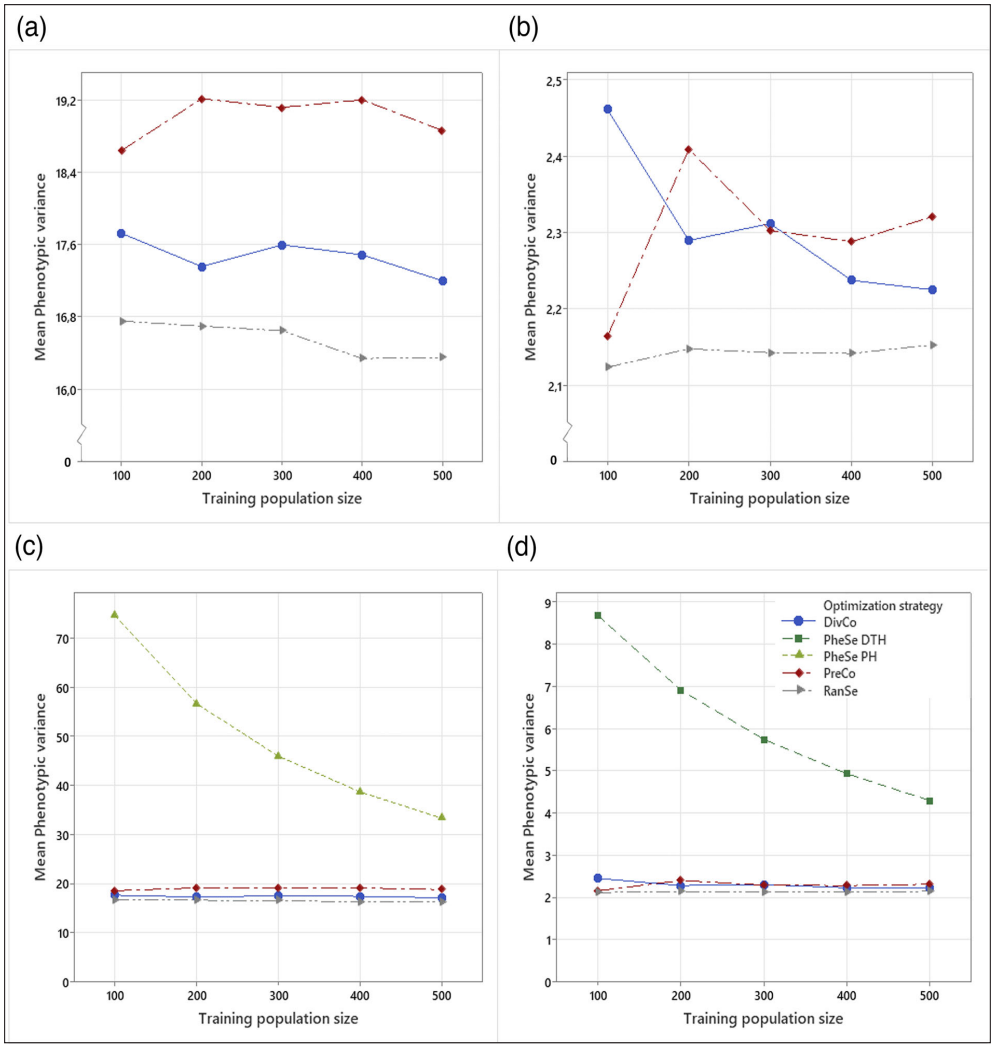



**FIGURE 5** Mean phenotypic diversity of the optimization strategies diversity core (DivCo), prediction core (PreCo), phenotypic selection (PheSe) and random selection (RanSe) in population size 100–500 for the traits (a) plant height (PH) without PheSe, (b) days to heading (DTH) without PheSe, (c) plant height with PheSe and (d) days to heading with PheSe [Color figure can be viewed at wileyonlinelibrary.com]

Another study (Akdemir et al., 2015) showed that the prediction core consistently gave better prediction abilities than random selection across population sizes and different traits, which is similar to the results as presented in this paper.

## 4.1 | Diversity core

The diversity core strategy worked as intended. It produced training populations with similar genetic diversity as the total candidate population and selected relatively equal number of lines from all four clusters. DivCo populations performed very similar to RanSe in both traits and all population sizes (Figure 3). DivCo populations showed lower genetic similarity than RanSe (Figure 4b), which along with the fact that it showed intermediate genetic and phenotypic diversity (Figures 4a and 5a,b) explains the low prediction abilities. The DivCo strategy is not optimal for our data because of the lack of population structure and genetic diversity in the testing population. Further research on this strategy in populations with more diversity and population structure would be useful to properly evaluate it. However, when exotic material is introduced into a breeding programme DivCo could be more useful. As the DivCo strategy does not depend on a specific testing population it is reasonable to think that it would give more stable prediction abilities.

**TABLE 7** *p*-values and contribution percentage from ANOVA with prediction ability as response variable population size (size), optimization strategy (strategy) and the interaction term Size*Strategy as factors

| | Days to heading | | Plant height | | Grain yield | |
|---|---|---|---|---|---|---|
| Source | Contribution | *p*-value | Contribution | *p*-value | Contribution | *p*-value |
| **Validation set 1** | | | | | | |
| Size | 12.86% | <.001 | 36.65% | <.001 | 29.77% | <.001 |
| Strategy | 1.57% | .034 | 5.61% | <.001 | .44% | .395 |
| Size*Strategy | 20.60% | <.001 | 13.70% | <.001 | 2.85% | .151 |
| Error | 64.98% | | 44.04% | | 66.95% | |
| **Validation set 2** | | | | | | |
| Size | 19.55% | <.001 | 45.35% | <.001 | 25.18% | <.001 |
| Strategy | .23% | .648 | .14% | .666 | 20.05% | <.001 |
| Size*Strategy | 4.23% | .049 | 4.75% | .001 | .74% | .864 |
| Error | 75.99% | | 49.76% | | 54.03% | |
| **Validation set 3** | | | | | | |
| Size | 28.50% | <.001 | 50.9 5% | <.001 | 9.99% | <.001 |
| Strategy | 8.20% | <.001 | .22% | .502 | 1.17% | .143 |
| Size*Strategy | 1.34% | .628 | 2.70% | .037 | 3.76% | .132 |
| Error | 61.96% | | 46.12% | | 85.09% | |

*Note*: The traits analysed were days to heading, plant height and grain yield. The wheat data used for training and testing populations are described in Section 2.7.

## 4.2 | Prediction core

The PreCo strategy worked as intended and produced training populations highly related to the testing population (Figure 4b). The PreCo populations also showed high genetic diversity (Figure 4a) and relatively high phenotypic variance (Figure 5a,b). This is likely because CD values in addition to minimizing PEV also maintains high genetic distance between individuals in the training population. The combination of high values for the three optimization criteria, and especially the genetic similarity likely explains why PreCo gave the highest prediction abilities. For the sake of this study, it would have been useful to also include populations only selected based on the PEV values. This would likely have decreased the genetic and phenotypic diversity. The PreCo strategy with the CD criteria works well when you know the genotypes you want to predict. However, further research is needed into the PreCo strategy to see whether these prediction abilities are stable across different testing populations. If the testing population is a good representation of the genetic diversity of the breeding programme, then the training population should work for the next breeding cycles as well.

## 4.3 | Phenotypic selection

The PheSe strategy selected training populations with very high phenotypic variance (Figure 5c,d). The days to heading populations showed similar genetic diversity as DivCo, but also the lowest genetic similarity (Figure 4). The plant height populations showed low genetic diversity but the highest genetic similarity. Both PheSe populations gave relatively high prediction abilities, indicating that phenotypic variance is an important criterion to optimize, despite their low genetic diversity in the plant height populations, and low genetic similarity in the days to heading populations. This can be either due to overfitting of the marker effects or increased diversity for the relevant markers. Our study suggests that PheSe is a good strategy for selecting training populations when no genotype data is available, and that the inclusion of material with low breeding value is important to increase prediction accuracy. In our study we maximized this by selecting 50% lines with low breeding values, whereas Zhao et al. concludes that 30% is enough to ensure high accuracy without underfitting of marker effects (Zhao et al., 2012).

## 4.4 | Wheat validation

The wheat validation sets were inconclusive in determining which strategy works best, as they rank differently in the different testing populations for different traits. We can see that DivCo worked best in population 1, PreCo in population 2 and PheSe in population 3. We did not do any further analysis into the optimization criteria of the wheat datasets. Further research can show why the optimization strategies worked differently for the different testing populations. A likely reason could be that the phenotypic data for the three populations were collected from different years, which increases the G × E effect. It is shown that the correlation between environments can vary a lot for the same trait (Cooper & DeLacy, 1994), which could explain the low prediction abilities in the validation sets.

## 4.5 | Conclusion

Of the three strategies analysed, prediction core had the highest average prediction ability in most population sizes for both traits and produced training populations with high genetic diversity, high genetic similarity to the testing population and high phenotypic variance compared with random selection. Genetic similarity along with population size were the most important criteria to optimize in the training populations. More research is needed to evaluate how well the prediction core strategy works over several breeding cycles, but our research points to prediction core as the best strategy to optimize training populations in cereals.

### ACKNOWLEDGEMENTS

This research was funded by the Research Council of Norway and Graminor (NFR # 286606) (the research funds for agriculture and the food industry - FFL/JA). The authors wish to thank Graminor breeders for providing the material for this research and the technical staff at Graminor and NMBU for helping in field experiments and data collection. We would also like to thank Prof. Åsmund Bjørnstad and Dr. Trond Buraas for sharing their experience and knowledge about oats. Further, we acknowledge the International Maize and Wheat Improvement Center for providing wheat material and support on genomic predictions.

### CONFLICTS OF INTEREST

All authors have no conflicts of interest to declare that are relevant to the content of this article.

### AUTHOR CONTRIBUTIONS

ESS collected and analysed historical data from Graminor; MA provided funding and support in data analysis; SW and JC provided support in genomic predictions; ML and AKS provided support in genetic analysis; CJ provided support in phenotypic analysis; ESS wrote the manuscript; and MA, SW, JC, ML, AKS and CJ participated in revising the scientific work and writing.

### DATA AVAILABILITY STATEMENT

Research data are not shared.

### ORCID

*Espen Sannes Sørensen* https://orcid.org/0000-0003-4646-8250
*Morten Lillemo* https://orcid.org/0000-0002-8594-8794

### REFERENCES

Adamack, A. T., & Gruber, B. (2014). PopGenReport: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, *5*(4), 384–387. https://doi.org/10.1111/2041-210X.12158

Akdemir, D. (2018). STPGA: Selection of Training Populations by Genetic Algorithm. R package version 5.2.1. https://CRAN.R-project.org/package=STPGA

Akdemir, D., & Isidro-Sánchez, J. (2019). Design of training populations for selective phenotyping in genomic prediction. *Scientific Reports*, *9*, 1446. https://doi.org/10.1038/s41598-018-38081-6

Akdemir, D., Sanchez, J. I., & Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selelection Evolution*, *47*, 38. https://doi.org/10.1186/s12711-015-0116-6

Ankamah-Yeboah, T., Janss, L. L., Jensen, J. D., Hjortshøj, R. L., & Rasmussen, S. K. (2020). Genomic selection using pedigree and marker-by-environment interaction for barley seed quality traits from two commercial breeding programs. *Frontiers in Plant Science*, *11*, 539. https://doi.org/10.3389/fpls.2020.00539

Baertschi, C., Cao, T. V., Bartholomé, J., Ospina, Y., Quintero, C., Frouin, J., Bouvet, J.-M., & Grenier, C. (2021). Impact of early genomic prediction for recurrent selection in an upland rice synthetic population. *G3 Genes|Genomes|Genetics*, *11*(12), jkab320. https://doi.org/10.1093/g3journal/jkab320

Berro, I., Lado, B., Nalin, R. S., Quincke, M., & Gutiérrez, L. (2019). Training population optimization for genomic selection. *The Plant Genome*, *12*(3), 190028. https://doi.org/10.3835/plantgenome2019.04.0028

Bhat, J. A., Ali, S., Salgotra, R. K., Mir, Z. A., Dutta, S., Jadon, V., Tyagi, A., Mushtaq, M., Jain, N., Singh, P. K., Singh, G. P., & Prabhu, K. V. (2016). Genomic selection in the era of next generation sequencing for complex traits in plant breeding. *Frontiers in Genetics*, *7*, 221. https://doi.org/10.3389/fgene.2016.00221

Ceplitis, A. (2014). Lantmännen spring oat breeding [PowerPoint slides]. American Oat Workers Conferance, Ottawa, 2014. https://oatnews.org/oatnews_pdfs/AOWC2014/breeders/Ceplitis_breedingAOWC2014.pdf

Cooper, M., & DeLacy, I. H. (1994). Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theoretical and Applied Genetics*, *88*(5), 561–572. https://doi.org/10.1007/BF01240919

Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964

Crossa, J., Campos, G., Pérez, P., Gianola, D., Burgueño, J., Araus, J. L., Makumbi, D., Singh, R. P., Dreisigacker, S., Yan, J., Arief, V., Banziger, M., & Braun, H. J. (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics*, *186*(2), 713–724. https://doi.org/10.1534/genetics.110.118521

Crossa, J., Jarquín, D., Franco, J., Pérez-Rodríguez, P., Burgueño, J., Saint-Pierre, C., Vikram, P., Sansaloni, C., Petroli, C., Akdemir, D., Sneller, C., Reynolds, M., Tattaris, M., Payne, T., Guzman, C., Peña, R. J., Wenzl, P., & Singh, S. (2016). Genomic prediction of Gene Bank wheat landraces. *G3: Genes, Genomes, Genetics*, *6*(7), 1819–1834. https://doi.org/10.1534/g3.116.029637

Crossa, J., Pérez, P., Hickey, J., Burgueño, J., Ornella, L., Ceròn-Rojas, J., Zhang, X., Dreisigacker, S., Babu, R., Li, Y., Bonnett, D., & Mathews, K. (2014). Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity*, *112*(1), 48–60. https://doi.org/10.1038/hdy.2013.16

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends in Plant Science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Desta, Z. A., & Ortiz, R. (2014). Genomic selection: Genome-wide prediction in plant improvement. *Trends in Plant Science*, *19*(9), 592–601. https://doi.org/10.1016/j.tplants.2014.05.006

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–336. https://doi.org/10.1007/s12686-011-9548-7

Franco, J., Crossa, J., Taba, S., & Shands, H. A. (2005). Sampling strategy for conserving genetic diversity when forming core subsets. *Crop Science*, *45*(3), 1035–1044. https://doi.org/10.2135/cropsci2004.0292

Haikka, H., Knürr, T., Manninen, O., Pietilä, L., Isolahti, M., Teperi, E., Mäntysaari, E. A., & Strandén, I. (2020). Genomic prediction of grain yield in commercial Finnish oat (Avena sativa) and barley (Hordeum vulgare) breeding programmes. *Plant Breeding*, *139*(3), 550–561. https://doi.org/10.1111/pbr.12807

Haikka, H., Manninen, O., Hautsalo, J., Pietilä, L., Jalli, M., & Veteläinen, M. (2020). Genome-wide association study and genomic prediction for *fusarium graminearum* resistance traits in Nordic oat (*Avena sativa* L.). *Agronomy*, *10*(2), 174. https://doi.org/10.3390/agronomy10020174

Heffner, E. L., Lorenz, A. J., Jannink, J. L., & Sorrells, M. E. (2010). Plant breeding with genomic selection: Gain per unit time and cost. *Crop Science*, *50*(5), 1681–1690. https://doi.org/10.2135/cropsci2009.11.0662

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*(5), 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Isidro, J., Jannink, J. L., Akdemir, D., Poland, J., Heslot, N., & Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *Theoretical and Applied Genetics*, *128*(1), 145–158. https://doi.org/10.1007/s00122-014-2418-4

Kaler, A. S., Purcell, L. C., Beissinger, T., & Gillman, J. D. (2022). Genomic prediction models for traits differing in heritability for soybean, rice, and maize. *BMC Plant Biology*, *22*, 87 (2022). https://doi.org/10.1186/s12870-022-03479-y

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*(8), 782–788. https://doi.org/10.1111/2041-210X.12067

Laloë, D. (1993). Precision and information in linear models of genetic evaluation. *Genetics Selection Evololution*, *25*, 557. https://doi.org/10.1186/1297-9686-25-6-557

Larkin, D. L., Lozada, D. N., & Mason, R. E. (2019). Genomic selection—Considerations for successful implementation in wheat breeding programs. *Agronomy*, *9*(9), 479. https://doi.org/10.3390/agronomy9090479

Latta, R. G., Bekele, W. A., Wight, C. P., & Tinker, N. A. (2019). Comparative linkage mapping of diploid, tetraploid, and hexaploid Avena species suggests extensive chromosome rearrangement in ancestral diploids. *Scientific Reports*, *9*(1), 12298. https://doi.org/10.1038/s41598-019-48639-7

Li, H., Rarsheed, A., Hickey, L. T., & He, Z. (2018). Rapid breeding cycles: Reaping the low hanging fruit of increasing genetic gain. *Trends in Plant Science*, *23*, 184–186. https://doi.org/10.1016/j.tplants.2018.01.007

Liu, X., Wang, H., Wang, H., Guo, Z., Xu, X., Liu, J., Wang, S., Li, W. X., Zou, C., Prasanna, B. M., Olsen, M. S., Huang, C., & Xu, Y. (2018). Factors affecting genomic selection revealed by empirical evidence in maize. *The Crop Journal*, *6*(4), 341–352. https://doi.org/10.1016/j.cj.2018.03.005

Lorenz, A., & Nice, L. (2017). Training Population Design and Resource Allocation for Genomic Selection in Plant Breeding. In R. Varshney, M. Roorkiwal, & M. Sorrells (Eds.). *Genomic selection for crop improvement*. Springer. https://doi.org/10.1007/978-3-319-63170-7_2

Mellers, G., Mackay, I., Cowan, S., Griffiths, I., Martinez-Martin, P., Poland, J. A., Bekele, W., Tinker, N. A., Bentley, A. R., & Howarth, C. J. (2020). Implementing within-cross genomic prediction to reduce oat breeding costs. *The Plant Genome*, *13*(1), e20004. https://doi.org/10.1002/tpg2.20004

Meuwissen, T. H., Hayes, N. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. (2021). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7–9, https://CRAN.R-project.org/package=e1071

Minitab 21 Statistical Software. (2010). *Computer software*. Minitab, Inc. www.minitab.com

Montesinos-López, O. A., Martín-Vallejo, J., Crossa, J., Gianola, D., Hernández-Suárez, C. M., Montesinos-López, A., Juliana, P., & Singh, R. (2019). New deep learning genomic-based prediction model for multiple traits with binary, ordinal, and continuous phenotypes. *G3: Genes, Genomes, Genetics*, *9*(5), 1545–1556. https://doi.org/10.1534/g3.119.300585

Patterson, H. D., & Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika*, *63*, 83–92. https://doi.org/10.1093/biomet/63.1.83

Pérez, P., & de los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, *198*(2), 483–495. https://doi.org/10.1534/genetics.114.164442

Rincent, R., Laloë, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., Rodríguez, V. M., Moreno-Gonzalez, J., Melchinger, A., Bauer, E., Schoen, C. C., Meyer, N., Giauffret, C., Bauland, C., Jamin, P., Laborde, J., Monod, H., Flament, P., Charcosset, A., & Moreau, L. (2012). Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: Comparison of methods in two diverse groups of maize Inbreds. (*Zea mays* L.). *Genetics*, *192*(2), 715–728. https://doi.org/10.1534/genetics.112.141473

Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2019). A practical introduction to the bootstrap: A versatile method to make inferences by using data-driven simulations. https://doi.org/10.31234/osf.io/h8ft7

Schmidt, P., Hartung, J., Rath, J., & Piepho, H. (2019). Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. *Crop Science*, *59*(2), 525–536. https://doi.org/10.2135/cropsci2018.06.0376

Tinker, N. A., Chao, S., Lazo, G. R., Oliver, R. E., Huang, Y. F., Poland, J. A., Jellen, E. N., Maughan, P. J., Kilian, A., & Jackson, E. W. (2014). A SNP genotyping Array for Hexaploid oat. *The Plant Genome*, *7*(3), plantgenome2014.03.0010. https://doi.org/10.3835/plantgenome2014.03.0010

Wang, X., Xu, Y., Hu, Z., & Xu, C. (2018). Genomic selection methods for crop improvement: Current status and prospects. *The Crop Journal*, *6*(4), 330–340. https://doi.org/10.1016/j.cj.2018.03.001

Ward, J. H. Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Werner, C. R., Gaynor, R. C., Gorjanc, G., Hickey, J. M., Kox, T., Abbadi, A., Leckband, G., Snowdon, R. J., & Stahl, A. (2020). How population structure impacts genomic selection accuracy in Cross-validation: Implications for practical breeding. *Frontiers in Plant Science*, *11*, 592977. https://doi.org/10.3389/fpls.2020.592977

Wickham, H., François, R., Henry, H. & Müller, K. (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.7. https://CRAN.R-project.org/package=dplyr

Xu, Y., Liu, X., Fu, J., Wang, H., Wang, J., Huang, C., Prasanna, B. M., Olsen, M. S., Wang, G., & Zhang, A. (2020). Enhancing genetic gain through genomic selection: From livestock to plants. *Plant Communications*, *1*(1), 100005. https://doi.org/10.1016/j.xplc.2019.100005

Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., Vicente, F. S., Olsen, M., Prasanna, B. M., Crossa, J., Yu, H., & Zhang, X. (2017). Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations. *Frontiers in Plant Science*, *8*, 1916. https://doi.org/10.3389/fpls.2017.01916

Zhao, Y., Gowda, M., Longin, F. H., Würschum, T., Ranc, N., & Reif, J. C. (2012). Impact of selective genotyping in the training population on accuracy and bias of genomic selection. *Theoretical and Applied Genetics*, *125*(4), 707–713. https://doi.org/10.1007/s00122-012-1862-2

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Sørensen, E. S., Jansen, C., Windju, S., Crossa, J., Sonesson, A. K., Lillemo, M., & Alsheikh, M. (2023). Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa*). *Plant Breeding*, *142*(1), 41–53. https://doi.org/10.1111/pbr.13061

# Paper II

# Identification of haplotypes associated with resistance to *Fusarium graminearum* in spring oat (*Avena sativa* L.).


## Authors' name

Espen Sannes Sørensen[1,2], Susanne Windju[1], Constantin Jansen[1], Anna Kristina Sonesson[3], Morten Lillemo[2], Muath Alsheikh[1,2]

## Institutes

[1]Graminor AS, Ridabu, Norway.

[2]Department of Plant Science, Norwegian University of Life Sciences, Ås, Norway.

[3]Nofima AS, Ås, Norway.

## Correspondence

Espen Sannes Sørensen. E-mail: espen.sorensen@nmbu.no.


## Key message

In this study 15 QTL-regions were associated to FHB resistance. Of these, five were validated in independent breeding populations showing consistent and significant effect. The presence of the resistance alleles was analyzed in a set of older and modern Nordic varieties.


## Keywords

*Fusarium graminearum; Genome-wide association study; QTL-regions; DON accumulation; Germination ability.*


## Conflict of interest

All authors have no conflict of interest to declare that are relevant to the content of this article.

## Contribution of authors

ESS analysed data from field trials, organized the germination analysis and performed the GWAS analysis, MA provided funding, CJ provided management of staff for phenotyping DON and field trials at Staur, SW, ML and AKS provided support in GWAS analysis, ESS wrote the manuscript, and MA, SW, ML, AKS and CJ participated in revising the scientific work and writing.

## Running title

Haplotypes associated with Fusarium resistance in oat.


## Acknowledgements

This research was funded by the Research council of Norway (The research funds for agriculture and food industry -FFL/JA) and Graminor ((NFR # 286606). The authors wish to thank Graminor for providing the material and facilities for this research, and the technical staff at Graminor and NMBU for helping in field experiments and data collection.

# Abstract

Fusarium head blight (FHB) is the predominant disease in oat in Norway caused by the fungus *Fusarium graminearum*. It causes yield loss, reduced seed quality, reduced germination ability and accumulation of deoxynivalenol (DON). The FHB resistance is quantitative, and most genes have small effect. Markers with verified effect in the breeding program could further enhance the resistance breeding. This study aims to use a large and diverse population of 541 lines to identify quantitative trait loci (QTL) associated to FHB resistance in a genome wide association study (GWAS) and verify their effect in independent breeding material. The material has been tested in six environments over three years and two locations in spawn inoculated and mist irrigated disease trials. The traits tested were germination ability and DON accumulation. A total of 15 significant QTL-regions were detected across 12 different linkage groups. Haplotypes for each region was constructed and the effect of the alleles in each environment were calculated, which identified the most likely resistant and susceptible alleles. Five QTL-regions were validated showing consistent effect in the GWAS population and the breeding material. Stacking of the resistant alleles of these regions from zero to five showed significant decrease in DON values and increased germination ability. The haplotype information of a set of historical and modern Nordic varieties were analysed, and the results could be used to select parents for future crossings. The validated haplotypes from this study can be used either to do marker assisted selection (MAS) or improve genomic prediction models in breeding programs.

# 1. Introduction

Oats is one of the most important cereal crops in Norway and is considered a good break crop in rotation with barley and wheat, as they have few diseases in common (Abrahamsen *et al.,* 2016). But one disease that infects all crops and pose major economic challenges, especially in oat, is Fusarium head blight (FHB) (Bernhoft *et al.,* 2013). FHB is caused by a wide range of *Fusarium* species. In Norway, *Fusarium graminearum* stands as the main species responsible for FHB in oats (Hofgaard *et al.,* 2016). This disease reduces yield, seed quality, germination capacity (Tekle *et al.,* 2013) and produces the mycotoxin deoxynivalenol (DON). DON is known to induce inflammation in the intestines in both humans and animals (Kang *et al.,* 2019). To mitigate health risks, Norway and the EU have set a threshold of 1.75 ppm DON for unprocessed oat (Commission regulation, 2006).

The need for resistant oat varieties in Norway is underscored by three key factors. Firstly, the temperate and humid conditions during Norwegian summers create favorable conditions for fungal growth and spore production of *F. graminearum* (Xu *et al.,* 2008). Secondly, the available fungicides have limited effect, providing only 30-50 % reduction in DON content when applied at the optimal timing (Felleskjøpet Agri SA, 2023). Thirdly, *F. graminearum* is shown to have the ability to sexually reproduce in Norway (Aamot *et al.,* 2015), which heightens the risk that the pathogen adapts and gains resistance to fungicides which could improve aggressiveness, virulence, and mycotoxin production (Becher *et al.,* 2010, de Chaves *et al.,* 2022).

Fusarium resistance can be divided into five primary classes (Hautsalo *et al.,* 2018); (i) resistance to initial infection (Schroeder & Christensen, 1963), (ii) resistance to disease spread, (iii) resistance to kernel infection, including germination ability (Mesterházy, 1995), (iv) tolerance (Mesterházy, 1995), and (v) resistance to mycotoxin accumulation (Miller *et al.,* 1985, Mesterházy *et al.,* 1999). Additionally, there are several passive avoidance mechanisms which exhibit strong correlation with FHB resistance. Two key mechanisms are plant height (PH), where taller plants increase the distance from the initial conidia spores from the soil to the heads (Hautsalo *et al.,* 2020), and days to heading (DTH), which affects the risk of plants flowering at the time of high disease pressure. (Tekle *et al.,* 2018).

Previous investigations that aimed to detect resistance QTL in oat found several QTL focusing on the traits FHB severity, DON accumulation and Germination ability (He *et al.,* 2013, Bjørnstad *et al.,* 2017 & Haikka *et al.,* 2020). The present study aims to identify and validate QTL along with linked markers for implementation in marker-assisted selection. This will be achieved through a comprehensive genome-wide association study (GWAS) focusing on DON accumulation and germination percentage. The study employs a large and diverse germplasm, subjected to disease trials conducted across

multiple environments. Furthermore, the investigation seeks to determine the presence of QTL in independent breeding populations.

# 2. Materials and methods

## 2.1. Materials

### 2.1.1. GWAS panel

The GWAS panel used in this study consisted of 541 oat lines and cultivars selected from a larger pool of 1,124 by using a combination of selection strategies evaluated by Sørensen *et al.* (2023). These strategies were based on marker information and was used to ensure that the GWAS panel had high relationship to the breeding material (Akdemir *et al.,* 2015), and high genetic diversity (Franco *et al.,* 2005). Additionally, lines that exhibit resistance and susceptibility based on historical data from Graminor's databases were included. Lines with very long and short straw as well as naked varieties were excluded to avoid association between the traits and FHB resistance. In total, 440 lines were from Norway, 40 from Sweden, 26 from Germany, nine from Netherlands, nine from Finland, eight from Canada, five from USA, two from Austria, one from Denmark and one from Australia (Supplementary Table S1).

### 2.1.2. Validation populations

Three breeding populations (V1, V2 & V3) were used to validate significant markers from the GWAS analysis. V1 and V2 consisted of 242 $F_{10}$ breeding lines from Graminor from the years 2020 and 2021 respectively. $F_{10}$ lines were used as they have not yet been selected for Fusarium resistance. V3 consisted of 230 lines where 112 were $F_{10}$ lines from 2022, 88 were $F_{11}$ from V2, 22 were $F_{12}$ from V1 and 8 were $F_{13}$ from a breeding population from 2019.

## 2.2. Field trials

Field experiments with the GWAS panel were conducted for three years from 2020 to 2022 in two locations, Vollebekk (59.66°N, 10.75°E) and Staur (60.73°N, 11.10°E). The GWAS panel trials were abbreviated as 20S, 20V, 21S, 21V, 22S & 22V for the individual years and locations, and Ov for the overall trial analysis. V1, V2 and V3 were tested in one year each, 2020, 2021 and 2022 respectively in both locations. The experimental design of the GWAS panel, V1 and V2 was alpha lattice with two replicates and sub-block size of 5. The experimental design of V3 was randomized complete block with two replicates (RCBD). All material were sown in spawn inoculated and mist irrigated disease trials (Tekle *et al.,* 2018).

## 2.3. Phenotyping

All plots were scored for DTH as the number of days from sowing to at least 50% of the heads had emerged, and PH as the number of centimeters from the ground to the top of the plants. DON was measured on milled seed samples with husks as parts per million (ppm) at Graminor with an Agraquant Deoxynivalenol Plus (0.25/5), 96 Wells ELISA kit developed by Romer Labs Ltd. Germination ability was measured at Graminor as percentage of germinated seeds (GP) using the "between paper" method described in point 5.6.2.1.1 of the ISTA protocol (ISTA, 2021). Two replicates of 50 seeds were used per plot. Plastic bags were used to retain moisture. Sample were stored in 5-10°C for 7 days, and approximately 20°C for 6-8 days before analysis. The papers used were of size 220 X 400 mm and 200 x 400 mm with a capillary capacity of 80 mm/10 min. DON and GP were collected for all trials except GP for V3 Staur in 2022.

## 2.4. Phenotypic data analysis

The Best linear unbiased estimators (BLUE) were calculated for each genotype in each trial and across trials (overall), with the META-R software (Alvarado *et al.,* 2016) using the models listed in Supplementary Table S4. The GWAS analysis was performed on data from single trials and overall, while the analysis for V1-3 only the overall values were used. Outliers were not excluded from the dataset. The BLUE DON values (Figure 1a) were log transformed to obtain close to normally distributed values (Figure 1b) (West, 2022) for the GWAS analysis. Both BLUE values of GP (Figure 1d) and logDON values were adjusted for the effect of DTH and PH to avoid false association to the correlated traits, by performing a regression analysis with GP and DON as response variables separately, and DTH and PH as explanatory variables (Nannuru *et al.,* 2022). The resulting adjusted DON (AdjDON) (Figure 1c) and GP (AdjGP) (Figure 1e) have zero correlation with PH and DTH and $R^2$ of 0.95 and 0.94 with the unadjusted logDON and GP overall values. AdjDON and AdjGP are used as phenotypes in the subsequent GWAS analysis.

## 2.5. Genotyping and data preparation

All lines of the GWAS panel were genotyped with a customized, unpublished 20K SNP-chip containing 18,598 markers including all markers from the publicly available 6K SNP-chip (Tinker *et al.,* 2014). The genetic data were filtered with a threshold of 10% for missing values/heterozygotes and 5% MAF based on the GWAS panel, resulting in 3071 polymorphic markers. V1-3 were genotyped using a different customized 7K-SNP chip (Polley *et al.,* 2023) containing 6,642 markers where 6587 were the most polymorphic markers from the Nordic 20K SNP-chip. An unpublished consensus map was used to assign markers to linkage groups (LG) representing the 21 oat chromosomes. The map is an

updated version of the genetic map developed by Chaffin *et al.* (2016), updated with six biparental populations from the Nordic breeding programs. Linkage disequilibrium (LD) between each pair of markers was calculated using the TASSEL statistical software (Bradbury *et al.,* 2007).

## 2.6. GWAS analysis

The GWAS was performed for the traits AdjGP and AdjDON for individual environments and overall values with the "farmCPU" method (Liu *et al.,* 2016) in the GAPIT3 package (Wang & Zhang. 2021) with the R statistical software (R Core Team, 2022). FarmCPU were chosen over MLMM based on the QQ-plot results (Supplementary Figure S2). FarmCPU has proven more efficient compared to other models for *F. graminearum* resistance in wheat (Nannuru *et al.,* 2022). This method is considered statistically powerful, it avoids overfitting and reduce the number of false positive and negatives compared to other models (Kaler *et al.,* 2020). The GWAS was not corrected for population structure as a "model selection" approach in GAPIT revealed that zero principal components were optimal, and a visual inspection of a PCA plot of the marker data supported this (Supplementary Figure S1). Markers with FDR adjusted p-value less than 0.05 were considered significant and were calculated as $\frac{p*n}{r}$ where p is the p-value, n is the number of markers tested, and r is the rank of the marker from lowest to highest p-value (Benjamini & Hochberg. 1995).

## 2.7. QTL-regions and Haplotype analysis

QTL-regions were determined as the significant markers from the GWAS that were on the same LG and in significant LD with each other. Markers that were not in LD on the same LG were considered as a separate region. Regions with only one marker from single experiments were not analysed further, while single marker detected using the overall data were retained. Haplotypes from each region were then formed by adding the significant SNP-marker information in the region together from lowest to highest centimorgan position on the consensus map.

The haplotype alleles for each QTL region were analysed in each environment of the GWAS panel and the overall values of V1-3 with a Games-Howell simultaneous pairwise comparison test (Games & Howell, 1976) using Minitab statistical software (Minitab, 2021) on the unadjusted DON and GP. Higher number allelic variations in the regions equals higher number of comparisons. This was done to determine which allele had significantly ($p < 0.05$) lower DON and higher GP than others in each environment. The ones that most frequently showed to significantly resistant than other in the same region were designated the resistant allele, and similar for the susceptible allele. Resistant alleles that showed consistently lower DON and higher GP than the susceptible allele across environments and populations were considered validated. The effects were summarized as percentage difference in

DON to get similar scales across environments. This was calculated as the difference in mean divided by the average between the alleles multiplied by 100.

### 2.8. Allele stacking and analysis of historical varieties

QTL-regions validated from the haplotype analysis were further analysed for their additive effect on the phenotypes DON and GP. Games-Howell tests were performed to see if there were significant decrease of DON and increase in GP with increased number of resistant alleles. This was done only in the overall GWAS panel. DTH and PH were also analysed to see if the resistant alleles had any effect on these traits.

To evaluate the trend of resistant alleles through time, and possibly identify resistance sources for crossings, a set of 74 varieties were selected to represent the most important Norwegian material from the last century (Supplementary Table S3). Each variety were given a year as an approximation of the year they became inbred lines determined as six years before release or six years after crossing depending on available information about the variety. The dataset was assembled in groups of different time periods, and the number of genotypes carrying different number of resistant and susceptible alleles were summarized and averaged for each time period.

### 2.9. BLAST search

Markers in the validated QTL-regions were BLASTED against the reference genome of OT3098 (PepsiCo, 2021) in the GrainGenes database (Yao *et al.,* 2022). This reference genome was chosen because it contains more annotations with information on gene function than other reference genomes. When the markers got more than one chromosome hit the lowest average E-value among the markers determined which chromosome they were assigned to. The region between the markers and 10 Mbp in each side were investigated for annotated genes described with an effect on disease resistance.

## 3. Results

### 3.1. Phenotypic correlations

PH showed significant positive Pearson correlation to GP in Staur and Vollebekk 2020, while the correlation was significantly negative in 2022 (Table 1). The overall values were non-significant and close to zero, as was most other trials. DTH showed significant positive correlation to DON in four experiments, significant negative correlation with GP in three, with r values of 0.2 and -0.25 respectively for the overall values of DON and GP. DON showed significant negative correlation to GP

in five of six experiments and the overall values, with r values between -0.13 in Vollebekk 2020 and -0.69 in Staur 2022.

***Table 1**: Pearson correlation between the BLUE values of traits plant height (PH), days to heading (DTH), DON accumulation in ppm (DON) and Germination percentage (GP) from each experiment (20S, 20V, 21S, 21V, 22S, 22V) and the overall values (Ov) of the GWAS panel and their level of significance level with α > 0.05 = *, > 0.01 = ** and > 0.001 ***.*

| | **20S** | **20V** | **21S** | **21V** | **22S** | **22V** | **Ov** |
|---|---|---|---|---|---|---|---|
| *PH v DON* | -0.23*** | 0.10* | 0.03 | -0.04 | 0.05 | -0.02 | -0.05 |
| *PH v GP* | 0.10* | 0.15*** | -0.04 | -0.05 | -0.03 | -0.20*** | -0.06 |
| *DTH v DON* | 0.00 | 0.29*** | 0.15*** | 0.33*** | 0.05 | 0.14** | 0.21*** |
| *DTH v GP* | -0.03 | 0.01 | -0.16*** | -0.34*** | 0.08* | -0.34*** | -0.25*** |
| *DON v GP* | -0.34*** | -0.13** | -0.02 | -0.28*** | -0.69*** | -0.49*** | -0.52*** |
| *DTH v PH* | 0.30*** | 0.29*** | 0.38*** | 0.10* | 0.29*** | 0.25*** | 0.30*** |

Trial statistics are shown in the supplementary Table S5. The heritability was relatively high for both DON and GP in all environments except for the GP of 2020 Vollebekk, while the overall heritability were 0.79 for DON and 0.62 for GP which were higher than any individual experiment. The genotype effect was significant below 0.05 for both traits in all experiments. The overall values of AdjDON and AdjGP were close to normally distributed (Figure 1c and e).

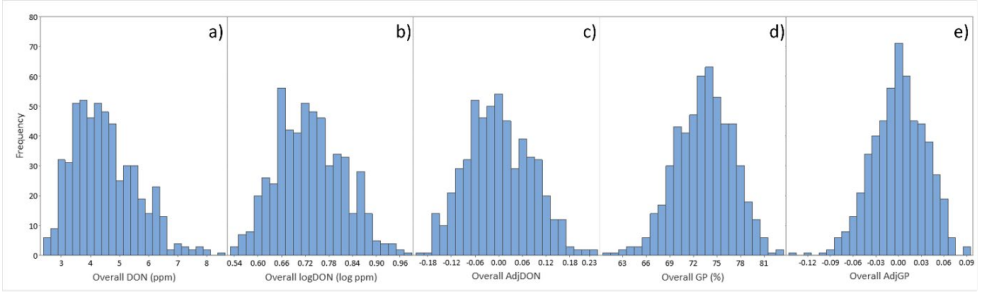



***Figure 1**: Distribution of overall phenotypic values of the GWAS panel for the traits DON values in ppm (a), logtransformed DON values(logDON) (b), logtransformed DON values adjusted for effect of days to heading and plant height (AdjDON) (c), Germination percentage (GP) (d) and Germination percentage adjusted for effect of Days to heading and plant height (AdjGP) (e) with number of genotypes in the Y-axis and phenotypic values in the X-axis.*

## 3.2. GWAS analysis and QTL-regions.

A total of 48 significant markers for FHB resistance were detected, 24 for adjDON, 22 for adjGP and two for both. Six LGs (3C, 4C, 8A, 10D, 11A and 14D) had only one significant marker for one trait from a single experiment and were not analyzed further. Four unmapped markers were detected, three of them were in LD with the significant markers on LG 1C. The unmapped markers were analyzed further. Two markers on 19A and 21D were excluded from further analysis as they were not in significant LD with the other markers in the LGs. The remaining 36 markers were assembled into 15 QTL-regions (Table 2). The markers on 5C, 7C-17A and 18D were split into separate QTL-regions as there were no significant LD between them.

***Table 2****: QTL-regions determined from significant markers from either single environment or overall GWAS for the traits log-transformed DON and GP adjusted for effect of days to heading and plant height (AdjDON& AdjGP). The table includes the size of the region in cM (Span), number of SNPs detected (n-SNP), the number of experiments they were detected in (n-exp), the traits associated to the region (Trait) and the range of -10log-transformed p-value of the markers (-LOG(p)). The R-HT is the most frequently resistant allele in the number of significance tests (R-HT tests), and the most frequent susceptible allele (S-HT) from the same number of tests (S-HT)*

| QTL/LG | Span(cM) | n-SNP | n-Exp | Trait | -LOG(p) | R-HT | R-HT tests | S-HT | S-HT tests |
|---|---|---|---|---|---|---|---|---|---|
| 1C | 89-92.8 | 4 | 3 | AdjGP & AdjDON | 3.84-14.22 | ACAG | 19/24 | GTCA | 10/24 |
| 2C | 74.2 | 2 | 2 | AdjGP | 3.86-6.04 | TC | 3/4 | CA | 3/4 |
| 5C-1 | 48.7-52.4 | 3 | 2 | AdjGP & AdjDON | 4.07-4.73 | CAC | 5/10 | TGA | 6/10 |
| 5C-2 | 86.6 | 1 | 1 | AdjGP | 4.42 | C | 6/6 | T | 6/6 |
| 6C | 48.7-65.3 | 2 | 2 | AdjGP | 5.3-6.43 | CA | 2/3 | CG | 2/3 |
| 7C-17A-1 | 28.7-72 | 3 | 3 | AdjGP & AdjDON | 3.69-4.94 | CGA | 18/25 | TGG | 11/25 |
| 7C-17A-2 | 74.9 | 1 | 1 | AdjDON | 4.32 | G | 1/1 | A | 1/1 |
| 9D | 21.3-35.4 | 3 | 4 | AdjGP & AdjDON | 5.2-5.42 | CTC | 11/17 | TCT | 17/17 |
| 12D | 51-61.1 | 3 | 2 | AdjGP & AdjDON | 4.15-4.93 | GGC | 19/23 | TGT | 12/23 |
| 15A | 87.9 | 1 | 1 | AdjGP & AdjDON | 5.12 | A | 4/4 | G | 4/4 |
| 18D-1 | 25-45.9 | 3 | 3 | AdjDON | 4.8-6.54 | AGT | 10/10 | GTC | 10/10 |
| 18D-2 | 97.7-99.5 | 2 | 2 | AdjGP & AdjDON | 4.33-4.87 | TG | 4/5 | CG | 3/5 |
| 19A | 30-54.1 | 2 | 2 | AdjDON | 5.11-5.31 | AA | 3/5 | CG | 3/5 |
| 20D | 39.6-70.9 | 2 | 2 | AdjGP | 4.09 | TG | 4/9 | CT | 5/9 |
| 21D | 41.9-87.8 | 4 | 3 | AdjGP & AdjDON | 4.02 | AACC | 10/13 | AACT | 5/13 |

### 3.3. Haplotype analysis and ANOVA validation.

The haplotype analysis revealed that of the 15 QTL-regions detected in the GWAS, five (1C, 7C-17A-2, 9D, 12D and 18D-1) showed a consistent effect and significant difference in unadjusted DON and GP in comparisons between the resistant and susceptible alleles for the overall phenotypes and at least two individual experiments. They also showed the same effect in at least two of three validation populations (Table 3). Of these five, 9D stands out positively, because the difference between the resistant and susceptible alleles were highly significant for DON in all experiments in the GWAS panel and V3 with an effect of 12-22% reduction in DON content.

Table 2 shows which alleles were the most frequently resistant (R-HT) and susceptible (S-HT) in the significance tests. Some QTL-regions had low number of significant comparisons (2C, 6C, 7C-17A-2, 15A, 18D-2 and 19A). Others had low frequency of resistant or susceptible alleles (1C, 5C-1, 7C-17A-1, 12D, 20D and 21D). Some combinations of markers were missing in the populations; hence all allelic variations were not tested in this study.

***Table 3**: List of the five most significant QTL-regions and the difference between the resistant and susceptible alleles listed in Table 2 in all environments of the GWAS panel (20S, 20V, 21S, 21V, 22S, 22V and Ov) and the validation populations (V1-3). DON is shown as percentage difference between alleles ((difference in mean / average between groups) × 100)) while GP is given as the difference in mean percentage points. α < 0.05 = *, < 0.01 = ** and < 0.001 ****

| QTL | 20S | 20V | 21S | 21V | 22S | 22V | Ov | V1 | V2 | V3 |
|---|---|---|---|---|---|---|---|---|---|---|
| **DON** | | | | | | | | | | |
| **1C** | -8 | -5 | -8 | -6 | -24*** | -11* | -12* | -43 | 31 | -14 |
| **7C-17A-1** | -17 | -31* | -18 | -26 | -29* | -19* | -23** | -32 | -6 | -33*** |
| **9D** | -22*** | -19*** | -14** | -14** | -19*** | -12*** | -16*** | -16 | -17 | -15* |
| **12D** | -15 | -26** | -19* | -8 | -22** | -15** | -17*** | NA | -15 | -14 |
| **18D-1** | -9 | -9* | -11** | -15** | -14** | -13*** | -12*** | -2 | -16* | -14** |
| **GP** | | | | | | | | | | |
| **1C** | 1.8 | 0.1 | 1.1 | 1.1 | 6.9*** | 7.9*** | 3.1*** | 1.5 | -1.8 | 5.6 |
| **7C-17A-1** | 2.2 | 2.9 | 0.1 | 3.8* | 4.9 | 9.4*** | 4.2*** | 4.0 | 0.6 | 12.2 |
| **9D** | 2.9*** | 0.9 | 0.5 | -0.2 | 3.1* | 2.7** | 1.6*** | 0.5 | 2.0 | 4.6* |
| **12D** | 5.3*** | 0.6 | -0.1 | -0.7 | 3.7* | 4.2** | 2.1*** | NA | 2.7 | 5.2 |
| **18D-1** | 2.4** | 0.7 | 0.2 | 0.1 | 2.9** | 3.2*** | 1.6*** | -0.2 | 2.2* | 4.9** |

## 3.4. Allele stacking.

The number of validated resistant alleles showed a reduction from a mean DON of 5.47 ppm with zero resistant alleles to 3.71 ppm with five (Figure 2a). GP increased from 71.5% for the group with zero resistant alleles to 76.4% for the one with five (Figure 2b). Both DON and GP had a linear increase in resistance from zero to five alleles, and $R^2$ between phenotype and number of alleles were 0.17 and 0.13 for overall DON and GP respectively which were both highly significant with $p < 0.001$ (data not shown). There was no significant reduction in DON from zero to one resistant allele, but there was a significant reduction from one to two, and two to three. But no significant reduction from three to four or five. For GP there were significant increase from zero to one, one to two, and two to three alleles, and from three to five, but not from three to four or four to five. The number of resistant alleles did not affect the PH (Figure 2d), but for DTH there were significant reductions from one, two and three to five alleles with approximately one day difference between one and five (Figure 2c).

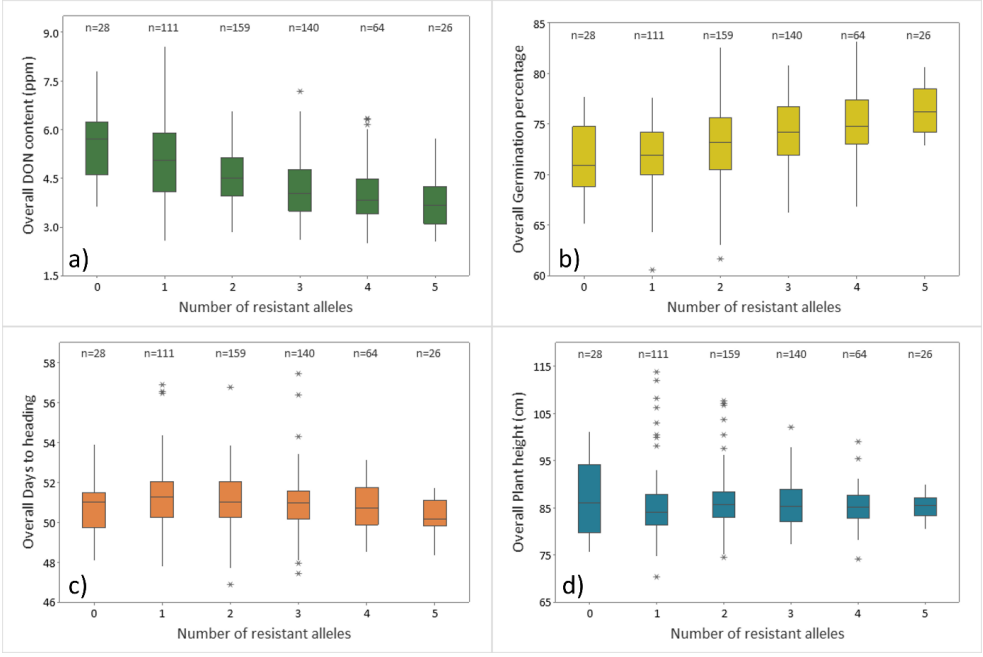



***Figure 2**: Boxplots of overall values from the GWAS panel for the traits DON accumulation in ppm (DON) (a), Germination percentage (GP) (b), days to heading (DTH) (c) and plant height (PH) (d) with phenotypes in the Y-axis and the number of resistant alleles from Table 3 in the X-axis. n equals the number of lines that carry different number of alleles.*

## 3.5. Analysis of historically important varieties.

The five validated haplotypes from Table 3 were analysed in a subset of material that includes older important Nordic varieties from 1895 to 1999 and modern Norwegian varieties from 2001 to 2017. The full list of varieties and their haplotypes are listed in Supplementary Table S3. The analysis showed that the average number of resistant alleles (Figure 3a) increased from 1.3 in the period of 1895-1920 to 3.1 in 2001-2009 with a small dip down to 2.7 in 2011-2017. In the oldest varieties 70% had one or less resistant alleles, while for most of the modern varieties none had less than two. The number of susceptible alleles (Figure 3b) has been reduced from an average of 1.7 in the period 1895-1920 to 0.3 in 2001-2009 with a small increase to 0.4 in 2011-2017. In the oldest varieties 60% had two or more susceptible alleles while 65% of the varieties from 2001-2017 had zero. There was almost no difference in resistant allele frequencies between the periods 1991-1999 and 2001-2009 with an increase in average of 0.2, while the number of susceptible alleles was reduced with an average of 0.5 in the same period.

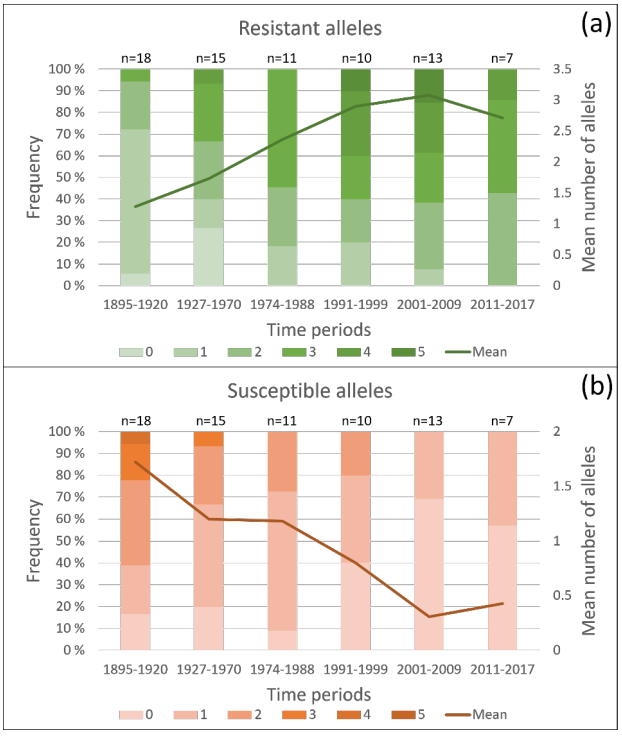



***Figure 3**: Number of resistant (a) and susceptible (b) alleles listed in Table 3 present in a selection of historical Nordic lines from 1891 to 2017. Each column represents a time period of approximately when the cultivars became inbred lines. The left Y-axis is the frequency in percentage and the right is*

*the average numbers of alleles. The different colours are the different number of alleles, and the solid line is the mean number of alleles in each period.*

### 3.6. Candidate genes

Based on BLAST searches in the OT3098 reference genome (PepsiCo, 2021) several disease resistance related genes were identified within the five QTL-regions. The QTL-regions 1C, 7C-17A-1, 9D, 12D and 18D-1 spanned 21.2, 47.3, 4.1, 23.2 and 4.2 Mbp, respectively. The genes found in these regions are named *RGA1-5*, *RPM1*, *RPS2*, *Pik-1, 2* & *6*, *RPP13*, *At3g14460*, *At1g50180*, *EDR2* and *EDR4*. QTL region 1C contained 20 candidate genes which is the highest among the regions with 16 of them close to the first marker. Region 7C-17A-1, 12D, 9D and 18D-1 contained 16, 11, five and three candidate genes respectively. The third SNP in the 7C-17A-1 haplotype did not match the same chromosome as the other two but a homologue, so it appears to not be part of the same QTL region. The full list of candidate genes found in each LG is listed in Supplementary Table S2.

## 4. Discussion

### 4.1. Quality of data

The phenotypic data showed a highly significant negative Pearson correlation between DON and GP (Table 1) which is expected from previous studies (Tekle *et al.,* 2012, Tekle *et al.,* 2018, Hautsalo *et al.,* 2020). It was expected that the agronomic traits DTH and PH would be highly correlated to DON and GP (Moreno-Amores *et al.,* 2020), which was the case for DTH but not for PH. This could be because the tallest varieties are also oldest and more susceptible either due to few resistant alleles or late heading. The disease pressure varied between the experiments, but this did not affect the difference in ranking as the Pearson correlation between the trials remained high for DON. The CV values for DON ranged between 20 and 43% for the individual experiments which shows that there is a large variation in DON contents within the trials. But given that the CV values are smaller than similar studies in oat (32.6-63.2) (Yan *et al.,* 2010, Haikka *et al.,* 2020), and that the heritability measures were relatively high means that the experiments were successful which ensured a good expression of genetic variation. GP had generally lower heritability than DON.

### 4.2. Comparison with previous QTL studies

The results from this study can be compared to three mapping studies on resistance to *F. graminearum* in oat. The first (He *et al.,* 2013) used two biparental populations based on crosses between accessions derived from *Avena sterillis*, North American varieties and a Norwegian variety. It detected QTL in three LG that might correspond to our results on 5C, 7C-17A and 9D. The second study (Bjørnstad *et al.,* 2017) used mostly North American breeding lines with a few varieties from

Northern Europe. They detected QTL on 6C, 7C-17A and 9D that might correspond to our results. The third (Haikka *et al.,* 2020) used most Finnish breeding lines and varieties. They did not detect any significant markers, but they did detect low p-values in regions on 1C and 9D potentially corresponding to our results. To summarize, there is ample evidence to support a major QTL for FHB resistance on 9D and 7C-17A based on our results and previous research. The remaining QTL-regions appear to be novel, except for regions on 1C, 5C and 6C which might have been detected in previous studies. More precise mapping of physical position of markers is needed to properly evaluate the overlapping regions in these studies.

A recent study from Norway suggests that the ranking of cultivars is partly similar in accumulation of DON and HT2+T2 which is the mycotoxin produced by *Fusarium langsethiae* (Hofgaard *et al.,* 2022). There was however evidence to suggests that some resistance is specifically associated to DON or HT2+T2. A recent GWAS study on *F. langsethiae* detected a significant QTL on LG 14D (Isidro-Sanchez *et al.,* 2020). This QTL were not significant in our study and could therefore be considered specific to *F. langsethiae.*

## 4.3. Candidate genes

There were several different disease resistance genes within the QTL-regions, and some of them were grouped together in smaller clusters. Networks of QTL as a defense response to *F. graminearum* have been previously reported in bread wheat (Kugler *et al.,* 2013) and durum wheat (Sari *et al.,* 2019) which also find the same genes as this study, specifically RGA1, RGA2, RGA4, RPP13 and At3g14460. A study of the *F. graminearum* fungus indicated the presence of *AVR-Pik* effector genes that helps in the infection of plant tissues (Hao *et al.,* 2020). A different study claimed that the *Pik-1* and *Pik-2* genes work as defence genes against these effectors in rice (Maidment *et al.,* 2023) Both of which were present in the QTL-regions of this study.

A recent study on possible DON detoxification genes in oat found two candidates named AsUGT1 and AsUGT2 (Khairullina *et al.,* 2022). These are UDP-Glucotransferases and were annotated in the Sang reference genome (Kamal *et al.,* 2022) found in the GrainGenes database (Yao *et al.,* 2022). The location of these genes does match one of the possible physical positions of the QTL-region on 1C from the BLAST search, but not the same as the one where 20 resistance genes were annotated.

## 4.4. Implications for resistance breeding

The stacking of the five resistant alleles indicates that a plateau of resistance is reached in the material with the stacking of three of validated resistance alleles, as further significant increase of resistance requires more than one allele. The analysis of the Nordic material shows that the number

of resistant alleles were low in the oldest varieties (1885-1920) and increased to almost three (1990-1999). But it did not increase further when breeding for FHB resistance with inoculated trials started in Norway (2001-2009). But the number of susceptible alleles were reduced to almost zero. It is possible that the screening of material for fusarium resistance resulted in reduction in most susceptible material, and that breeding priorities changed to other traits like yield and quality. Increase in resistance before 2001-2009 probably did not come from targeted breeding against *F. graminearum*, although it is likely that highly susceptible genotypes would have been discarded due to visible symptoms in epidemic years. It is also likely that selection for yield and test weight in years with high disease pressure would have improved the FHB resistance. Future breeding strategies could be to select crosses based on the haplotype information provided in this study, use MAS to select for the QTL with largest effect in early generations and use genomic prediction to select for both small and large effect QTL using genome wide SNP arrays.

## 4.5. Conclusion

This study identified 15 significant QTL-regions involved in Fusarium resistance in oats, validated five that showed consistent effect across environments and populations and identified resistant and susceptible haplotype alleles. The additive effects of the five QTL-regions reduced the DON content by 38%. Several genes associated with resistance against *F. graminearum* in other crops were found to be located within these regions. Our study has confirmed that Fusarium resistance is made up of multiple QTL across the genome with varied effects. An approach to resistance breeding could be to use MAS to select for the QTL with larger effects in early generations and to develop genomic prediction models and make selections that include QTL with both large and small effects. The haplotype information provided in this study could also be used to select crosses for improved fusarium resistance.

# References


Aamot, H.U., Ward, T.J., Brodal, G., Vrålstad, T., Larsen, G. B., Klemsdal, S. S., Uhlig, S. & Hofgaard, I. S. (2015) Genetic and phenotypic diversity within the *Fusarium graminearum* species complex in Norway. *European Journal of Plant Pathology, 142*, 501–519. https://doi.org/10.1007/s10658-015-0629-4.

Abrahamsen, U., Brodal, G. & Waalen, W. (2016) Virkning av ulike forgrøder på neste års avling av hvete. NIBIO BOK 2 (1) Jord- og Plantekultur 2016, 106-111. URL: https://nibio.brage.unit.no/nibio-xmlui/handle/11250/2577326.

Akdemir, D., Sanchez, J. I., & Jannink, J. L. (2015) Optimization of genomic selection training populations with a genetic algorithm. *Genetics, selection, evolution : GSE*, *47*(1), 38. https://doi.org/10.1186/s12711-015-0116-6.

Alvarado, G., López, M., Vargas, M., Pacheco, A., Rodríguez, F., Burgueño, J. & Crossa, J. (2016) META-R (Multi Environment Trial Analysis with R for Windows.) Version 6.0. International Maize and Wheat Improvement Center.

Becher, R., Hettwer, U., Karlovsky, P., Deising, H. B., & Wirsel, S. G. (2010) Adaptation of Fusarium graminearum to tebuconazole yielded descendants diverging for levels of fitness, fungicide resistance, virulence, and mycotoxin production. *Phytopathology*, *100*(5), 444–453. https://doi.org/10.1094/PHYTO-100-5-0444.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bernhoft, A. Eriksen, G. S., Sundheim, L., Berntssen, M., Brantsæter, A. L., Brodal, G., Fæste, C. K., Hofgaard, I. S., Rafoss, T., Sivertsen, T. & Tronsmo, A. M. (2013) Risk Assessment of Mycotoxins. in Cereal Grain in Norway. *Opinion of the Scientific Steering Committee of the Norwegian Scientific Committee for Food Safety; VKM (Norwegian Scientific Committee for Food Safety):* Oslo, Norway, 2013. URL: https://vkm.no/english/riskassessments/allpublications/riskassessmentofmycotoxinsincerealgraininnorway.4.27ef9ca915e07938c3b2bad8.html.

Bjørnstad, Å., He, X., Tekle, S., Klos, K., Huang, Y-F, Tinker, N. A., Dong, Y. & Skinnes, H. (2017) Genetic variation and associations involving Fusarium head blight and deoxynivalenol accumulation in cultivated oat (*Avena sativa* L.). *Plant Breeding*, *136*, 620– 636. https://doi.org/10.1111/pbr.12502

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S. (2007) TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*(19), 2633-2635. https://doi.org/10.1093/bioinformatics/btm308.

de Chaves, M.A., Reginatto, P., da Costa, B.S., de Paschoal, R. I., Teixeira, M. L. & Fuentefria, A. M. (2022) Fungicide Resistance in *Fusarium graminearum* Species Complex. *Current Microbiology* 79, 62. https://doi.org/10.1007/s00284-021-02759-4.

Commission regulation (EC) No 1881/2006 of 19 December 2006. Setting maximum levels for certain contaminants in foodstuffs. Annex section 2: Mycotoxins, 2.4 Deoxynivalenol, 2.4.2 Unprocessed durum wheat and oats. http://data.europa.eu/eli/reg/2006/1881/2023-01-01.

Chaffin, A. S., Huang, Y.-F., Smith, S., Bekele, W. A., Babiker, E., Gnanesh, B. N., Foresman, B. J., Blanchard, S. G., Jay, J. J., Reid, R. W., Wight, C. P., Chao, S., Oliver, R., Islamovic, E., Kolb, F. L., McCartney, C., Mitchell Fetch, J. W., Beattie, A. D., Bjørnstad, Å., Bonman, J. M., Langdon, T., Howarth, C. J., Brouwer, C. R., Jellen, E. N., Klos, K. E., Poland, J. A., Hsieh, T.-F., Brown, R., Jackson, E., Schlueter,

J. A. & Tinker, N. A. (2016) A Consensus Map in Cultivated Hexaploid Oat Reveals Conserved Grass Synteny with Substantial Subgenome Rearrangement. *The Plant Genome, 9*(2), plantgenome2015.10.0102. https://doi.org/10.3835/plantgenome2015.10.0102.

Franco, J., Crossa, J., Taba, S. & Shands, H. A. (2005) Sampling strategy for conserving genetic diversity when forming core subsets. *Crop Science. 45*(3), 1035–1044. https://doi.org/10.2135/cropsci2004.0292.

Felleskjøpet Agri SA, Plantevernkatalogen 2023 (2023) Page 26. URL: https://ipaper.ipapercms.dk/TidRom/felleskjopet/plantevern/.

Games P. A., Howell J. F. (1976) Pairwise multiple comparison procedures with unequal N's and/or variances: A Monte Carlo study. *Journal of Educational Statistics*, 1, 113–125. https://doi.org/10.2307/1164979.

Haikka, H., Manninen, O., Hautsalo, J., Pietilä, L., Jalli, M. & Veteläinen M. (2020) Genome-wide Association Study and Genomic Prediction for *Fusarium graminearum* Resistance Traits in Nordic Oat (*Avena sativa* L.). *Agronomy, 10*(2), 174. https://doi.org/10.3390/agronomy10020174.

Hao, G., McCormick, S., Usgaard, T., Tiley, H., & Vaughan, M. M. (2020) Characterization of Three *Fusarium graminearum* Effectors and Their Roles During Fusarium Head Blight. *Frontiers in plant science*, *11*, 579553. https://doi.org/10.3389/fpls.2020.579553.

Hautsalo, J., Jalli, M., Manninen, O. (2018) Evaluation of resistance to *Fusarium graminearum* in oats. *Euphytica 214,* 139. https://doi.org/10.1007/s10681-018-2222-3

Hautsalo, J., Jauhiainen, L., Hannukkala, A., Manninen, O., Veteläinen, M., Pietilä, L., Peltoniemi, K. & Jalli, M. (2020) Resistance to Fusarium head blight in oats based on analyses of multiple field and greenhouse studies. *European Journal of Plant Pathology*, 158, 15–33. https://doi.org/10.1007/s10658-020-02039-0.

He, X., Skinnes, H., Oliver, R.E., Jackson, E. W. & Bjørnstad, Å. (2013) Linkage mapping and identification of QTL affecting deoxynivalenol (DON) content (*Fusarium* resistance) in oats (*Avena sativa* L.). *Theoretical and Applied Genetics,* 126, 2655–2670. https://doi.org/10.1007/s00122-013-2163-0.

Hofgaard, I. S., Aamot, H. U., Torp, T., Jestoi, M., Lattanzio, V. M. T., Klemsdal, S. S., Waalwijk, C., Van der Lee, T. & Brodal, G. (2016) Associations between *Fusarium* species and mycotoxins in oats and spring wheat from farmers' fields in Norway over a six-year period. *World Mycotoxin Journal 9*, 365-378. https://doi.org/10.3920/WMJ2015.2003.

Hofgaard, I. S., Brodal, G., Almvik, M., Lillemo, M., Russenes, A. L., Edwards, S. G., & Aamot, H. U. (2022). Different Resistance to DON versus HT2 + T2 Producers in Nordic Oat Varieties. *Toxins*, *14*(5), 313. https://doi.org/10.3390/toxins14050313.

International Rules for Seed Testing, Chapter 5, (2021) i–5-56 (64)

Isidro-Sánchez, J., D'Arcy Cusack, K., Verheecke-Vaessen, C., Kahla, A., Bekele, W., Doohan, F., Magan, N., & Medina, A. (2020) Genome-wide association mapping of Fusarium langsethiae infection and mycotoxin accumulation in oat (Avena sativa L.). *The plant genome*, *13*(2), e20023. https://doi.org/10.1002/tpg2.20023.

Kaler, A. S., Gillman, J.D., Beissinger, T. & Purcell, L.C. (2020) Comparing different statistical models and multiple testing corrections for association mapping in Soybean and Maize. *Frontier of Plant Science, 10*, 1794. https://doi.org/10.3389/fpls.2019.01794.

Kang, R., Li, R., Dai, P., Li, Z., Li, Y., & Li, C. (2019) Deoxynivalenol induced apoptosis and inflammation of IPEC-J2 cells by promoting ROS production. *Environmental pollution*, *251*, 689–698. https://doi.org/10.1016/j.envpol.2019.05.026.

Kamal, N., Tsardakas Renhuldt, N., Bentzer, J. *et al.* (2022) The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature* 606, 113–119. https://doi.org/10.1038/s41586-022-04732-y.

Khairullina, A., Tsardakas Renhuldt, N., Wiesenberger, G., Bentzer, J., Collinge, D. B., Adam, G., & Bülow, L. (2022) Identification and Functional Characterisation of Two Oat UDP-Glucosyltransferases Involved in Deoxynivalenol Detoxification. *Toxins*, *14*(7), 446. https://doi.org/10.3390/toxins14070446.

Kugler, K.G., Siegwart, G., Nussbaumer, T. Ametz, C., Spannagl, M., Steiner, B., Lemmens, M., Mayer, K. F. X., Buerstmayr, H. & Schweiger, W. (2013) Quantitative trait loci-dependent analysis of a gene co-expression network associated with Fusarium head blight resistance in bread wheat (*Triticum aestivum*L.). *BMC Genomics, 14*, 728. https://doi.org/10.1186/1471-2164-14-728

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016) Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS genetics*, *12*(2), e1005767. https://doi.org/10.1371/journal.pgen.1005767.

Maidment, J. H. R., Shimizu, M., Bentham, A. R., Vera, S., Franceschetti, M., Longya, A., Stevenson, C. E. M., De la Concepcion, J. C., Białas, A., Kamoun, S., Terauchi, R. & Banfield, M. J. (2023) Effector target-guided engineering of an integrated domain expands the disease resistance profile of a rice NLR immune receptor *eLife, 12*, e81123. https://doi.org/10.7554/eLife.81123.

Mesterházy, A. (1995) Types and components of resistance to *Fusarium* head blight of wheat. *Plant Breeding, 114*: 377-386. https://doi.org/10.1111/j.1439-0523.1995.tb00816.x.

Mesterházy, Á., Bartók, T., Mirocha, C.G. and Komoróczy, R. (1999) Nature of wheat resistance to *Fusarium* head blight and the role of deoxynivalenol for breeding. *Plant Breeding, 118*: 97-110. https://doi.org/10.1046/j.1439-0523.1999.118002097.x.

Miller, J. D., Young, J. C., & Sampson, D. R. (1985) Deoxynivalenol and Fusarium Head Blight Resistance in Spring Cereals. Journal of Phytopathology, 113: 359-367. https://doi.org/10.1111/j.1439-0434.1985.tb04837.x.

Minitab 21 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

Moreno-Amores, J. Michel, S., Miedaner, T., Longin, C. F. H. & Buerstmayr, H. (2020) Genomic predictions for Fusarium head blight resistance in a diverse durum wheat panel: an effective incorporation of plant height and heading date as covariates. *Euphyitca, 216,* 22. https://doi.org/10.1007/s10681-019-2551-x.

Nannuru, V.K.R., Windju, S.S., Belova, T., Dieseth, J, A., Alsheikh, M., Dong, Y., McCartney, C. A., Henriques, M. A., Buertsmayr, H., Michel, S., Meuwissen, T. H. E. & Lillemo, M. (2022). Genetic architecture of fusarium head blight disease resistance and associated traits in Nordic spring wheat. *Theoretical & Applied Genetics 135*, 2247–2263. https://doi.org/10.1007/s00122-022-04109-9.

PepsiCo (2021) Avena sativa – OT3098 v2. https://wheat.pw.usda.gov/jb?data=/ggds/oat-ot3098v2-pepsico.

Polley, A., Plieske, J., Grafarend-Belau, E., Manninen, O., Vetelainen, M., Ceplitis, A., Vallenback, P., Alsheikh, M., Gnad, H. & Ganal, M. W. (2023) A new 7K SNP genotyping array for oat. *Oat newsletter, 60:4.* URL: https://oatnews.org/oatnews_pdfs/2020/oatnews_2023_Ganal.pdf.

R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Sari, E., Cabral, A. L., Polley, B., Tan, Y., Hsueh, E., Konkin, D. J., Knox, R. E., Ruan, Y., & Fobert, P. R. (2019) Weighted gene co-expression network analysis unveils gene networks associated with the Fusarium head blight resistance in tetraploid wheat. *BMC genomics*, *20*(1), 925. https://doi.org/10.1186/s12864-019-6161-8.

Schroeder, H. W. & Christensen, J. J. (1963) Factors Affecting Resistance of Wheat to Scab Caused by Gibberella zeae. *Phytopathology, 53*, 831-838.

Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., & Kizek, R. (2010). Deoxynivalenol and its toxicity. *Interdisciplinary toxicology*, *3*(3), 94–99. https://doi.org/10.2478/v10102-010-0019-x.

Sørensen, E. S., Jansen, C., Windju, S., Crossa, J., Sonesson, A. K., Lillemo, M., & Alsheikh, M. (2023) Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa*). *Plant Breeding*, 142(1), 41–53. https://doi.org/10.1111/pbr.13061.

Tekle, S., Dill-Macky, R., Skinnes, H., Tronsmo, A. M. & Bjørnstad, Å. (2012) Infection process of *Fusarium graminearum* in oats (*Avena sativa* L.). *European Journal Plant Pathololgy, 132*, 431–442. https://doi.org/10.1007/s10658-011-9888-x.

Tekle, S., Skinnes, H. & Bjørnstad, Å. (2013) The germination problem of oat seed lots affected by *Fusarium* head blight. *European Journal of Plant Pathology*, *135*, 147–158. https://doi.org/10.1007/s10658-012-0074-6.

Tekle, S., Lillemo, M., Skinnes, H., Reitan, L., Buraas, T. & Bjørnstad, A. (2018) Screening of oat accessions for fusarium head blight resistance using spawn-inoculated field experiments. *Crop Science 58*, 143-151. https://doi.org/10.2135/cropsci2017.04.0264.

Tinker, N. A., Chao, S., Lazo, G.R., Oliver, R. E., Huang, Y-F., Poland, J. A., Jellen, E. N., Maughan, P. J., Kilian, A. & Jackson, E. W. (2014) A SNP Genotyping Array for Hexaploid Oat. *The Plant Genome, 7*: plantgenome2014.03.0010. https://doi.org/10.3835/plantgenome2014.03.0010.

West, R. M. (2022) Best practice in statistics: The use of log transformation. *Annals of Clinical Biochemistry*, 59(3):162-165. https://doi.org/10.1177/00045632211050 53.

Wang, J. & Zhang, Z. (2021). GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics, Proteomics & Bioinformatics, 19(4)*, 629-640. https://doi.org/10.1016/j.gpb.2021.08.005.

Xu, X-M., Nicholsen, P., Thomsett, M. A., Simpson, D., Cooke, B. M., Doohan, F. M., Brennan, J., Monaghan, S., Moretti, A., Mule, G., Hornok, L., Beki, E., Tatnell, J., Ritieni, A. & Edwards, S. G. (2008) Relationship between the fungal complex causing fusarium head blight in wheat and environmental conditions. *Phytopathology, 98,* 69-78. https://doi.org/10.1094/PHYTO-98-1-0069.

Yan, W., Fregeau-Reid, J., Rioux, S., Pageau, D., Xue, A., Martin, R., Fedak, G., de Haan, B., Lajeunesse, J. & Savard, M. (2010) Response of oat genotypes to Fusarium head blight in Eastern Canada. *Crop Science, 50,* 134–142. https://doi.org/10.2135/cropsci2009.03.0129.

Yao E., Blake V. C, Cooper L., Wight C. P., Michel S., Cagirici H. B., Lazo G. R., Birkett C. L., Waring D. J., Jannink J.-L., Holmes, I., Waters A. J., Eickholt D. P. & Sen, T. Z. (2022) "GrainGenes: A Data-Rich Repository for Small Grains Genetics and Genomics", Database, Volume 2022, baac034, https://doi.org/10.1093/database/baac034.

Supplementary material

# Identification of haplotypes associated with resistance to *Fusarium graminearum* in spring oat (*Avena sativa* L.).

Manuscript

Espen Sannes Sørensen[1,2], Susanne Windju[1], Constantin Jansen[1], Anna Kristina Sonesson[3], Morten Lillemo[2], Muath Alsheikh[1,2]

[1]Graminor AS, Ridabu, Norway.
[2]Department of Plant Science, Norwegian University of Life Sciences, Ås, Norway.
[3]Nofima AS, Ås, Norway.

Corresponding author:

Espen Sannes Sørensen
E-mail: espen.sorensen@nmbu.no.
Orcid-ID: 0000-0003-4646-8250

***Table S2**: Name of the genotypes used in the GWAS analysis, country of origin and name of population.*

| Name | Origin | Population |
|---|---|---|
| Kareela | Australia | GWAS panel |
| Elison | Austria | GWAS panel |
| Eneko | Austria | GWAS panel |
| AC_Assiniboia | Canada | GWAS panel |
| CDC_Boyer | Canada | GWAS panel |
| CDC_Nasser | Canada | GWAS panel |
| CDC_Orrin | Canada | GWAS panel |
| Hi Fi | Canada | GWAS panel |
| Leggett | Canada | GWAS panel |
| Pro Fi | Canada | GWAS panel |
| Sol Fi | Canada | GWAS panel |
| Palu Abed | Denmark | GWAS panel |
| Akseli | Finland | GWAS panel |
| Aslak | Finland | GWAS panel |
| Fiia | Finland | GWAS panel |
| Hannes | Finland | GWAS panel |
| Marika | Finland | GWAS panel |
| Roope | Finland | GWAS panel |
| Steinar | Finland | GWAS panel |
| Yty | Finland | GWAS panel |
| Aarre | Finland | GWAS panel |
| Aragon | Germany | GWAS panel |
| Bauer | Germany | GWAS panel |
| Bessin | Germany | GWAS panel |
| Breeding line 1 | Germany | GWAS panel |
| Breeding line 2 | Germany | GWAS panel |
| Breeding line 3 | Germany | GWAS panel |
| Breeding line 4 | Germany | GWAS panel |
| Breeding line 5 | Germany | GWAS panel |
| Breeding line 6 | Germany | GWAS panel |
| Breeding line 7 | Germany | GWAS panel |
| Breeding line 8 | Germany | GWAS panel |
| Breeding line 9 | Germany | GWAS panel |
| Breeding line 10 | Germany | GWAS panel |
| Caddy | Germany | GWAS panel |
| Canary | Germany | GWAS panel |
| Curly | Germany | GWAS panel |
| Delfin | Germany | GWAS panel |
| Flemingsplus | Germany | GWAS panel |
| Kaplan | Germany | GWAS panel |
| Moby | Germany | GWAS panel |
| Neklan | Germany | GWAS panel |

| | | |
|---|---|---|
| Poseidon | Germany | GWAS panel |
| Proxy | Germany | GWAS panel |
| Revisor | Germany | GWAS panel |
| Scorpion | Germany | GWAS panel |
| Symphony | Germany | GWAS panel |
| Breeding line 11 | Netherlands | GWAS panel |
| Breeding line 12 | Netherlands | GWAS panel |
| Breeding line 13 | Netherlands | GWAS panel |
| Breeding line 14 | Netherlands | GWAS panel |
| Breeding line 15 | Netherlands | GWAS panel |
| Breeding line 16 | Netherlands | GWAS panel |
| Breeding line 17 | Netherlands | GWAS panel |
| Liberto | Netherlands | GWAS panel |
| Mustang | Netherlands | GWAS panel |
| Avetron | Norway | GWAS panel |
| Biri | Norway | GWAS panel |
| Breeding line 18 | Norway | GWAS panel |
| Breeding line 19 | Norway | GWAS panel |
| Breeding line 20 | Norway | GWAS panel |
| Breeding line 21 | Norway | GWAS panel |
| Breeding line 22 | Norway | GWAS panel |
| Breeding line 23 | Norway | GWAS panel |
| Breeding line 24 | Norway | GWAS panel |
| Breeding line 25 | Norway | GWAS panel |
| Breeding line 26 | Norway | GWAS panel |
| Breeding line 27 | Norway | GWAS panel |
| Breeding line 28 | Norway | GWAS panel |
| Breeding line 29 | Norway | GWAS panel |
| Breeding line 30 | Norway | GWAS panel |
| Breeding line 31 | Norway | GWAS panel |
| Breeding line 32 | Norway | GWAS panel |
| Breeding line 33 | Norway | GWAS panel |
| Breeding line 34 | Norway | GWAS panel |
| Breeding line 35 | Norway | GWAS panel |
| Breeding line 36 | Norway | GWAS panel |
| Breeding line 37 | Norway | GWAS panel |
| Breeding line 38 | Norway | GWAS panel |
| Breeding line 39 | Norway | GWAS panel |
| Breeding line 40 | Norway | GWAS panel |
| Breeding line 41 | Norway | GWAS panel |
| Breeding line 42 | Norway | GWAS panel |
| Breeding line 43 | Norway | GWAS panel |
| Breeding line 44 | Norway | GWAS panel |
| Breeding line 45 | Norway | GWAS panel |
| Breeding line 46 | Norway | GWAS panel |
| Breeding line 47 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 48 | Norway | GWAS panel |
| Breeding line 49 | Norway | GWAS panel |
| Breeding line 50 | Norway | GWAS panel |
| Breeding line 51 | Norway | GWAS panel |
| Breeding line 52 | Norway | GWAS panel |
| Breeding line 53 | Norway | GWAS panel |
| Breeding line 54 | Norway | GWAS panel |
| Breeding line 55 | Norway | GWAS panel |
| Breeding line 56 | Norway | GWAS panel |
| Breeding line 57 | Norway | GWAS panel |
| Breeding line 58 | Norway | GWAS panel |
| Breeding line 59 | Norway | GWAS panel |
| Breeding line 60 | Norway | GWAS panel |
| Breeding line 61 | Norway | GWAS panel |
| Breeding line 62 | Norway | GWAS panel |
| Breeding line 63 | Norway | GWAS panel |
| Breeding line 64 | Norway | GWAS panel |
| Breeding line 65 | Norway | GWAS panel |
| Breeding line 66 | Norway | GWAS panel |
| Breeding line 67 | Norway | GWAS panel |
| Breeding line 68 | Norway | GWAS panel |
| Breeding line 69 | Norway | GWAS panel |
| Breeding line 70 | Norway | GWAS panel |
| Breeding line 71 | Norway | GWAS panel |
| Breeding line 72 | Norway | GWAS panel |
| Breeding line 73 | Norway | GWAS panel |
| Breeding line 74 | Norway | GWAS panel |
| Breeding line 75 | Norway | GWAS panel |
| Breeding line 76 | Norway | GWAS panel |
| Breeding line 77 | Norway | GWAS panel |
| Breeding line 78 | Norway | GWAS panel |
| Breeding line 79 | Norway | GWAS panel |
| Breeding line 80 | Norway | GWAS panel |
| Breeding line 81 | Norway | GWAS panel |
| Breeding line 82 | Norway | GWAS panel |
| Breeding line 83 | Norway | GWAS panel |
| Breeding line 84 | Norway | GWAS panel |
| Breeding line 85 | Norway | GWAS panel |
| Breeding line 86 | Norway | GWAS panel |
| Breeding line 87 | Norway | GWAS panel |
| Breeding line 88 | Norway | GWAS panel |
| Breeding line 89 | Norway | GWAS panel |
| Breeding line 90 | Norway | GWAS panel |
| Breeding line 91 | Norway | GWAS panel |
| Breeding line 92 | Norway | GWAS panel |
| Breeding line 93 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 94 | Norway | GWAS panel |
| Breeding line 95 | Norway | GWAS panel |
| Breeding line 96 | Norway | GWAS panel |
| Breeding line 97 | Norway | GWAS panel |
| Breeding line 98 | Norway | GWAS panel |
| Breeding line 99 | Norway | GWAS panel |
| Breeding line 100 | Norway | GWAS panel |
| Breeding line 101 | Norway | GWAS panel |
| Breeding line 102 | Norway | GWAS panel |
| Breeding line 103 | Norway | GWAS panel |
| Breeding line 104 | Norway | GWAS panel |
| Breeding line 105 | Norway | GWAS panel |
| Breeding line 106 | Norway | GWAS panel |
| Breeding line 107 | Norway | GWAS panel |
| Breeding line 108 | Norway | GWAS panel |
| Breeding line 109 | Norway | GWAS panel |
| Breeding line 110 | Norway | GWAS panel |
| Breeding line 111 | Norway | GWAS panel |
| Breeding line 112 | Norway | GWAS panel |
| Breeding line 113 | Norway | GWAS panel |
| Breeding line 114 | Norway | GWAS panel |
| Breeding line 115 | Norway | GWAS panel |
| Breeding line 116 | Norway | GWAS panel |
| Breeding line 117 | Norway | GWAS panel |
| Breeding line 118 | Norway | GWAS panel |
| Breeding line 119 | Norway | GWAS panel |
| Breeding line 120 | Norway | GWAS panel |
| Breeding line 121 | Norway | GWAS panel |
| Breeding line 122 | Norway | GWAS panel |
| Breeding line 123 | Norway | GWAS panel |
| Breeding line 124 | Norway | GWAS panel |
| Breeding line 125 | Norway | GWAS panel |
| Breeding line 126 | Norway | GWAS panel |
| Breeding line 127 | Norway | GWAS panel |
| Breeding line 128 | Norway | GWAS panel |
| Breeding line 129 | Norway | GWAS panel |
| Breeding line 130 | Norway | GWAS panel |
| Breeding line 131 | Norway | GWAS panel |
| Breeding line 132 | Norway | GWAS panel |
| Breeding line 133 | Norway | GWAS panel |
| Breeding line 134 | Norway | GWAS panel |
| Breeding line 135 | Norway | GWAS panel |
| Breeding line 136 | Norway | GWAS panel |
| Breeding line 137 | Norway | GWAS panel |
| Breeding line 138 | Norway | GWAS panel |
| Breeding line 139 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 140 | Norway | GWAS panel |
| Breeding line 141 | Norway | GWAS panel |
| Breeding line 142 | Norway | GWAS panel |
| Breeding line 143 | Norway | GWAS panel |
| Breeding line 144 | Norway | GWAS panel |
| Breeding line 145 | Norway | GWAS panel |
| Breeding line 146 | Norway | GWAS panel |
| Breeding line 147 | Norway | GWAS panel |
| Breeding line 148 | Norway | GWAS panel |
| Breeding line 149 | Norway | GWAS panel |
| Breeding line 150 | Norway | GWAS panel |
| Breeding line 151 | Norway | GWAS panel |
| Breeding line 152 | Norway | GWAS panel |
| Breeding line 153 | Norway | GWAS panel |
| Breeding line 154 | Norway | GWAS panel |
| Breeding line 155 | Norway | GWAS panel |
| Breeding line 156 | Norway | GWAS panel |
| Breeding line 157 | Norway | GWAS panel |
| Breeding line 158 | Norway | GWAS panel |
| Breeding line 159 | Norway | GWAS panel |
| Breeding line 160 | Norway | GWAS panel |
| Breeding line 161 | Norway | GWAS panel |
| Breeding line 162 | Norway | GWAS panel |
| Breeding line 163 | Norway | GWAS panel |
| Breeding line 164 | Norway | GWAS panel |
| Breeding line 165 | Norway | GWAS panel |
| Breeding line 166 | Norway | GWAS panel |
| Breeding line 167 | Norway | GWAS panel |
| Breeding line 168 | Norway | GWAS panel |
| Breeding line 169 | Norway | GWAS panel |
| Breeding line 170 | Norway | GWAS panel |
| Breeding line 171 | Norway | GWAS panel |
| Breeding line 172 | Norway | GWAS panel |
| Breeding line 173 | Norway | GWAS panel |
| Breeding line 174 | Norway | GWAS panel |
| Breeding line 175 | Norway | GWAS panel |
| Breeding line 176 | Norway | GWAS panel |
| Breeding line 177 | Norway | GWAS panel |
| Breeding line 178 | Norway | GWAS panel |
| Breeding line 179 | Norway | GWAS panel |
| Breeding line 180 | Norway | GWAS panel |
| Breeding line 181 | Norway | GWAS panel |
| Breeding line 182 | Norway | GWAS panel |
| Breeding line 183 | Norway | GWAS panel |
| Breeding line 184 | Norway | GWAS panel |
| Breeding line 185 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 186 | Norway | GWAS panel |
| Breeding line 187 | Norway | GWAS panel |
| Breeding line 188 | Norway | GWAS panel |
| Breeding line 189 | Norway | GWAS panel |
| Breeding line 190 | Norway | GWAS panel |
| Breeding line 191 | Norway | GWAS panel |
| Breeding line 192 | Norway | GWAS panel |
| Breeding line 193 | Norway | GWAS panel |
| Breeding line 194 | Norway | GWAS panel |
| Breeding line 195 | Norway | GWAS panel |
| Breeding line 196 | Norway | GWAS panel |
| Breeding line 197 | Norway | GWAS panel |
| Breeding line 198 | Norway | GWAS panel |
| Breeding line 199 | Norway | GWAS panel |
| Breeding line 200 | Norway | GWAS panel |
| Breeding line 201 | Norway | GWAS panel |
| Breeding line 202 | Norway | GWAS panel |
| Breeding line 203 | Norway | GWAS panel |
| Breeding line 204 | Norway | GWAS panel |
| Breeding line 205 | Norway | GWAS panel |
| Breeding line 206 | Norway | GWAS panel |
| Breeding line 207 | Norway | GWAS panel |
| Breeding line 208 | Norway | GWAS panel |
| Breeding line 209 | Norway | GWAS panel |
| Breeding line 210 | Norway | GWAS panel |
| Breeding line 211 | Norway | GWAS panel |
| Breeding line 212 | Norway | GWAS panel |
| Breeding line 213 | Norway | GWAS panel |
| Breeding line 214 | Norway | GWAS panel |
| Breeding line 215 | Norway | GWAS panel |
| Breeding line 216 | Norway | GWAS panel |
| Breeding line 217 | Norway | GWAS panel |
| Breeding line 218 | Norway | GWAS panel |
| Breeding line 219 | Norway | GWAS panel |
| Breeding line 220 | Norway | GWAS panel |
| Breeding line 221 | Norway | GWAS panel |
| Breeding line 222 | Norway | GWAS panel |
| Breeding line 223 | Norway | GWAS panel |
| Breeding line 224 | Norway | GWAS panel |
| Breeding line 225 | Norway | GWAS panel |
| Breeding line 226 | Norway | GWAS panel |
| Breeding line 227 | Norway | GWAS panel |
| Breeding line 228 | Norway | GWAS panel |
| Breeding line 229 | Norway | GWAS panel |
| Breeding line 230 | Norway | GWAS panel |
| Breeding line 231 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 232 | Norway | GWAS panel |
| Breeding line 233 | Norway | GWAS panel |
| Breeding line 234 | Norway | GWAS panel |
| Breeding line 235 | Norway | GWAS panel |
| Breeding line 236 | Norway | GWAS panel |
| Breeding line 237 | Norway | GWAS panel |
| Breeding line 238 | Norway | GWAS panel |
| Breeding line 239 | Norway | GWAS panel |
| Breeding line 240 | Norway | GWAS panel |
| Breeding line 241 | Norway | GWAS panel |
| Breeding line 242 | Norway | GWAS panel |
| Breeding line 243 | Norway | GWAS panel |
| Breeding line 244 | Norway | GWAS panel |
| Breeding line 245 | Norway | GWAS panel |
| Breeding line 246 | Norway | GWAS panel |
| Breeding line 247 | Norway | GWAS panel |
| Breeding line 248 | Norway | GWAS panel |
| Breeding line 249 | Norway | GWAS panel |
| Breeding line 250 | Norway | GWAS panel |
| Breeding line 251 | Norway | GWAS panel |
| Breeding line 252 | Norway | GWAS panel |
| Breeding line 253 | Norway | GWAS panel |
| Breeding line 254 | Norway | GWAS panel |
| Breeding line 255 | Norway | GWAS panel |
| Breeding line 256 | Norway | GWAS panel |
| Breeding line 257 | Norway | GWAS panel |
| Breeding line 258 | Norway | GWAS panel |
| Breeding line 259 | Norway | GWAS panel |
| Breeding line 260 | Norway | GWAS panel |
| Breeding line 261 | Norway | GWAS panel |
| Breeding line 262 | Norway | GWAS panel |
| Breeding line 263 | Norway | GWAS panel |
| Breeding line 264 | Norway | GWAS panel |
| Breeding line 265 | Norway | GWAS panel |
| Breeding line 266 | Norway | GWAS panel |
| Breeding line 267 | Norway | GWAS panel |
| Breeding line 268 | Norway | GWAS panel |
| Breeding line 269 | Norway | GWAS panel |
| Breeding line 270 | Norway | GWAS panel |
| Breeding line 271 | Norway | GWAS panel |
| Breeding line 272 | Norway | GWAS panel |
| Breeding line 273 | Norway | GWAS panel |
| Breeding line 274 | Norway | GWAS panel |
| Breeding line 275 | Norway | GWAS panel |
| Breeding line 276 | Norway | GWAS panel |
| Breeding line 277 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 278 | Norway | GWAS panel |
| Breeding line 279 | Norway | GWAS panel |
| Breeding line 280 | Norway | GWAS panel |
| Breeding line 281 | Norway | GWAS panel |
| Breeding line 282 | Norway | GWAS panel |
| Breeding line 283 | Norway | GWAS panel |
| Breeding line 284 | Norway | GWAS panel |
| Breeding line 285 | Norway | GWAS panel |
| Breeding line 286 | Norway | GWAS panel |
| Breeding line 287 | Norway | GWAS panel |
| Breeding line 288 | Norway | GWAS panel |
| Breeding line 289 | Norway | GWAS panel |
| Breeding line 290 | Norway | GWAS panel |
| Breeding line 291 | Norway | GWAS panel |
| Breeding line 292 | Norway | GWAS panel |
| Breeding line 293 | Norway | GWAS panel |
| Breeding line 294 | Norway | GWAS panel |
| Breeding line 295 | Norway | GWAS panel |
| Breeding line 296 | Norway | GWAS panel |
| Breeding line 297 | Norway | GWAS panel |
| Breeding line 298 | Norway | GWAS panel |
| Breeding line 299 | Norway | GWAS panel |
| Breeding line 300 | Norway | GWAS panel |
| Breeding line 301 | Norway | GWAS panel |
| Breeding line 302 | Norway | GWAS panel |
| Breeding line 303 | Norway | GWAS panel |
| Breeding line 304 | Norway | GWAS panel |
| Breeding line 305 | Norway | GWAS panel |
| Breeding line 306 | Norway | GWAS panel |
| Breeding line 307 | Norway | GWAS panel |
| Breeding line 308 | Norway | GWAS panel |
| Breeding line 309 | Norway | GWAS panel |
| Breeding line 310 | Norway | GWAS panel |
| Breeding line 311 | Norway | GWAS panel |
| Breeding line 312 | Norway | GWAS panel |
| Breeding line 313 | Norway | GWAS panel |
| Breeding line 314 | Norway | GWAS panel |
| Breeding line 315 | Norway | GWAS panel |
| Breeding line 316 | Norway | GWAS panel |
| Breeding line 317 | Norway | GWAS panel |
| Breeding line 318 | Norway | GWAS panel |
| Breeding line 319 | Norway | GWAS panel |
| Breeding line 320 | Norway | GWAS panel |
| Breeding line 321 | Norway | GWAS panel |
| Breeding line 322 | Norway | GWAS panel |
| Breeding line 323 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 324 | Norway | GWAS panel |
| Breeding line 325 | Norway | GWAS panel |
| Breeding line 326 | Norway | GWAS panel |
| Breeding line 327 | Norway | GWAS panel |
| Breeding line 328 | Norway | GWAS panel |
| Breeding line 329 | Norway | GWAS panel |
| Breeding line 330 | Norway | GWAS panel |
| Breeding line 331 | Norway | GWAS panel |
| Breeding line 332 | Norway | GWAS panel |
| Breeding line 333 | Norway | GWAS panel |
| Breeding line 334 | Norway | GWAS panel |
| Breeding line 335 | Norway | GWAS panel |
| Breeding line 336 | Norway | GWAS panel |
| Breeding line 337 | Norway | GWAS panel |
| Breeding line 338 | Norway | GWAS panel |
| Breeding line 339 | Norway | GWAS panel |
| Breeding line 340 | Norway | GWAS panel |
| Breeding line 341 | Norway | GWAS panel |
| Breeding line 342 | Norway | GWAS panel |
| Breeding line 343 | Norway | GWAS panel |
| Breeding line 344 | Norway | GWAS panel |
| Breeding line 345 | Norway | GWAS panel |
| Breeding line 346 | Norway | GWAS panel |
| Breeding line 347 | Norway | GWAS panel |
| Breeding line 348 | Norway | GWAS panel |
| Breeding line 349 | Norway | GWAS panel |
| Breeding line 350 | Norway | GWAS panel |
| Breeding line 351 | Norway | GWAS panel |
| Breeding line 352 | Norway | GWAS panel |
| Breeding line 353 | Norway | GWAS panel |
| Breeding line 354 | Norway | GWAS panel |
| Breeding line 355 | Norway | GWAS panel |
| Breeding line 356 | Norway | GWAS panel |
| Breeding line 357 | Norway | GWAS panel |
| Breeding line 358 | Norway | GWAS panel |
| Breeding line 359 | Norway | GWAS panel |
| Breeding line 360 | Norway | GWAS panel |
| Breeding line 361 | Norway | GWAS panel |
| Breeding line 362 | Norway | GWAS panel |
| Breeding line 363 | Norway | GWAS panel |
| Breeding line 364 | Norway | GWAS panel |
| Breeding line 365 | Norway | GWAS panel |
| Breeding line 366 | Norway | GWAS panel |
| Breeding line 367 | Norway | GWAS panel |
| Breeding line 368 | Norway | GWAS panel |
| Breeding line 369 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 370 | Norway | GWAS panel |
| Breeding line 371 | Norway | GWAS panel |
| Breeding line 372 | Norway | GWAS panel |
| Breeding line 373 | Norway | GWAS panel |
| Breeding line 374 | Norway | GWAS panel |
| Breeding line 375 | Norway | GWAS panel |
| Breeding line 376 | Norway | GWAS panel |
| Breeding line 377 | Norway | GWAS panel |
| Breeding line 378 | Norway | GWAS panel |
| Breeding line 379 | Norway | GWAS panel |
| Breeding line 380 | Norway | GWAS panel |
| Breeding line 381 | Norway | GWAS panel |
| Breeding line 382 | Norway | GWAS panel |
| Breeding line 383 | Norway | GWAS panel |
| Breeding line 384 | Norway | GWAS panel |
| Breeding line 385 | Norway | GWAS panel |
| Breeding line 386 | Norway | GWAS panel |
| Breeding line 387 | Norway | GWAS panel |
| Breeding line 388 | Norway | GWAS panel |
| Breeding line 389 | Norway | GWAS panel |
| Breeding line 390 | Norway | GWAS panel |
| Breeding line 391 | Norway | GWAS panel |
| Breeding line 392 | Norway | GWAS panel |
| Breeding line 393 | Norway | GWAS panel |
| Breeding line 394 | Norway | GWAS panel |
| Breeding line 395 | Norway | GWAS panel |
| Breeding line 396 | Norway | GWAS panel |
| Breeding line 397 | Norway | GWAS panel |
| Breeding line 398 | Norway | GWAS panel |
| Breeding line 399 | Norway | GWAS panel |
| Breeding line 400 | Norway | GWAS panel |
| Breeding line 401 | Norway | GWAS panel |
| Breeding line 402 | Norway | GWAS panel |
| Breeding line 403 | Norway | GWAS panel |
| Breeding line 404 | Norway | GWAS panel |
| Breeding line 405 | Norway | GWAS panel |
| Breeding line 406 | Norway | GWAS panel |
| Breeding line 407 | Norway | GWAS panel |
| Breeding line 408 | Norway | GWAS panel |
| Breeding line 409 | Norway | GWAS panel |
| Breeding line 410 | Norway | GWAS panel |
| Breeding line 411 | Norway | GWAS panel |
| Breeding line 412 | Norway | GWAS panel |
| Breeding line 413 | Norway | GWAS panel |
| Breeding line 414 | Norway | GWAS panel |
| Breeding line 415 | Norway | GWAS panel |

| | | |
|---|---|---|
| Breeding line 416 | Norway | GWAS panel |
| Breeding line 417 | Norway | GWAS panel |
| Breeding line 418 | Norway | GWAS panel |
| Breeding line 419 | Norway | GWAS panel |
| Breeding line 420 | Norway | GWAS panel |
| Breeding line 421 | Norway | GWAS panel |
| Breeding line 422 | Norway | GWAS panel |
| Breeding line 423 | Norway | GWAS panel |
| Breeding line 424 | Norway | GWAS panel |
| Breeding line 425 | Norway | GWAS panel |
| Breeding line 426 | Norway | GWAS panel |
| Breeding line 427 | Norway | GWAS panel |
| Breeding line 428 | Norway | GWAS panel |
| Breeding line 429 | Norway | GWAS panel |
| Breeding line 430 | Norway | GWAS panel |
| Breeding line 431 | Norway | GWAS panel |
| Dovre | Norway | GWAS panel |
| Eidsvoll | Norway | GWAS panel |
| Flisa | Norway | GWAS panel |
| Frigg | Norway | GWAS panel |
| Gere | Norway | GWAS panel |
| Gimse | Norway | GWAS panel |
| Grane | Norway | GWAS panel |
| Haga | Norway | GWAS panel |
| Hurdal | Norway | GWAS panel |
| Hurum | Norway | GWAS panel |
| Kapp | Norway | GWAS panel |
| Lena | Norway | GWAS panel |
| Martin | Norway | GWAS panel |
| Moholt | Norway | GWAS panel |
| Munin | Norway | GWAS panel |
| Nes | Norway | GWAS panel |
| Odal | Norway | GWAS panel |
| Pol | Norway | GWAS panel |
| Ringsaker | Norway | GWAS panel |
| Roverud | Norway | GWAS panel |
| Skarnes | Norway | GWAS panel |
| Vinger | Norway | GWAS panel |
| Våler | Norway | GWAS panel |
| Årnes | Norway | GWAS panel |
| Avanti | Sweden | GWAS panel |
| Aveny | Sweden | GWAS panel |
| Bambu | Sweden | GWAS panel |
| Belinda | Sweden | GWAS panel |
| Breeding line 432 | Sweden | GWAS panel |
| Breeding line 433 | Sweden | GWAS panel |

| | | |
|---|---|---|
| Breeding line 434 | Sweden | GWAS panel |
| Breeding line 435 | Sweden | GWAS panel |
| Breeding line 436 | Sweden | GWAS panel |
| Breeding line 437 | Sweden | GWAS panel |
| Breeding line 438 | Sweden | GWAS panel |
| Breeding line 439 | Sweden | GWAS panel |
| Breeding line 440 | Sweden | GWAS panel |
| Breeding line 441 | Sweden | GWAS panel |
| Breeding line 442 | Sweden | GWAS panel |
| Breeding line 443 | Sweden | GWAS panel |
| Breeding line 444 | Sweden | GWAS panel |
| Breeding line 445 | Sweden | GWAS panel |
| Cilla | Sweden | GWAS panel |
| Fatima | Sweden | GWAS panel |
| Frode | Sweden | GWAS panel |
| Fyris | Sweden | GWAS panel |
| Galant | Sweden | GWAS panel |
| Guld | Sweden | GWAS panel |
| Gullregn II | Sweden | GWAS panel |
| Gunhild | Sweden | GWAS panel |
| Hvit Odal | Sweden | GWAS panel |
| Ingeborg | Sweden | GWAS panel |
| Klock Extra | Sweden | GWAS panel |
| Matilda | Sweden | GWAS panel |
| Nike | Sweden | GWAS panel |
| Ramiro | Sweden | GWAS panel |
| Sang | Sweden | GWAS panel |
| Selma | Sweden | GWAS panel |
| Svala | Sweden | GWAS panel |
| Svea | Sweden | GWAS panel |
| SW Kerstin | Sweden | GWAS panel |
| Titus | Sweden | GWAS panel |
| Örn | Sweden | GWAS panel |
| Breeding line 446 | USA | GWAS panel |
| Breeding line 447 | USA | GWAS panel |
| Breeding line 448 | USA | GWAS panel |
| Breeding line 449 | USA | GWAS panel |
| Milton | USA | GWAS panel |
| Robust | USA | GWAS panel |
| Breeding line 450 | Norway | Validation population 1 |
| Breeding line 451 | Norway | Validation population 1 |
| Breeding line 452 | Norway | Validation population 1 |
| Breeding line 453 | Norway | Validation population 1 |
| Breeding line 454 | Norway | Validation population 1 |
| Breeding line 455 | Norway | Validation population 1 |
| Breeding line 456 | Norway | Validation population 1 |

| | | |
|---|---|---|
| Breeding line 457 | Norway | Validation population 1 |
| Breeding line 458 | Norway | Validation population 1 |
| Breeding line 459 | Norway | Validation population 1 |
| Breeding line 460 | Norway | Validation population 1 |
| Breeding line 461 | Norway | Validation population 1 |
| Breeding line 462 | Norway | Validation population 1 |
| Breeding line 463 | Norway | Validation population 1 |
| Breeding line 464 | Norway | Validation population 1 |
| Breeding line 465 | Norway | Validation population 1 |
| Breeding line 466 | Norway | Validation population 1 |
| Breeding line 467 | Norway | Validation population 1 |
| Breeding line 468 | Norway | Validation population 1 |
| Breeding line 469 | Norway | Validation population 1 |
| Breeding line 470 | Norway | Validation population 1 |
| Breeding line 471 | Norway | Validation population 1 |
| Breeding line 472 | Norway | Validation population 1 |
| Breeding line 473 | Norway | Validation population 1 |
| Breeding line 474 | Norway | Validation population 1 |
| Breeding line 475 | Norway | Validation population 1 |
| Breeding line 476 | Norway | Validation population 1 |
| Breeding line 477 | Norway | Validation population 1 |
| Breeding line 478 | Norway | Validation population 1 |
| Breeding line 479 | Norway | Validation population 1 |
| Breeding line 480 | Norway | Validation population 1 |
| Breeding line 481 | Norway | Validation population 1 |
| Breeding line 482 | Norway | Validation population 1 |
| Breeding line 483 | Norway | Validation population 1 |
| Breeding line 484 | Norway | Validation population 1 |
| Breeding line 485 | Norway | Validation population 1 |
| Breeding line 486 | Norway | Validation population 1 |
| Breeding line 487 | Norway | Validation population 1 |
| Breeding line 488 | Norway | Validation population 1 |
| Breeding line 489 | Norway | Validation population 1 |
| Breeding line 490 | Norway | Validation population 1 |
| Breeding line 491 | Norway | Validation population 1 |
| Breeding line 492 | Norway | Validation population 1 |
| Breeding line 493 | Norway | Validation population 1 |
| Breeding line 494 | Norway | Validation population 1 |
| Breeding line 495 | Norway | Validation population 1 |
| Breeding line 496 | Norway | Validation population 1 |
| Breeding line 497 | Norway | Validation population 1 |
| Breeding line 498 | Norway | Validation population 1 |
| Breeding line 499 | Norway | Validation population 1 |
| Breeding line 500 | Norway | Validation population 1 |
| Breeding line 501 | Norway | Validation population 1 |
| Breeding line 502 | Norway | Validation population 1 |

| | | |
|---|---|---|
| Breeding line 503 | Norway | Validation population 1 |
| Breeding line 504 | Norway | Validation population 1 |
| Breeding line 505 | Norway | Validation population 1 |
| Breeding line 506 | Norway | Validation population 1 |
| Breeding line 507 | Norway | Validation population 1 |
| Breeding line 508 | Norway | Validation population 1 |
| Breeding line 509 | Norway | Validation population 1 |
| Breeding line 510 | Norway | Validation population 1 |
| Breeding line 511 | Norway | Validation population 1 |
| Breeding line 512 | Norway | Validation population 1 |
| Breeding line 513 | Norway | Validation population 1 |
| Breeding line 514 | Norway | Validation population 1 |
| Breeding line 515 | Norway | Validation population 1 |
| Breeding line 516 | Norway | Validation population 1 |
| Breeding line 517 | Norway | Validation population 1 |
| Breeding line 518 | Norway | Validation population 1 |
| Breeding line 519 | Norway | Validation population 1 |
| Breeding line 520 | Norway | Validation population 1 |
| Breeding line 521 | Norway | Validation population 1 |
| Breeding line 522 | Norway | Validation population 1 |
| Breeding line 523 | Norway | Validation population 1 |
| Breeding line 524 | Norway | Validation population 1 |
| Breeding line 525 | Norway | Validation population 1 |
| Breeding line 526 | Norway | Validation population 1 |
| Breeding line 527 | Norway | Validation population 1 |
| Breeding line 528 | Norway | Validation population 1 |
| Breeding line 529 | Norway | Validation population 1 |
| Breeding line 530 | Norway | Validation population 1 |
| Breeding line 531 | Norway | Validation population 1 |
| Breeding line 532 | Norway | Validation population 1 |
| Breeding line 533 | Norway | Validation population 1 |
| Breeding line 534 | Norway | Validation population 1 |
| Breeding line 535 | Norway | Validation population 1 |
| Breeding line 536 | Norway | Validation population 1 |
| Breeding line 537 | Norway | Validation population 1 |
| Breeding line 538 | Norway | Validation population 1 |
| Breeding line 539 | Norway | Validation population 1 |
| Breeding line 540 | Norway | Validation population 1 |
| Breeding line 541 | Norway | Validation population 1 |
| Breeding line 542 | Norway | Validation population 1 |
| Breeding line 543 | Norway | Validation population 1 |
| Breeding line 544 | Norway | Validation population 1 |
| Breeding line 545 | Norway | Validation population 1 |
| Breeding line 546 | Norway | Validation population 1 |
| Breeding line 547 | Norway | Validation population 1 |
| Breeding line 548 | Norway | Validation population 1 |

| | | |
|---|---|---|
| Breeding line 549 | Norway | Validation population 1 |
| Breeding line 550 | Norway | Validation population 1 |
| Breeding line 551 | Norway | Validation population 1 |
| Breeding line 552 | Norway | Validation population 1 |
| Breeding line 553 | Norway | Validation population 1 |
| Breeding line 554 | Norway | Validation population 1 |
| Breeding line 555 | Norway | Validation population 1 |
| Breeding line 556 | Norway | Validation population 1 |
| Breeding line 557 | Norway | Validation population 1 |
| Breeding line 558 | Norway | Validation population 1 |
| Breeding line 559 | Norway | Validation population 1 |
| Breeding line 560 | Norway | Validation population 1 |
| Breeding line 561 | Norway | Validation population 1 |
| Breeding line 562 | Norway | Validation population 1 |
| Breeding line 563 | Norway | Validation population 1 |
| Breeding line 564 | Norway | Validation population 1 |
| Breeding line 565 | Norway | Validation population 1 |
| Breeding line 566 | Norway | Validation population 1 |
| Breeding line 567 | Norway | Validation population 1 |
| Breeding line 568 | Norway | Validation population 1 |
| Breeding line 569 | Norway | Validation population 1 |
| Breeding line 570 | Norway | Validation population 1 |
| Breeding line 571 | Norway | Validation population 1 |
| Breeding line 572 | Norway | Validation population 1 |
| Breeding line 573 | Norway | Validation population 1 |
| Breeding line 574 | Norway | Validation population 1 |
| Breeding line 575 | Norway | Validation population 1 |
| Breeding line 576 | Norway | Validation population 1 |
| Breeding line 577 | Norway | Validation population 1 |
| Breeding line 578 | Norway | Validation population 1 |
| Breeding line 579 | Norway | Validation population 1 |
| Breeding line 580 | Norway | Validation population 1 |
| Breeding line 581 | Norway | Validation population 1 |
| Breeding line 582 | Norway | Validation population 1 |
| Breeding line 583 | Norway | Validation population 1 |
| Breeding line 584 | Norway | Validation population 1 |
| Breeding line 585 | Norway | Validation population 1 |
| Breeding line 586 | Norway | Validation population 1 |
| Breeding line 587 | Norway | Validation population 1 |
| Breeding line 588 | Norway | Validation population 1 |
| Breeding line 589 | Norway | Validation population 1 |
| Breeding line 590 | Norway | Validation population 1 |
| Breeding line 591 | Norway | Validation population 1 |
| Breeding line 592 | Norway | Validation population 1 |
| Breeding line 593 | Norway | Validation population 1 |
| Breeding line 594 | Norway | Validation population 1 |

| | | |
|---|---|---|
| Breeding line 595 | Norway | Validation population 1 |
| Breeding line 596 | Norway | Validation population 1 |
| Breeding line 597 | Norway | Validation population 1 |
| Breeding line 598 | Norway | Validation population 1 |
| Breeding line 599 | Norway | Validation population 1 |
| Breeding line 600 | Norway | Validation population 1 |
| Breeding line 601 | Norway | Validation population 1 |
| Breeding line 602 | Norway | Validation population 1 |
| Breeding line 603 | Norway | Validation population 1 |
| Breeding line 604 | Norway | Validation population 1 |
| Breeding line 605 | Norway | Validation population 1 |
| Breeding line 606 | Norway | Validation population 1 |
| Breeding line 607 | Norway | Validation population 1 |
| Breeding line 608 | Norway | Validation population 1 |
| Breeding line 609 | Norway | Validation population 1 |
| Breeding line 610 | Norway | Validation population 1 |
| Breeding line 611 | Norway | Validation population 1 |
| Breeding line 612 | Norway | Validation population 1 |
| Breeding line 613 | Norway | Validation population 1 |
| Breeding line 614 | Norway | Validation population 1 |
| Breeding line 615 | Norway | Validation population 1 |
| Breeding line 616 | Norway | Validation population 1 |
| Breeding line 617 | Norway | Validation population 1 |
| Breeding line 618 | Norway | Validation population 1 |
| Breeding line 619 | Norway | Validation population 1 |
| Breeding line 620 | Norway | Validation population 1 |
| Breeding line 621 | Norway | Validation population 1 |
| Breeding line 622 | Norway | Validation population 1 |
| Breeding line 623 | Norway | Validation population 1 |
| Breeding line 624 | Norway | Validation population 1 |
| Breeding line 625 | Norway | Validation population 1 |
| Breeding line 626 | Norway | Validation population 1 |
| Breeding line 627 | Norway | Validation population 1 |
| Breeding line 628 | Norway | Validation population 1 |
| Breeding line 629 | Norway | Validation population 1 |
| Breeding line 630 | Norway | Validation population 1 |
| Breeding line 631 | Norway | Validation population 1 |
| Breeding line 632 | Norway | Validation population 1 |
| Breeding line 633 | Norway | Validation population 1 |
| Breeding line 634 | Norway | Validation population 1 |
| Breeding line 635 | Norway | Validation population 1 |
| Breeding line 636 | Norway | Validation population 1 |
| Breeding line 637 | Norway | Validation population 1 |
| Breeding line 638 | Norway | Validation population 1 |
| Breeding line 639 | Norway | Validation population 1 |
| Breeding line 640 | Norway | Validation population 1 |

| | | |
|---|---|---|
| Breeding line 641 | Norway | Validation population 1 |
| Breeding line 642 | Norway | Validation population 1 |
| Breeding line 643 | Norway | Validation population 1 |
| Breeding line 644 | Norway | Validation population 1 |
| Breeding line 645 | Norway | Validation population 1 |
| Breeding line 646 | Norway | Validation population 1 |
| Breeding line 647 | Norway | Validation population 1 |
| Breeding line 648 | Norway | Validation population 1 |
| Breeding line 649 | Norway | Validation population 1 |
| Breeding line 650 | Norway | Validation population 1 |
| Breeding line 651 | Norway | Validation population 1 |
| Breeding line 652 | Norway | Validation population 1 |
| Breeding line 653 | Norway | Validation population 1 |
| Breeding line 654 | Norway | Validation population 1 |
| Breeding line 655 | Norway | Validation population 1 |
| Breeding line 656 | Norway | Validation population 1 |
| Breeding line 657 | Norway | Validation population 1 |
| Breeding line 658 | Norway | Validation population 1 |
| Breeding line 659 | Norway | Validation population 1 |
| Breeding line 660 | Norway | Validation population 1 |
| Breeding line 661 | Norway | Validation population 1 |
| Breeding line 662 | Norway | Validation population 1 |
| Breeding line 663 | Norway | Validation population 1 |
| Breeding line 664 | Norway | Validation population 1 |
| Breeding line 665 | Norway | Validation population 1 |
| Breeding line 666 | Norway | Validation population 1 |
| Breeding line 667 | Norway | Validation population 1 |
| Breeding line 668 | Norway | Validation population 1 |
| Breeding line 669 | Norway | Validation population 1 |
| Breeding line 670 | Norway | Validation population 1 |
| Breeding line 671 | Norway | Validation population 1 |
| Breeding line 672 | Norway | Validation population 1 |
| Breeding line 673 | Norway | Validation population 1 |
| Breeding line 674 | Norway | Validation population 1 |
| Breeding line 675 | Norway | Validation population 1 |
| Breeding line 676 | Norway | Validation population 1 |
| Breeding line 677 | Norway | Validation population 1 |
| Breeding line 678 | Norway | Validation population 1 |
| Breeding line 679 | Norway | Validation population 1 |
| Breeding line 680 | Norway | Validation population 1 |
| Breeding line 681 | Norway | Validation population 1 |
| Breeding line 682 | Norway | Validation population 1 |
| Breeding line 683 | Norway | Validation population 1 |
| Breeding line 684 | Norway | Validation population 1 |
| Breeding line 685 | Norway | Validation population 1 |
| Breeding line 686 | Norway | Validation population 1 |

| | | |
|---|---|---|
| Breeding line 687 | Norway | Validation population 1 |
| Breeding line 688 | Norway | Validation population 1 |
| Breeding line 689 | Norway | Validation population 1 |
| Breeding line 690 | Norway | Validation population 1 |
| Breeding line 691 | Norway | Validation population 1 |
| Breeding line 692 | Norway | Validation population 1 |
| Breeding line 693 | Norway | Validation population 1 |
| Breeding line 694 | Norway | Validation population 1 |
| Breeding line 695 | Norway | Validation population 1 |
| Breeding line 696 | Norway | Validation population 1 |
| Breeding line 697 | Norway | Validation population 1 |
| Breeding line 698 | Norway | Validation population 1 |
| Breeding line 699 | Norway | Validation population 1 |
| Breeding line 700 | Norway | Validation population 1 |
| Breeding line 701 | Norway | Validation population 1 |
| Breeding line 702 | Norway | Validation population 1 |
| Breeding line 703 | Norway | Validation population 1 |
| Breeding line 704 | Norway | Validation population 1 |
| Breeding line 705 | Norway | Validation population 1 |
| Breeding line 706 | Norway | Validation population 1 |
| Breeding line 707 | Norway | Validation population 1 |
| Breeding line 708 | Norway | Validation population 1 |
| Breeding line 709 | Norway | Validation population 1 |
| Breeding line 710 | Norway | Validation population 1 |
| Breeding line 711 | Norway | Validation population 1 |
| Breeding line 712 | Norway | Validation population 2 |
| Breeding line 713 | Norway | Validation population 2 |
| Breeding line 714 | Norway | Validation population 2 |
| Breeding line 715 | Norway | Validation population 2 |
| Breeding line 716 | Norway | Validation population 2 |
| Breeding line 717 | Norway | Validation population 2 |
| Breeding line 718 | Norway | Validation population 2 |
| Breeding line 719 | Norway | Validation population 2 |
| Breeding line 720 | Norway | Validation population 2 |
| Breeding line 721 | Norway | Validation population 2 |
| Breeding line 722 | Norway | Validation population 2 |
| Breeding line 723 | Norway | Validation population 2 |
| Breeding line 724 | Norway | Validation population 2 |
| Breeding line 725 | Norway | Validation population 2 |
| Breeding line 726 | Norway | Validation population 2 |
| Breeding line 727 | Norway | Validation population 2 |
| Breeding line 728 | Norway | Validation population 2 |
| Breeding line 729 | Norway | Validation population 2 |
| Breeding line 730 | Norway | Validation population 2 |
| Breeding line 731 | Norway | Validation population 2 |
| Breeding line 732 | Norway | Validation population 2 |

| | | |
|---|---|---|
| Breeding line 733 | Norway | Validation population 2 |
| Breeding line 734 | Norway | Validation population 2 |
| Breeding line 735 | Norway | Validation population 2 |
| Breeding line 736 | Norway | Validation population 2 |
| Breeding line 737 | Norway | Validation population 2 |
| Breeding line 738 | Norway | Validation population 2 |
| Breeding line 739 | Norway | Validation population 2 |
| Breeding line 740 | Norway | Validation population 2 |
| Breeding line 741 | Norway | Validation population 2 |
| Breeding line 742 | Norway | Validation population 2 |
| Breeding line 743 | Norway | Validation population 2 |
| Breeding line 744 | Norway | Validation population 2 |
| Breeding line 745 | Norway | Validation population 2 |
| Breeding line 746 | Norway | Validation population 2 |
| Breeding line 747 | Norway | Validation population 2 |
| Breeding line 748 | Norway | Validation population 2 |
| Breeding line 749 | Norway | Validation population 2 |
| Breeding line 750 | Norway | Validation population 2 |
| Breeding line 751 | Norway | Validation population 2 |
| Breeding line 752 | Norway | Validation population 2 |
| Breeding line 753 | Norway | Validation population 2 |
| Breeding line 754 | Norway | Validation population 2 |
| Breeding line 755 | Norway | Validation population 2 |
| Breeding line 756 | Norway | Validation population 2 |
| Breeding line 757 | Norway | Validation population 2 |
| Breeding line 758 | Norway | Validation population 2 |
| Breeding line 759 | Norway | Validation population 2 |
| Breeding line 760 | Norway | Validation population 2 |
| Breeding line 761 | Norway | Validation population 2 |
| Breeding line 762 | Norway | Validation population 2 |
| Breeding line 763 | Norway | Validation population 2 |
| Breeding line 764 | Norway | Validation population 2 |
| Breeding line 765 | Norway | Validation population 2 |
| Breeding line 766 | Norway | Validation population 2 |
| Breeding line 767 | Norway | Validation population 2 |
| Breeding line 768 | Norway | Validation population 2 |
| Breeding line 769 | Norway | Validation population 2 |
| Breeding line 770 | Norway | Validation population 2 |
| Breeding line 771 | Norway | Validation population 2 |
| Breeding line 772 | Norway | Validation population 2 |
| Breeding line 773 | Norway | Validation population 2 |
| Breeding line 774 | Norway | Validation population 2 |
| Breeding line 775 | Norway | Validation population 2 |
| Breeding line 776 | Norway | Validation population 2 |
| Breeding line 777 | Norway | Validation population 2 |
| Breeding line 778 | Norway | Validation population 2 |

| | | |
|---|---|---|
| Breeding line 779 | Norway | Validation population 2 |
| Breeding line 780 | Norway | Validation population 2 |
| Breeding line 781 | Norway | Validation population 2 |
| Breeding line 782 | Norway | Validation population 2 |
| Breeding line 783 | Norway | Validation population 2 |
| Breeding line 784 | Norway | Validation population 2 |
| Breeding line 785 | Norway | Validation population 2 |
| Breeding line 786 | Norway | Validation population 2 |
| Breeding line 787 | Norway | Validation population 2 |
| Breeding line 788 | Norway | Validation population 2 |
| Breeding line 789 | Norway | Validation population 2 |
| Breeding line 790 | Norway | Validation population 2 |
| Breeding line 791 | Norway | Validation population 2 |
| Breeding line 792 | Norway | Validation population 2 |
| Breeding line 793 | Norway | Validation population 2 |
| Breeding line 794 | Norway | Validation population 2 |
| Breeding line 795 | Norway | Validation population 2 |
| Breeding line 796 | Norway | Validation population 2 |
| Breeding line 797 | Norway | Validation population 2 |
| Breeding line 798 | Norway | Validation population 2 |
| Breeding line 799 | Norway | Validation population 2 |
| Breeding line 800 | Norway | Validation population 2 |
| Breeding line 801 | Norway | Validation population 2 |
| Breeding line 802 | Norway | Validation population 2 |
| Breeding line 803 | Norway | Validation population 2 |
| Breeding line 804 | Norway | Validation population 2 |
| Breeding line 805 | Norway | Validation population 2 |
| Breeding line 806 | Norway | Validation population 2 |
| Breeding line 807 | Norway | Validation population 2 |
| Breeding line 808 | Norway | Validation population 2 |
| Breeding line 809 | Norway | Validation population 2 |
| Breeding line 810 | Norway | Validation population 2 |
| Breeding line 811 | Norway | Validation population 2 |
| Breeding line 812 | Norway | Validation population 2 |
| Breeding line 813 | Norway | Validation population 2 |
| Breeding line 814 | Norway | Validation population 2 |
| Breeding line 815 | Norway | Validation population 2 |
| Breeding line 816 | Norway | Validation population 2 |
| Breeding line 817 | Norway | Validation population 2 |
| Breeding line 818 | Norway | Validation population 2 |
| Breeding line 819 | Norway | Validation population 2 |
| Breeding line 820 | Norway | Validation population 2 |
| Breeding line 821 | Norway | Validation population 2 |
| Breeding line 822 | Norway | Validation population 2 |
| Breeding line 823 | Norway | Validation population 2 |
| Breeding line 824 | Norway | Validation population 2 |

| | | |
|---|---|---|
| Breeding line 825 | Norway | Validation population 2 |
| Breeding line 826 | Norway | Validation population 2 |
| Breeding line 827 | Norway | Validation population 2 |
| Breeding line 828 | Norway | Validation population 2 |
| Breeding line 829 | Norway | Validation population 2 |
| Breeding line 830 | Norway | Validation population 2 |
| Breeding line 831 | Norway | Validation population 2 |
| Breeding line 832 | Norway | Validation population 2 |
| Breeding line 833 | Norway | Validation population 2 |
| Breeding line 834 | Norway | Validation population 2 |
| Breeding line 835 | Norway | Validation population 2 |
| Breeding line 836 | Norway | Validation population 2 |
| Breeding line 837 | Norway | Validation population 2 |
| Breeding line 838 | Norway | Validation population 2 |
| Breeding line 839 | Norway | Validation population 2 |
| Breeding line 840 | Norway | Validation population 2 |
| Breeding line 841 | Norway | Validation population 2 |
| Breeding line 842 | Norway | Validation population 2 |
| Breeding line 843 | Norway | Validation population 2 |
| Breeding line 844 | Norway | Validation population 2 |
| Breeding line 845 | Norway | Validation population 2 |
| Breeding line 846 | Norway | Validation population 2 |
| Breeding line 847 | Norway | Validation population 2 |
| Breeding line 848 | Norway | Validation population 2 |
| Breeding line 849 | Norway | Validation population 2 |
| Breeding line 850 | Norway | Validation population 2 |
| Breeding line 851 | Norway | Validation population 2 |
| Breeding line 852 | Norway | Validation population 2 |
| Breeding line 853 | Norway | Validation population 2 |
| Breeding line 854 | Norway | Validation population 2 |
| Breeding line 855 | Norway | Validation population 2 |
| Breeding line 856 | Norway | Validation population 2 |
| Breeding line 857 | Norway | Validation population 2 |
| Breeding line 858 | Norway | Validation population 2 |
| Breeding line 859 | Norway | Validation population 2 |
| Breeding line 860 | Norway | Validation population 2 |
| Breeding line 861 | Norway | Validation population 2 |
| Breeding line 862 | Norway | Validation population 2 |
| Breeding line 863 | Norway | Validation population 2 |
| Breeding line 864 | Norway | Validation population 2 |
| Breeding line 865 | Norway | Validation population 2 |
| Breeding line 866 | Norway | Validation population 2 |
| Breeding line 867 | Norway | Validation population 2 |
| Breeding line 868 | Norway | Validation population 2 |
| Breeding line 869 | Norway | Validation population 2 |
| Breeding line 870 | Norway | Validation population 2 |

| | | |
|---|---|---|
| Breeding line 871 | Norway | Validation population 2 |
| Breeding line 872 | Norway | Validation population 2 |
| Breeding line 873 | Norway | Validation population 2 |
| Breeding line 874 | Norway | Validation population 2 |
| Breeding line 875 | Norway | Validation population 2 |
| Breeding line 876 | Norway | Validation population 2 |
| Breeding line 877 | Norway | Validation population 2 |
| Breeding line 878 | Norway | Validation population 2 |
| Breeding line 879 | Norway | Validation population 2 |
| Breeding line 880 | Norway | Validation population 2 |
| Breeding line 881 | Norway | Validation population 2 |
| Breeding line 882 | Norway | Validation population 2 |
| Breeding line 883 | Norway | Validation population 2 |
| Breeding line 884 | Norway | Validation population 2 |
| Breeding line 885 | Norway | Validation population 2 |
| Breeding line 886 | Norway | Validation population 2 |
| Breeding line 887 | Norway | Validation population 2 |
| Breeding line 888 | Norway | Validation population 2 |
| Breeding line 889 | Norway | Validation population 2 |
| Breeding line 890 | Norway | Validation population 2 |
| Breeding line 891 | Norway | Validation population 2 |
| Breeding line 892 | Norway | Validation population 2 |
| Breeding line 893 | Norway | Validation population 2 |
| Breeding line 894 | Norway | Validation population 2 |
| Breeding line 895 | Norway | Validation population 2 |
| Breeding line 896 | Norway | Validation population 2 |
| Breeding line 897 | Norway | Validation population 2 |
| Breeding line 898 | Norway | Validation population 2 |
| Breeding line 899 | Norway | Validation population 2 |
| Breeding line 900 | Norway | Validation population 2 |
| Breeding line 901 | Norway | Validation population 2 |
| Breeding line 902 | Norway | Validation population 2 |
| Breeding line 903 | Norway | Validation population 2 |
| Breeding line 904 | Norway | Validation population 2 |
| Breeding line 905 | Norway | Validation population 2 |
| Breeding line 906 | Norway | Validation population 2 |
| Breeding line 907 | Norway | Validation population 2 |
| Breeding line 908 | Norway | Validation population 2 |
| Breeding line 909 | Norway | Validation population 2 |
| Breeding line 910 | Norway | Validation population 2 |
| Breeding line 911 | Norway | Validation population 2 |
| Breeding line 912 | Norway | Validation population 2 |
| Breeding line 913 | Norway | Validation population 2 |
| Breeding line 914 | Norway | Validation population 2 |
| Breeding line 915 | Norway | Validation population 2 |
| Breeding line 916 | Norway | Validation population 2 |

| | | |
|---|---|---|
| Breeding line 917 | Norway | Validation population 2 |
| Breeding line 918 | Norway | Validation population 2 |
| Breeding line 919 | Norway | Validation population 2 |
| Breeding line 920 | Norway | Validation population 2 |
| Breeding line 921 | Norway | Validation population 2 |
| Breeding line 922 | Norway | Validation population 2 |
| Breeding line 923 | Norway | Validation population 2 |
| Breeding line 924 | Norway | Validation population 2 |
| Breeding line 925 | Norway | Validation population 2 |
| Breeding line 926 | Norway | Validation population 2 |
| Breeding line 927 | Norway | Validation population 2 |
| Breeding line 928 | Norway | Validation population 2 |
| Breeding line 929 | Norway | Validation population 2 |
| Breeding line 930 | Norway | Validation population 2 |
| Breeding line 931 | Norway | Validation population 2 |
| Breeding line 932 | Norway | Validation population 2 |
| Breeding line 933 | Norway | Validation population 2 |
| Breeding line 934 | Norway | Validation population 2 |
| Breeding line 935 | Norway | Validation population 2 |
| Breeding line 936 | Norway | Validation population 2 |
| Breeding line 937 | Norway | Validation population 2 |
| Breeding line 938 | Norway | Validation population 2 |
| Breeding line 939 | Norway | Validation population 2 |
| Breeding line 940 | Norway | Validation population 2 |
| Breeding line 941 | Norway | Validation population 2 |
| Breeding line 942 | Norway | Validation population 2 |
| Breeding line 943 | Norway | Validation population 2 |
| Breeding line 944 | Norway | Validation population 2 |
| Breeding line 945 | Norway | Validation population 2 |
| Breeding line 946 | Norway | Validation population 2 |
| Breeding line 947 | Norway | Validation population 2 |
| Breeding line 948 | Norway | Validation population 2 |
| Breeding line 949 | Norway | Validation population 2 |
| Breeding line 950 | Norway | Validation population 2 |
| Breeding line 951 | Norway | Validation population 2 |
| Breeding line 952 | Norway | Validation population 2 |
| Breeding line 953 | Norway | Validation population 2 |
| Breeding line 954 | Norway | Validation population 2 |
| Breeding line 955 | Norway | Validation population 2 |
| Breeding line 956 | Norway | Validation population 2 |
| Breeding line 957 | Norway | Validation population 2 |
| Breeding line 958 | Norway | Validation population 2 |
| Breeding line 959 | Norway | Validation population 2 |
| Breeding line 960 | Norway | Validation population 2 |
| Breeding line 961 | Norway | Validation population 2 |
| Breeding line 962 | Norway | Validation population 2 |

| | | |
|---|---|---|
| Breeding line 963 | Norway | Validation population 2 |
| Breeding line 964 | Norway | Validation population 2 |
| Breeding line 965 | Norway | Validation population 2 |
| Breeding line 966 | Norway | Validation population 2 |
| Breeding line 967 | Norway | Validation population 2 |
| Breeding line 968 | Norway | Validation population 2 |
| Breeding line 969 | Norway | Validation population 2 |
| Breeding line 970 | Norway | Validation population 2 |
| Breeding line 971 | Norway | Validation population 2 |
| Breeding line 972 | Norway | Validation population 2 |
| Breeding line 973 | Norway | Validation population 3 |
| Breeding line 974 | Norway | Validation population 3 |
| Breeding line 975 | Norway | Validation population 3 |
| Breeding line 976 | Norway | Validation population 3 |
| Breeding line 977 | Norway | Validation population 3 |
| Breeding line 978 | Norway | Validation population 3 |
| Breeding line 979 | Norway | Validation population 3 |
| Breeding line 980 | Norway | Validation population 3 |
| Breeding line 981 | Norway | Validation population 3 |
| Breeding line 982 | Norway | Validation population 3 |
| Breeding line 983 | Norway | Validation population 3 |
| Breeding line 984 | Norway | Validation population 3 |
| Breeding line 985 | Norway | Validation population 3 |
| Breeding line 986 | Norway | Validation population 3 |
| Breeding line 987 | Norway | Validation population 3 |
| Breeding line 988 | Norway | Validation population 3 |
| Breeding line 989 | Norway | Validation population 3 |
| Breeding line 990 | Norway | Validation population 3 |
| Breeding line 991 | Norway | Validation population 3 |
| Breeding line 992 | Norway | Validation population 3 |
| Breeding line 993 | Norway | Validation population 3 |
| Breeding line 994 | Norway | Validation population 3 |
| Breeding line 995 | Norway | Validation population 3 |
| Breeding line 996 | Norway | Validation population 3 |
| Breeding line 997 | Norway | Validation population 3 |
| Breeding line 998 | Norway | Validation population 3 |
| Breeding line 999 | Norway | Validation population 3 |
| Breeding line 1000 | Norway | Validation population 3 |
| Breeding line 1001 | Norway | Validation population 3 |
| Breeding line 1002 | Norway | Validation population 3 |
| Breeding line 1003 | Norway | Validation population 3 |
| Breeding line 1004 | Norway | Validation population 3 |
| Breeding line 1005 | Norway | Validation population 3 |
| Breeding line 1006 | Norway | Validation population 3 |
| Breeding line 1007 | Norway | Validation population 3 |
| Breeding line 1008 | Norway | Validation population 3 |

| | | |
|---|---|---|
| Breeding line 1009 | Norway | Validation population 3 |
| Breeding line 1010 | Norway | Validation population 3 |
| Breeding line 1011 | Norway | Validation population 3 |
| Breeding line 1012 | Norway | Validation population 3 |
| Breeding line 1013 | Norway | Validation population 3 |
| Breeding line 1014 | Norway | Validation population 3 |
| Breeding line 1015 | Norway | Validation population 3 |
| Breeding line 1016 | Norway | Validation population 3 |
| Breeding line 1017 | Norway | Validation population 3 |
| Breeding line 1018 | Norway | Validation population 3 |
| Breeding line 1019 | Norway | Validation population 3 |
| Breeding line 1020 | Norway | Validation population 3 |
| Breeding line 1021 | Norway | Validation population 3 |
| Breeding line 1022 | Norway | Validation population 3 |
| Breeding line 1023 | Norway | Validation population 3 |
| Breeding line 1024 | Norway | Validation population 3 |
| Breeding line 1025 | Norway | Validation population 3 |
| Breeding line 1026 | Norway | Validation population 3 |
| Breeding line 1027 | Norway | Validation population 3 |
| Breeding line 1028 | Norway | Validation population 3 |
| Breeding line 1029 | Norway | Validation population 3 |
| Breeding line 1030 | Norway | Validation population 3 |
| Breeding line 1031 | Norway | Validation population 3 |
| Breeding line 1032 | Norway | Validation population 3 |
| Breeding line 1033 | Norway | Validation population 3 |
| Breeding line 1034 | Norway | Validation population 3 |
| Breeding line 1035 | Norway | Validation population 3 |
| Breeding line 1036 | Norway | Validation population 3 |
| Breeding line 1037 | Norway | Validation population 3 |
| Breeding line 1038 | Norway | Validation population 3 |
| Breeding line 1039 | Norway | Validation population 3 |
| Breeding line 1040 | Norway | Validation population 3 |
| Breeding line 1041 | Norway | Validation population 3 |
| Breeding line 1042 | Norway | Validation population 3 |
| Breeding line 1043 | Norway | Validation population 3 |
| Breeding line 1044 | Norway | Validation population 3 |
| Breeding line 1045 | Norway | Validation population 3 |
| Breeding line 1046 | Norway | Validation population 3 |
| Breeding line 1047 | Norway | Validation population 3 |
| Breeding line 1048 | Norway | Validation population 3 |
| Breeding line 1049 | Norway | Validation population 3 |
| Breeding line 1050 | Norway | Validation population 3 |
| Breeding line 1051 | Norway | Validation population 3 |
| Breeding line 1052 | Norway | Validation population 3 |
| Breeding line 1053 | Norway | Validation population 3 |
| Breeding line 1054 | Norway | Validation population 3 |

| | | |
|---|---|---|
| Breeding line 1055 | Norway | Validation population 3 |
| Breeding line 1056 | Norway | Validation population 3 |
| Breeding line 1057 | Norway | Validation population 3 |
| Breeding line 1058 | Norway | Validation population 3 |
| Breeding line 1059 | Norway | Validation population 3 |
| Breeding line 1060 | Norway | Validation population 3 |
| Breeding line 1061 | Norway | Validation population 3 |
| Breeding line 1062 | Norway | Validation population 3 |
| Breeding line 1063 | Norway | Validation population 3 |
| Breeding line 1064 | Norway | Validation population 3 |
| Breeding line 1065 | Norway | Validation population 3 |
| Breeding line 1066 | Norway | Validation population 3 |
| Breeding line 1067 | Norway | Validation population 3 |
| Breeding line 1068 | Norway | Validation population 3 |
| Breeding line 1069 | Norway | Validation population 3 |
| Breeding line 1070 | Norway | Validation population 3 |
| Breeding line 1071 | Norway | Validation population 3 |
| Breeding line 1072 | Norway | Validation population 3 |
| Breeding line 1073 | Norway | Validation population 3 |
| Breeding line 1074 | Norway | Validation population 3 |
| Breeding line 1075 | Norway | Validation population 3 |
| Breeding line 1076 | Norway | Validation population 3 |
| Breeding line 1077 | Norway | Validation population 3 |
| Breeding line 1078 | Norway | Validation population 3 |
| Breeding line 1079 | Norway | Validation population 3 |
| Breeding line 1080 | Norway | Validation population 3 |
| Breeding line 1081 | Norway | Validation population 3 |
| Breeding line 1082 | Norway | Validation population 3 |
| Breeding line 1083 | Norway | Validation population 3 |
| Breeding line 1084 | Norway | Validation population 3 |
| Breeding line 1085 | Norway | Validation population 3 |
| Breeding line 1086 | Norway | Validation population 3 |
| Breeding line 1087 | Norway | Validation population 3 |
| Breeding line 1088 | Norway | Validation population 3 |
| Breeding line 1089 | Norway | Validation population 3 |
| Breeding line 1090 | Norway | Validation population 3 |
| Breeding line 1091 | Norway | Validation population 3 |
| Breeding line 1092 | Norway | Validation population 3 |
| Breeding line 1093 | Norway | Validation population 3 |
| Breeding line 1094 | Norway | Validation population 3 |
| Breeding line 1095 | Norway | Validation population 3 |
| Breeding line 1096 | Norway | Validation population 3 |
| Breeding line 1097 | Norway | Validation population 3 |
| Breeding line 1098 | Norway | Validation population 3 |
| Breeding line 1099 | Norway | Validation population 3 |
| Breeding line 1100 | Norway | Validation population 3 |

| | | |
|---|---|---|
| Breeding line 1101 | Norway | Validation population 3 |
| Breeding line 1102 | Norway | Validation population 3 |
| Breeding line 1103 | Norway | Validation population 3 |
| Breeding line 1104 | Norway | Validation population 3 |
| Breeding line 1105 | Norway | Validation population 3 |
| Breeding line 1106 | Norway | Validation population 3 |
| Breeding line 1107 | Norway | Validation population 3 |
| Breeding line 1108 | Norway | Validation population 3 |
| Breeding line 1109 | Norway | Validation population 3 |
| Breeding line 1110 | Norway | Validation population 3 |
| Breeding line 1111 | Norway | Validation population 3 |
| Breeding line 1112 | Norway | Validation population 3 |
| Breeding line 1113 | Norway | Validation population 3 |
| Breeding line 1114 | Norway | Validation population 3 |
| Breeding line 1115 | Norway | Validation population 3 |
| Breeding line 1116 | Norway | Validation population 3 |
| Breeding line 1117 | Norway | Validation population 3 |
| Breeding line 1118 | Norway | Validation population 3 |
| Breeding line 1119 | Norway | Validation population 3 |
| Breeding line 1120 | Norway | Validation population 3 |
| Breeding line 1121 | Norway | Validation population 3 |
| Breeding line 1122 | Norway | Validation population 3 |
| Breeding line 1123 | Norway | Validation population 3 |
| Breeding line 1124 | Norway | Validation population 3 |
| Breeding line 1125 | Norway | Validation population 3 |
| Breeding line 1126 | Norway | Validation population 3 |
| Breeding line 1127 | Norway | Validation population 3 |
| Breeding line 1128 | Norway | Validation population 3 |
| Breeding line 1129 | Norway | Validation population 3 |
| Breeding line 1130 | Norway | Validation population 3 |
| Breeding line 1131 | Norway | Validation population 3 |
| Breeding line 1132 | Norway | Validation population 3 |
| Breeding line 1133 | Norway | Validation population 3 |
| Breeding line 1134 | Norway | Validation population 3 |
| Breeding line 1135 | Norway | Validation population 3 |
| Breeding line 1136 | Norway | Validation population 3 |
| Breeding line 1137 | Norway | Validation population 3 |
| Breeding line 1138 | Norway | Validation population 3 |
| Breeding line 1139 | Norway | Validation population 3 |
| Breeding line 1140 | Norway | Validation population 3 |
| Breeding line 1141 | Norway | Validation population 3 |
| Breeding line 1142 | Norway | Validation population 3 |
| Breeding line 1143 | Norway | Validation population 3 |
| Breeding line 1144 | Norway | Validation population 3 |
| Breeding line 1145 | Norway | Validation population 3 |
| Breeding line 1146 | Norway | Validation population 3 |

| | | |
|---|---|---|
| Breeding line 1147 | Norway | Validation population 3 |
| Breeding line 1148 | Norway | Validation population 3 |
| Breeding line 1149 | Norway | Validation population 3 |
| Breeding line 1150 | Norway | Validation population 3 |
| Breeding line 1151 | Norway | Validation population 3 |
| Breeding line 1152 | Norway | Validation population 3 |
| Breeding line 1153 | Norway | Validation population 3 |
| Breeding line 1154 | Norway | Validation population 3 |
| Breeding line 1155 | Norway | Validation population 3 |
| Breeding line 1156 | Norway | Validation population 3 |
| Breeding line 1157 | Norway | Validation population 3 |
| Breeding line 1158 | Norway | Validation population 3 |
| Breeding line 1159 | Norway | Validation population 3 |
| Breeding line 1160 | Norway | Validation population 3 |
| Breeding line 1161 | Norway | Validation population 3 |
| Breeding line 1162 | Norway | Validation population 3 |
| Breeding line 1163 | Norway | Validation population 3 |
| Breeding line 1164 | Norway | Validation population 3 |
| Breeding line 1165 | Norway | Validation population 3 |
| Breeding line 1166 | Norway | Validation population 3 |
| Breeding line 1167 | Norway | Validation population 3 |
| Breeding line 1168 | Norway | Validation population 3 |
| Breeding line 1169 | Norway | Validation population 3 |
| Breeding line 1170 | Norway | Validation population 3 |
| Breeding line 1171 | Norway | Validation population 3 |
| Breeding line 1172 | Norway | Validation population 3 |
| Breeding line 1173 | Norway | Validation population 3 |
| Breeding line 1174 | Norway | Validation population 3 |
| Breeding line 1175 | Norway | Validation population 3 |
| Breeding line 1176 | Norway | Validation population 3 |
| Breeding line 1177 | Norway | Validation population 3 |
| Breeding line 1178 | Norway | Validation population 3 |
| Breeding line 1179 | Norway | Validation population 3 |
| Breeding line 1180 | Norway | Validation population 3 |
| Breeding line 1181 | Norway | Validation population 3 |
| Breeding line 1182 | Norway | Validation population 3 |
| Breeding line 1183 | Norway | Validation population 3 |
| Breeding line 1184 | Norway | Validation population 3 |
| Breeding line 1185 | Norway | Validation population 3 |
| Breeding line 1186 | Norway | Validation population 3 |
| Breeding line 1187 | Norway | Validation population 3 |
| Breeding line 1188 | Norway | Validation population 3 |
| Breeding line 1189 | Norway | Validation population 3 |
| Breeding line 1190 | Norway | Validation population 3 |
| Breeding line 1191 | Norway | Validation population 3 |
| Breeding line 1192 | Norway | Validation population 3 |

| | | |
|---|---|---|
| Breeding line 1193 | Norway | Validation population 3 |
| Breeding line 1194 | Norway | Validation population 3 |
| Breeding line 1195 | Norway | Validation population 3 |
| Breeding line 1196 | Norway | Validation population 3 |
| Breeding line 1197 | Norway | Validation population 3 |
| Breeding line 1198 | Norway | Validation population 3 |
| Breeding line 1199 | Norway | Validation population 3 |
| Breeding line 1200 | Norway | Validation population 3 |
| Breeding line 1201 | Norway | Validation population 3 |
| Breeding line 1202 | Norway | Validation population 3 |

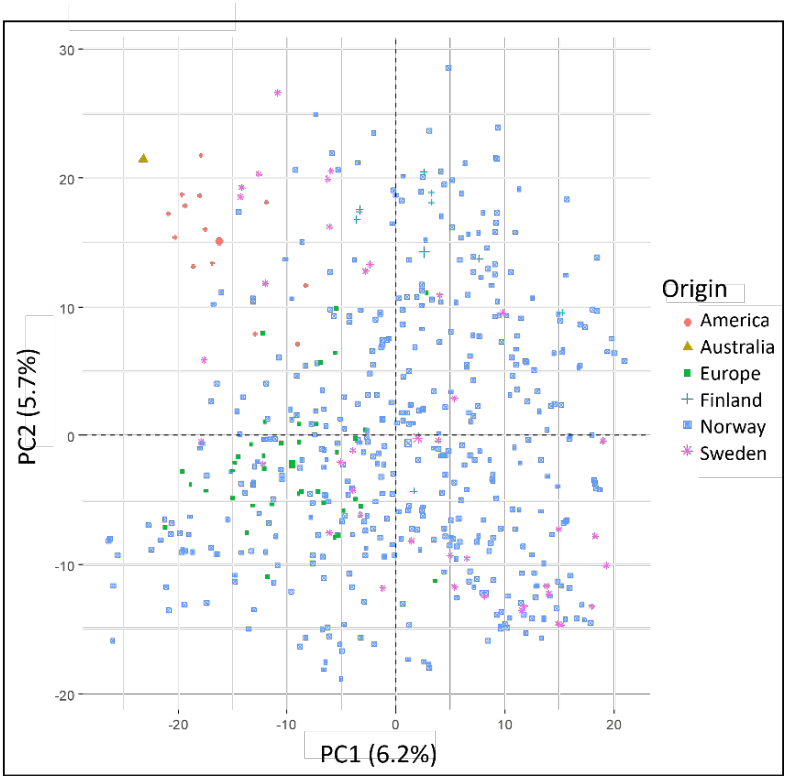



***Figure S4**: Plot of the first two principal components of the marker information for the GWAS panel sorted into their region of origin.*

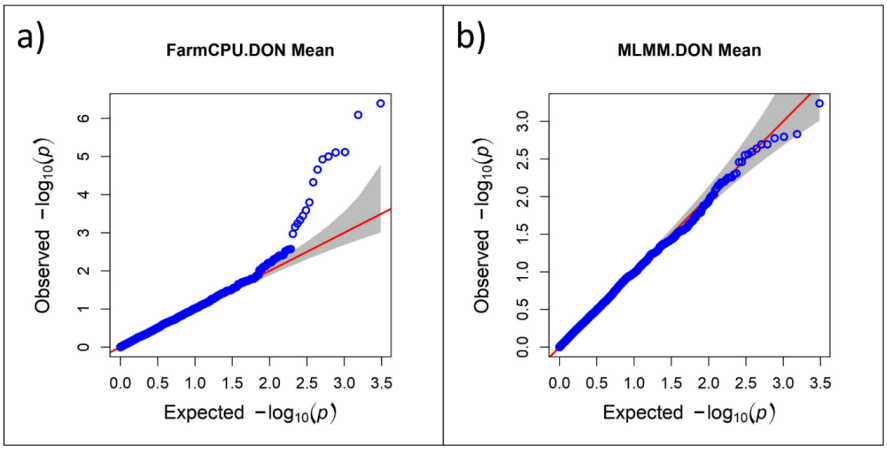



***Figure S5****: QQ-plots from GWAS analysis using Overall AdjDON data with the FarmCPU model (a) and the MLMM model (b) showing the deviation between the expected (X-axis) and observed (Y-axis) - $log_{10}$ p-values.*

***Table S3****: Candidate genes for fusarium resistance found within the validated QTL regions from Table 3.*

| QTL regions | 1C | 7C-17A-1 | 9D | 12D | 18D-1 |
|---|---|---|---|---|---|
| Distance between SNPs | 21.2 Mb | 47.3Mb | 4.1Mb | 23.2Mb | 4.2Mb |
| Copies of disease related genes | | | | | |
| Putative disease resistance protein RGA1 | 6 | 0 | 2 | 1 | 0 |
| Putative disease resistance protein RGA2 | 0 | 5 | 0 | 2 | 0 |
| Putative disease resistance protein RGA3 | 1 | 0 | 0 | 0 | 0 |
| Putative disease resistance protein RGA4 | 2 | 0 | 0 | 1 | 0 |
| Putative disease resistance protein RGA5 | 2 | 3 | 0 | 1 | 1 |
| Disease resistance protein RPM1 | 1 | 1 | 1 | 1 | 0 |
| Disease resistance protein RPS2 | 1 | 0 | 0 | 3 | 0 |
| Disease resistance protein Pik-1 | 0 | 0 | 0 | 0 | 1 |
| Disease resistance protein Pik-2 | 3 | 3 | 0 | 1 | 0 |
| Putative disease resistance protein RPP13-like protein 1 | 1 | 1 | 1 | 1 | 0 |
| Putative disease resistance protein RPP13-like protein 4 | 0 | 0 | 0 | 0 | 1 |
| Disease resistance protein Pik6-NP | 1 | 0 | 0 | 0 | 0 |
| Putative disease resistance protein At3g14460 | 1 | 2 | 0 | 0 | 0 |
| Enhanced disease resistance EDR4 | 1 | 0 | 0 | 0 | 0 |
| Disease resistance protein At1g50180 | 0 | 1 | 0 | 0 | 0 |
| Enhanced disease resistance EDR2-like | 0 | 0 | 1 | 0 | 0 |

***Table S3**: List of varieties with historical importance to the Norwegian germplasm with their name, country of origin, approximate year of $F_{10}$ generation, and validated QTL region from Table 3 with their haplotypes. The letters correspond to one of the four nucleotides A, T, C, G, while N is unknown identity.*

| Name | Origin | Year | 1C | 7C-17A-1 | 9D | 12D | 18D-1 |
|---|---|---|---|---|---|---|---|
| Probsteier | Denmark | 1891 | GCAG | CGA | TNT | GAC | GTC |
| Milton | USA | 1894 | GTCA | CGA | TCT | GAC | GTC |
| Stormogul | Sweden | 1895 | GCAG | TAA | TCT | TAT | AGT |
| Svala | Sweden | 1900 | GCAG | CGG | CTC | GAC | AGT |
| Seger | Sweden | 1902 | GTCA | CGA | TTT | TAT | GTC |
| Fyris | Sweden | 1905 | GTCA | CGG | CTC | GGT | GTC |
| FØRÆDLADDALAHAVR | Sweden | 1907 | GTCA | CGA | TTT | TAT | AGT |
| Grenader | Norway | 1912 | GTCA | CGA | TCT | GAC | GTC |
| ThorsMøystad | Norway | 1913 | GCAA | CGG | TCT | TGT | AGT |
| Gullregn_II | Sweden | 1914 | GTCA | CGA | TCT | GAC | GTC |
| Perle | Norway | 1914 | GCAG | CGG | CTC | GAC | AGT |
| Stjærn | Sweden | 1914 | GTCA | CGA | TTT | TAT | GTC |
| Vasa | Finland | 1917 | GTCA | CGA | TTT | GAC | GTC |
| Engelbrekt | Sweden | 1918 | GTCA | CAA | TCT | TGT | GTC |
| Nidar | Norway | 1918 | GCAG | CGG | CTC | GGT | AGT |
| Kytø | Finland | 1919 | GTCA | CGA | TTT | GAC | GTC |
| Hvit_Odal | Sweden | 1920 | GTCA | CGG | CTT | TGT | AGT |
| Orn | Sweden | 1920 | GTCA | CAA | CTC | GGC | AGT |
| KlockExtra | Denmark | 1927 | GCAG | CGA | TTT | TAT | GTC |
| Bambu | Sweden | 1928 | GTCA | CAA | TTT | TAT | GTC |
| Hird | Norway | 1934 | GTCA | CGG | TTT | GGC | AGT |
| Sol_II | Sweden | 1936 | GTCA | CAA | CTC | GGC | AGT |
| PaluAbed | Denmark | 1939 | GTCA | CGG | CTT | GAT | AGT |
| Blenda | Sweden | 1944 | GCAG | CGA | TTT | GGC | AGT |
| Norum | Norway | 1944 | GCAG | CGG | CTC | TAT | AGT |
| Voll | Norway | 1950 | GNCN | CGG | CTC | GGC | AGT |
| Hannes | Finland | 1958 | GTCA | TGA | TCT | GAC | GTC |
| Titus | Sweden | 1958 | GTCA | CGA | TTT | TAT | AGT |
| Pol | Norway | 1961 | GCAG | TGA | TCT | TAT | GTC |
| Selma | Sweden | 1962 | GTCA | CAA | CTT | GAC | GTC |
| Gråkall | Norway | 1966 | GTCA | CGG | CTC | GGC | AGT |
| Sang | Sweden | 1968 | GTCA | CGA | CTC | GAC | GTC |
| Svea | Sweden | 1970 | GTCA | CGA | CTC | GGC | AGT |
| Mustang | Netherlands | 1974 | GTCA | CGA | TTT | GGC | AGT |
| Moholt | Norway | 1976 | GTCA | CGG | CTC | GAC | AGT |
| Kapp | Norway | 1980 | ACAG | TGA | TCT | GGC | AGT |
| Lena | Norway | 1980 | GTCA | CAA | CTC | GAC | GTC |
| Magne | Sweden | 1981 | GTCA | CGA | TTT | GGC | AGT |
| Martin | Norway | 1982 | ACAG | CGA | TCT | GGC | GTC |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Frigg | Sweden | 1983 | GCAG | CGA | TCT | GAC | AGT |
| Freja | Sweden | 1985 | ACAG | CAA | TCT | GGC | AGT |
| Grane | Norway | 1986 | GCAG | CGG | CTT | GAC | AGT |
| Munin | Norway | 1986 | GCAG | CGA | TCT | GGC | GTC |
| Matilda | Sweden | 1988 | ACAG | CGA | TTC | GGC | GTC |
| Bikini | Norway | 1991 | GCAG | CGA | CTC | GGC | AGT |
| Biri | Norway | 1991 | GTCA | CAA | CTC | GGC | AGT |
| Belinda | Sweden | 1992 | ACAG | CGA | CTC | GGC | AGT |
| Hugin | Norway | 1993 | GTCA | CAA | CTT | GGC | AGT |
| Ingeborg | Sweden | 1998 | GCAG | TGG | TCT | GAC | AGT |
| Roverud | Norway | 1998 | GCAG | CAG | CTC | GAC | AGT |
| Eidsvoll | Norway | 1999 | GCAG | CGA | TCT | GGC | AGT |
| Flisa | Norway | 1999 | GTCA | CAA | CTC | GAC | GTC |
| Gere | Norway | 1999 | ACAG | CGA | TCT | GGC | AGT |
| Hurdal | Norway | 1999 | ACAG | CAG | CTC | GGC | AGT |
| Nudist | Norway | 2001 | GCAG | CGA | CTC | GGC | AGT |
| Ringsaker | Norway | 2002 | GTCA | CAA | CTC | GGC | AGT |
| Nes | Norway | 2003 | GCAG | CGA | CTC | GGC | AGT |
| Odal | Norway | 2003 | ACAG | CGA | CTC | GGC | AGT |
| Haga | Norway | 2004 | GCAG | CAA | CTT | GAC | AGT |
| Skarnes | Norway | 2004 | GCAG | CAA | CTC | GGC | AGT |
| Vinger | Norway | 2004 | GCAG | CGA | CTT | GGC | AGT |
| Hurum | Norway | 2007 | GCAG | CGA | CTC | GAC | GTC |
| Avetron | Norway | 2008 | ACAG | CAA | CTT | GAC | AGT |
| Gimse | Norway | 2008 | ACAG | CGA | TCT | GGC | AGT |
| Dovre | Norway | 2009 | GTCA | CAA | CTC | GAC | AGT |
| Våler | Norway | 2009 | ACAG | CGA | CTC | GGC | AGT |
| Årnes | Norway | 2009 | GCAG | CGG | CTC | GAC | AGT |
| Staur | Norway | 2012 | GCCG | CAA | CTC | GGC | AGT |
| Bingen | Norway | 2014 | GCAG | CGA | CTC | GAC | GTC |
| Mo | Norway | 2014 | NCAG | CGA | CTC | GAC | AGT |
| Ridabu | Norway | 2014 | GCAG | CGA | CTC | GAC | GTC |
| Brandval | Norway | 2016 | ACAG | CGA | TCT | GGC | AGT |
| Romedal | Norway | 2016 | GCAG | CAA | CTT | GGC | AGT |
| Vallset | Norway | 2016 | GCAG | CGA | CTT | GGC | AGT |

***Table S4**: Models for calculating BLUE-values for each field design where μ is the mean, ε is the error which is assumed to be independent and normally distributed with a mean of zero and variance of $\sigma^2$, and $Y_{ijkl}$ is the phenotype of the ith environment (Env) of the jth replication (Rep) of the kth incomplete block (Block) of the lth genotype (Gen) for the multiple environment models. For the single environments models $Y_{ijk}$ is the phenotype of the ith replicate (Rep) of the jth incomplete block (Block) of the kth genotype (Gen).*

| Experimental design | Model to calculate BLUE-values |
|---|---|
| Single trial alpha | $Y_{ijk} = \mu + Rep_i + Block_j\ (Rep_i) + Gen_k + \varepsilon_{ijk}$ [Eq.1] |
| Overall trials alpha | $Y_{ijkl} = \mu + Env_i + Rep_j\ (Env_i) + Block_k(Env_i\ Rep_i) + Gen_l + Env_i \times Gen_l + \varepsilon_{ijkl}$ [Eq.2] |
| Single trials RCBD | $Y_{ik} = \mu + Rep_i + Gen_k + \varepsilon_{ik}$ [Eq.3] |
| Overall trials RCBD | $Y_{ijkl} = \mu + Env_i + Rep_j\ (Env_i) + Gen_l + Env_i \times Gen_l + \varepsilon_{ijkl}$ [Eq.4] |

***Table S5**: Trial statistics of individual experiments (20S, 20V, 21S, 21V, 22S, 22V) and overall (Ov) for the traits DON and GP including broad sense Heritability ($H^2$), Variance (Var) and significance level (Sig) of the factors Genotype (Gen), Environment (Env) Genotype x Environment (Gen x Env), and Residual (Res), Grand Mean and Coefficient of variance (CV).*

| Statistics | 20S | 20V | 21S | 21V | 22S | 22V | Ov |
|---|---|---|---|---|---|---|---|
| | | | | **DON** | | | |
| **$H^2$** | 0.69 | 0.68 | 0.49 | 0.73 | 0.59 | 0.67 | 0.79 |
| **Gen Var** | 2.41 | 1.39 | 0.62 | 0.48 | 2.57 | 3.31 | 1.08 |
| **Env Var** | | | | | | | 6.00 |
| **GenxEnv Var** | | | | | | | 0.71 |
| **Res Var** | 2.18 | 1.33 | 1.29 | 0.37 | 3.51 | 3.24 | 1.99 |
| **Grand Mean** | 3.45 | 3.78 | 3.04 | 2.01 | 6.13 | 9.11 | 4.59 |
| **CV** | 42.78 | 30.44 | 37.35 | 30.10 | 30.56 | 19.77 | 30.75 |
| **Gen Sig** | 2.26E-38 | 1.98E-30 | 2.17E-13 | 1.36E-44 | 2.13E-20 | 2.32E-31 | 3.87E-150 |
| **Env Sig** | | | | | | | 0.02 |
| **GenxEnv Sig** | | | | | | | 5.31E-38 |
| | | | | **GP** | | | |
| **$H^2$** | 0.47 | 0.20 | 0.41 | 0.57 | 0.57 | 0.61 | 0.62 |
| **Gen Var** | 17.15 | 3.90 | 12.60 | 23.85 | 33.02 | 35.24 | 8.51 |
| **Env Var** | | | | | | | 27.61 |
| **GenxEnv Var** | | | | | | | 12.01 |
| **Res Var** | 38.85 | 32.01 | 35.69 | 36.48 | 48.95 | 44.30 | 39.53 |
| **Grand Mean** | 79.11 | 80.19 | 74.98 | 70.43 | 69.06 | 65.19 | 73.14 |
| **CV** | 7.88 | 7.06 | 7.97 | 8.57 | 10.13 | 10.21 | 8.60 |
| **Gen Sig** | 1.88E-12 | 0.02 | 2.51E-08 | 4.74E-19 | 3.90E-19 | 1.41E-21 | 4.07E-50 |
| **Env Sig** | | | | | | | 0.08 |
| **GenxEnv Sig** | | | | | | | 2.69E-30 |

# Paper III

# Evaluation of genomic prediction for Fusarium resistance in the Norwegian oat breeding program


## Authors' name
Espen Sannes Sørensen[1,2], Susanne Windju[1], Constantin Jansen[1], Anna Kristina Sonesson[3], Morten Lillemo[2], Muath Alsheikh[1,2]

## Institutes
[1]Graminor AS, Ridabu, Norway.

[2]Department of Plant Science, Norwegian University of Life Sciences, Ås, Norway.

[3]Nofima AS, Ås, Norway.

## Correspondence
Espen Sannes Sørensen. E-mail: espen.sorensen@nmbu.no.


## Key message
In this study five validates QTL-regions associated to FHB resistance were added as fixed effects to improve prediction ability for genomic prediction of DON and Germination percentage in oat. Each region contributed to increased prediction ability in the training population, and reduced DON content in selection of elite material compared to a basic model. But the results from three breeding populations showed inconsistent results.


## Keywords
Fusarium resistsance*;* Genomic prediction; QTL-regions; Fixed effects; Oat


## Conflict of interest
All authors have no conflict of interest to declare that are relevant to the content of this article.

## Contribution of authors
ESS analysed data from field trials, organized the germination analysis and performed the GWAS analysis, MA provided funding, CJ provided management of staff for phenotyping DON and field trials at Staur, SW, ML and AKS provided support in GWAS analysis, ESS wrote the manuscript, and MA, SW, ML, AKS and CJ participated in revising the scientific work and writing.

## Running title
Evaluation of genomic prediction for FHB resistance in oat.


## Acknowledgements
This research was funded by the Research council of Norway (The research funds for agriculture and food industry -FFL/JA) and Graminor ((NFR # 286606). The authors wish to thank Graminor for providing the material and facilities for this research, and the technical staff at Graminor and NMBU for helping in field experiments and data collection.

# Abstract

Fusarium head blight (FHB) is the predominant disease in oat in Norway caused by the fungus *Fusarium graminearum*. It causes yield loss, reduced seed quality, reduced germination ability and accumulation of deoxynivalenol (DON). The FHB resistance is quantitative, and most genes have small effect. A previous GWAS study identified five QTL-regions with large effect associated with FHB resistance validated across environments and populations. This study aims to use these QTL regions as fixed effects to improve genomic prediction for DON and Germination ability in oat. A large and diverse training population of 541 lines was sown in field trials for three years and two locations obtain accurate phenotypic data. The trials were spawn inoculated and mist irrigated to ensure a high and even disease pressure. The results showed relatively high prediction ability for both DON and germination ability, both in the training population and in three validation population, especially when heritability of the trait was high. The added effect of QTL-regions increased prediction ability individually and collectively in the training set but showed inconsistent results in the validation populations where prediction ability sometime was reduced compared to a basic model. A selection of elite material identified that the targeted resistant alleles did increase in frequency, but that some of the other resistance alleles from the GWAS were reduced in selections indicating that there might be a linkage between the resistant and non-resistant alleles of these QTL-regions due to the family structure of the populations. The complexity of FHB resistance in oat with multiple small effect QTL, several types of resistance, high GxE interaction, correlated traits and potential epistatic effects emphasise the need to develop more complex models in order to capture these effects. The results from this study will be used for implementation of genomic selection in the Norwegian breeding program.

# 1. Introduction

In Norway, oat is one of the three major crops along with wheat and barley (Statistisk sentralbyrå, 2023). Oat has been considered a good break crop in rotation with barley and wheat as they have few diseases in common (Abrahamsen *et al.,* 2016). There are several diseases that impacts oat production in Norway like powdery mildew and leaf blotch, However, Fusarium head blight (FHB) has been highlighted the major disease with big economic impact on the entire oat industry (Bernhoft *et al.,* 2013). FHB is a fungal disease caused by several different Fusarium species, and causes yield loss, reduced seed quality and germination ability, and accumulation of mycotoxins that are harmful to both humans and animals. The most important Fusarium pathogen in Norway is *Fusarium graminearum*, which produces the mycotoxin deoxynivalenol (DON) for which the EU and Norway have set a limit on unprocessed grain to 1750 µg/kg, and farmers gets reduced price if the shipments exceed the limits. In the period of 2010-2012 there were high infection pressure and elevated DON content, which caused approximately 40% of the shipments these years to exceed this limit (Felleskjøpet agri SA, 2016). The available fungicides in Norway have only a partial effect on the disease, reducing DON content with maximum 50% if applied at the optimal time (Felleskjøpet Agri SA, 2023). The pathogen reproduces sexually (Aamot *et al.,* 2015) and can therefore change in aggressiveness, virulence and potentially achieve fungicide resistance (Becher *et al.,* 2010, de Chaves *et al.,* 2022). The use of resistant varieties along with other management options could be a more effective way to reduce the risk of infection. The resistance is quantitative in nature and race non-specific (He *et al.,* 2013, Bjørnstad *et al.,* 2017) which likely contributes to durability of the resistance. With a predicted future Norwegian climate with more rainfall and humid conditions that are favourable for the fungus, it is important that the resistance breeding continues in order to prevent large economic losses due to FHB epidemics.

Resistance to Fusarium can be categorized into five main types (Hautsalo *et al.,* 2018); (i) resistance to initial infection (Schroeder & Christensen, 1963), (ii) resistance to disease spread, (iii) resistance to kernel infection, including germination ability (Mesterházy, 1995), (iv) tolerance (Mesterházy, 1995), and (v) resistance to mycotoxin accumulation (Miller *et al.,* 1985, Mesterházy *et al.,* 1999). Additionally, there are passive avoidance mechanisms that the plants use to avoid infection. Three of these are plant height, where taller plants increase the distance to the initial infection, days to heading, which makes the plants flower in periods of lower disease pressure, and anther extrusion, which ensures that the anthers are not stuck between palea and lemma, leaving openings for the fungus to enter the florets and grow on the dead anther tissue. Fusarium resistance is measured in several different ways. The four main approaches are to (i) measure mycotoxin content, (ii) count

number of infected spikelets, (iii) measure germination percentage (GP), and (iv) quantify the amount of fusarium DNA in a qPCR. These traits are correlated and deal with different parts of the resistance mechanisms.

Traditionally, oat breeding in Norway has been done with phenotypic selection where inbred lines from each cross are selected and evaluated, first in head rows for simple traits like height and earliness and later for yield and quality traits. For the past 12 years, Norwegian material has been screened for fusarium resistance in generations $F_{10}$-$F_{12}$ in spawn-inoculated and mist-irrigated disease trials at two locations. Such trials are expensive, but efficient to provoke large variation in DON content and GP and thus increase selection accuracy for FHB resistance.

Genomic selection (Meuwissen *et al.,* 2001) is a breeding strategy that could help select for FHB resistance at an early stage, while reducing the need to do large expensive field testing. It uses statistical models to predict the phenotypes of new individuals using only their genotypic information. The models are trained with a training population that has been genotyped and phenotyped for the given trait. Marker effects will be estimated based on a prior distribution, giving each marker a small effect that is summed into the total predicted breeding value of the genotype. The accuracy of the genomic prediction models is determined by the number of individuals used in the training and their relatedness to the breeding material, the accuracy of the phenotyping, the heritability of the trait and the density of the markers. Fusarium resistance is a complex trait with many small-effect quantitative trait loci (QTL) that contribute to resistance, and large genotype x environment interactions that give the trait relatively low heritability (He *et al.,* 2013, Hautsalo *et al.,* 2018, Hautsalo *et al.,* 2020). There is also a low accuracy in phenotyping compared to other traits like days to heading and plant height, which makes fusarium resistance a difficult trait to predict.

There are several ways to improve genomic prediction models. Some examples of successful improvement of FHB in wheat are to include known QTL information (Rutkoski *et al.,* 2012), correlated traits (Gaire *et al.,* 2022) or environmental factors (Zhang *et al.,* 2021) as co-factors in the models. There are few examples of genomic prediction for fusarium resistance in oat, and in one that have been conducted do not include co-variates in the models (Haikka et al. 2020). This study will utilize a large and diverse training population tested in several environments to predict new breeding material for DON accumulation and germination ability by including QTL information from a previous mapping study in the prediction models. The aim is to evaluate genomic prediction models for DON and GP both within the training population and in the Norwegian oat breeding program, and to give general advice on the utilization of fixed effect models for breeding programs.

# 2. Materials and methods

## 2.1. Materials

### 2.1.1. Training population

The training population used in this study consisted of 541 oat lines and cultivars selected from a larger pool of approximately 1,124 lines and cultivars by using a combination of the three selection strategies described by Sørensen *et al.* (2023). These strategies aimed to maintain high phenotypic and genotypic diversity in smaller populations and ensured that the training population remained highly related to the material in the Norwegian breeding program. In total, 440 lines were from Norway, 40 from Sweden, 26 from Germany, nine from Netherland, eight from Canada, eight from Finland, five from USA, two from Austria, one from Denmark and one from Australia (Supplementary 1, Table 1). Some lines were excluded to avoid strong association between correlated traits and FHB resistance. These were (1) very tall plants, as increased height increases the risk for lodging which in turn creates favorable humid microclimate for the pathogen to thrive, (2) very short plants as it shortens the distance from the heads to the primary inoculum, and (3) lines without husks as more fungus and thus DON is present on the husks (Brodal *et al.,* 2020). These traits could give large effect to loci present in families that share these traits, or large effect to QTL for these traits that are not present in the breeding material. These traits include plant height and presence of husks.

### 2.1.2. Validation populations

Three breeding populations were used to validate the results from the cross-validation, and will be called V1, V2 and V3 from this point. V1 and V2 consist of 242 $F_{10}$ breeding lines from Graminor from the years 2020 and 2021 respectively. $F_{10}$ lines have not been selected for Fusarium resistance. V3 consisted of 230 lines where 112 were $F_{10}$ lines from 2022, 88 were $F_{11}$ from V2, 22 were $F_{12}$ from V1 and 8 were $F_{13}$ from the breeding population of 2019.

## 2.2. Field trials

Field experiments of the training population were conducted for three years from 2020 to 2022 in two locations, Vollebekk (59.66°N, 10.75°E) and Staur (60.73°N, 11.10°E). V1, 2 and 3 were tested in 2020, 2021 and 2022 respectively in both locations. The experimental design of the training populations, V1 and V2 were alpha lattice design with two replicates and sub-block size of 5. The experimental design of V3 was randomized complete block (RCBD). All material were sown in spawn inoculated and mist irrigated disease trials (Tekle *et al.,* 2018).

## 2.3. Phenotyping

All plots were scored for Days to Heading (DTH) as the number of days from sowing to at least 50% of the heads have emerged, and Plant Height (PH) in centimeters from the ground to the top of the plant. DON accumulation was measured on milled seed samples with husks as parts per million (ppm) in the lab at Graminor with an Agraquant Deoxynivalenol Plus (0.25/5), 96 Wells ELISA kit developed by Romer Labs Ltd. Germination ability was measured at Graminor as percentage of germinated (GP) seeds using the "between paper" method described in point 5.6.2.1.1 of the ISTA protocol (ISTA, 2021). Two replicates of 50 seeds were used per plot. The paper rolls were put in plastic bags to retain moisture, stored in cool temperatures between 5 and 10°C for 7 days, then moved to room temperature for 6-8 days before analysis. Paper size was 220 X 400 mm and 200 x 400 mm with a capillary capacity of 80 mm/10 min and produced by Munktell Filter AB. DON and GP were collected for all trials except GP for V3 Staur.

## 2.4. Phenotypic data analysis

The Best linear unbiased estimators (BLUE) were calculated for each genotype in each trial and across trials (Overall), with the following models in the META-R software (Alvarado *et al.,* 2016):

Overall trials alpha: $Y_{ijkl} = \mu + Env_i + Rep_j\,(Env_i) + Block_k(Env_i\,Rep_i\,) + Gen_l + Env_i \times Gen_l + \varepsilon_{ijkl}$ [Eq.1]

Single trials RCBD: $Y_{ik} = \mu + Rep_i + Gen_k + \varepsilon_{ik}$ [Eq.2]

Overall trials RCBD: $Y_{ijkl} = \mu + Env_i + Rep_j\,(Env_i) + Gen_l + Env_i \times Gen_l + \varepsilon_{ijkl}$ [Eq.3]

$Y_{ijkl}$ is the phenotype of the *i*th envrionment (Env) of the *j*th replication (Rep) of the *k*th incomplete block (Block) of the *l*th variety (Gen). $Y_{ik}$ is the phenotype of the *i*th replicate (Rep) of the *k*th variety (Gen). $\mu$ is the mean, $\varepsilon$ is the error which is assumed to be independant and normally distributed with a mean of zero and variance of $\sigma^{2}$.

## 2.5. Genotyping and data preparation

The training population was genotyped with a customized, unpublished 20 kSNP array containing 18,598 markers. The array contained all markers from the publicly available 6 k SNP array (Tinker *et al.,* 2014). The validation populations were genotyped by a different 7kSNP array (Polley *et al.,* 2023), which contained the most polymorphic markers from the previous 20K array. The genetic data were filtered based on the training population with a threshold of 10% for missing values/heterozygotes and 5% MAF based on the training population, resulting in 3071 polymorphic markers.

## 2.6. Previous association mapping

A genome wide association study (GWAS) was done on the training population using phenotypic data from individual experiments and V1-3 to identify and validate important QTL-regions. The GWAS detected markers on 15 different QTL-regions, and resistant and susceptible alleles were identified for each region. Five regions were validated in a separate haplotype analysis which showed consistent significant effect across environments. The QTL regions are from this point referred to as haplotypes and named 1C, 7C-1, 9D, 12D and 18D-1 in reference to the linkage groups they were detected in. This study aims to use the allele information for each of these haplotypes in separate models and in combination to improve predictions.

## 2.7. Genomic prediction

### 2.7.1. Models

In this study, the Bayesian Ridge Regression model was used as the baseline model for the predictions and is given as:

$$Y_i = \mu + \sum_{k=1}^{p} X_{ik}\beta_k + \varepsilon_i \text{ [Model 1]}$$

Where Y is the BLUEof the *i*th individual, μ is the intercept, X is the marker value of the *k*th marker of the *i*th individual, β is the random marker effect of the *k*th marker, and ε is the error term of the *i*th individual. Both β and ε are assumed to be normally distributed with a mean of 0.β has a variance of $\sigma^2$ of the *k*th marker effects and ε has a variance of $\sigma^2$ of the *i*th individual.

The fixed effect models are written as:

$$Y_i = \mu + \sum_{k=1}^{p} X_{ik}\beta_k + \sum_{h=1}^{q} Z_{ij}u_j + \varepsilon_i \text{ [Model 2]}$$

Where the markers are modelled the same as in Model 1 with the additional terms of Z which is the allele values of the *j*th haplotype of the *i*th individual, and u, which is the fixed effect of the *j*th haplotype. The haplotypes are coded as 1, 0 and -1 for the resistant, neutral and susceptible alleles, respectively, in accordance with the results given by the previous haplotype analysis. The single haplotype models are named M1C, M7C-1, M9D, M12D and M18D-1 after the respective haplotypes used as fixed effects in the models. Two models use multiple haplotypes, named M5HT which uses the five haplotypes from the single haplotype models, and M15HT which also uses the additional 10 haplotypes detected in the GWAS. The genomic prediction was done using only the overall values for all experiments available for each population, and accuracy was collected by correlating the predicted values with the overall observed values of the validation populations. The DON values were log transformed to obtain normally distributed values for the GS analysis.

### 2.7.2. Cross-validation and prediction accuracy

The models were used in a five-fold cross-validation setup where the training set was randomly split into five groups and each group in turn was given missing phenotypes while the rest were used to train the models. This procedure was repeated 30 times, and Pearson correlations between predicted and observed phenotypes were collected and will from here on be referred to as prediction ability (PA). The models were evaluated with a Tukey pairwise comparison to see if they yielded significantly different mean PA. The same models were used to predict the phenotypes of the three validation populations with the full training set.

## 2.8. Selections

In the Norwegian breeding program 1/3 of the lines from $F_{10}$ are usually advanced for the next cycle of retesting. In this study we used this as a scenario to select 1/3 of the lines with lowest predicted logDON values for each cross-validation replication and validation population. The selections were then analysed for average observed DON, GP, DTH and PH, and for the frequency of the 15 resistant alleles from the previous GWAS in each model. The phenotypes and allele frequencies were then compared to the base BRR model to see how the fixed effects models affected the lines selected. The results from the training population were compared with Tukey pairwise comparisons using Minitab statistical software (Minitab, 2021) to see if there were significant differences. In addition, a selection based on the observed values were done to compare the allele frequencies of phenotypic and genomic selection. Selections were not done for GP because of the low heritability and genetic significance in V1 and V2 (Table 1).

# 3. Results

## 3.1. Phenotypic analysis and correlations

The trial statistics show that there were large differences between the years of the testing populations with low infections rates in 2020 and 2021 where V1 and V2 were tested respectively with a grand mean of 2.28 ppm DON in both populations, compared to 7.05 ppm for V3 tested in 2022. The heritability was also lower in V1 and V2 than V3 for both traits, and exceptionally so for GP with heritability of 0.19 and 0.18 for V1 and V2, and 0.79 for V3. In V1, the environmental variance was very large (123.42) for GP, while in V2 there was a large residual variance (94.78). The Genotype main effect is highly significant for DON in all populations, but only in the TP and V3 for GP.

***Table 4**: Trial statistics of the overall values of DON and GP for the validation populations (V1-3) and training population (TP) including broad sense Heritability ($H^2$), Variance (Var) and significance level (Sig) of the factors Genotype (Gen), Environment (Env) Genotype x Environment (Gen x Env), and Residual (Res), Grand Mean, Minimum values (Min), Maximum value (Max) and Coefficient of variance (CV).*

| | DON | | | | GP | | | |
|---|---|---|---|---|---|---|---|---|
| | **V1** | **V2** | **V3** | **TP** | **V1** | **V2** | **V3** | **TP** |
| **$H^2$** | 0.48 | 0.54 | 0.81 | 0.79 | 0.19 | 0.18 | 0.79 | 0.62 |
| **Gen Var** | 0.22 | 0.33 | 2.05 | 1.08 | 2.97 | 5.58 | 33.10 | 8.51 |
| **Env Var** | 0.14 | 0.68 | 0.31 | 6.00 | 123.42 | 2.18 | 20.11 | 27.61 |
| **Gen x Env Var** | 0.13 | 0.06 | 0.62 | 0.71 | 10.27 | 4.59 | 10.34 | 12.01 |
| **Res Var** | 0.71 | 1.00 | 4.41 | 1.99 | 29.17 | 94.78 | 67.50 | 39.53 |
| **Grand Mean** | 2.28 | 2.28 | 7.05 | 4.59 | 79.33 | 72.25 | 68.51 | 73.14 |
| **Min** | 0.86 | 0.87 | 3.27 | 2.48 | 65.82 | 56.25 | 33.96 | 60.51 |
| **Max** | 6.21 | 6.1 | 12.00 | 8.56 | 88.14 | 85.75 | 87.96 | 83.15 |
| **CV** | 37.07 | 43.83 | 29.8 | 30.75 | 6.81 | 13.47 | 11.99 | 8.60 |
| **Gen Sig** | 5.11E-07 | 6.74E-11 | 2.40E-23 | 3.87E-150 | 0.11 | 0.10 | 4.82E-11 | 4.07E-50 |
| **Env Sig** | 0.37 | 0.02 | 0.16 | 0.02 | 0.03 | 0.79 | 0.06 | 0.08 |
| **Gen x Env Sig** | 5.52E-03 | 0.22 | 1.30 | 5.31E-38 | 9.54E-06 | 0.36 | 0.02 | 2.69E-30 |

The correlations for the validation populations (Table 2) show that even though the heritability and genotypic significance were low for GP in V1 and 2, there was still a significant negative correlation between DON and GP in those populations (-0.13 and -0.14). DON showed a consistent positive significant correlation (0.18-0.24) with DTH in all populations, while GP showed significant negative correlation to DTH in V2 and V3. There was a significant negative correlation between PH and DON in V2 and V3 (0.15 and 0.17), and a significant positive correlation between GP and PH in V1 (0.21).

***Table 5**: Pearson correlations between the overall values of days to heading (DTH), plant height (PH), DON accumulation (DON) and germination percentage (GP) in the three validation populations (V1-3)*

| Population | Trait 1 | Trait 2 | Correlation | p-value |
|---|---|---|---|---|
| V1 | DON | DTH | 0.212 | 0.001 |
| | DON | PH | -0.062 | 0.317 |
| | DON | GP | -0.133 | 0.032 |
| | GP | DTH | -0.001 | 0.988 |
| | GP | PH | 0.212 | 0.001 |
| V2 | DON | DTH | 0.237 | 0 |
| | DON | PH | -0.156 | 0.012 |
| | DON | GP | -0.141 | 0.023 |
| | GP | DTH | -0.235 | 0 |
| | GP | PH | -0.062 | 0.319 |
| V3 | DON | DTH | 0.179 | 0.007 |
| | DON | PH | -0.17 | 0.01 |
| | DON | GP | -0.488 | 0 |
| | GP | DTH | -0.165 | 0.013 |
| | GP | PH | 0.02 | 0.767 |

## 3.2. Prediction ability

### 3.2.1. Training population

A significance test of the PA of DON and GP in the training population using the Bayesian ridge regression (BRR), BayesA, BayesB, BayesC and Bayesian LASSO showed no significant differences (data not shown). BRR were chosen as the basic model for further analysis because it assumes that most markers have small effect on the phenotypes which matched previous assumptions of the quantitative nature of FHB resistance in oat (He *et al.,* 2013, Bjørnstad *et al.,* 2017).

Predictions of logDON (Figure 1a) with single haplotype models showed significantly increased PA compared to BRR from a mean of 0.493 to 0.505, 0.509 and 0.504 with M7C-1, M9D and M12D respectively. M1C increased mean PA to 0.498 but were not significantly different from BRR, while M18-1 did not differ from the BRR in mean PA. For the multiple haplotypes models the M5HT showed significantly higher PA than any single haplotype model with a PA of 0.524, while M15HT showed significantly higher PA than M5HT with a mean of 0.545.

Predictions of GP (Figure 1b) with single haplotype models showed no statistical difference in PA compared to BRR. For the multiple haplotype models M5HT showed a mean PA of 0.538 which was significantly higher than BRR of 0.523 while M15HT showed a mean PA of 0.558 which was significantly higher than M5HT.

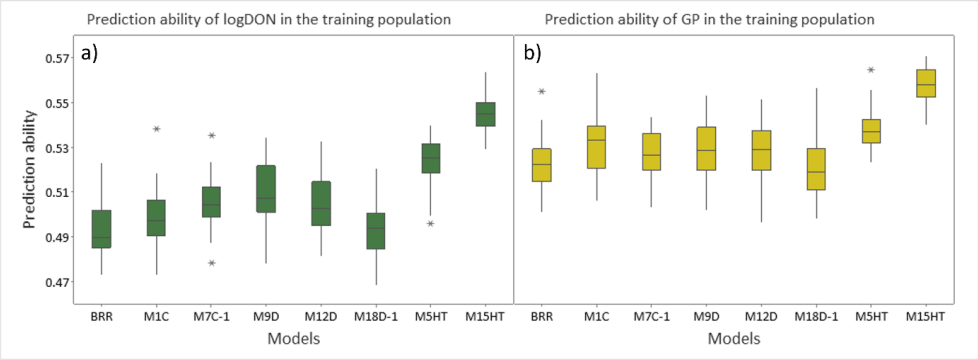



***Figure 6**: Prediction ability of log-transformed DON values (logDON) and Germination percentage (GP) in the training population using 5-fold cross-validation. Each model was replicated 30 times.*

### 3.2.2. Validation populations

The same models were applied to the validation populations and the results are showed in Figure 2. This section will highlight the notable differences between the fixed effect models and BRR. Notable difference here is defined as equal or bigger difference than 0.01 in PA. Some of the results were inconsistent across populations and traits, but the main trends for the different models were the following. M1C did not differ notably from BRR in terms of PA, except for an increase for logDON in V1 and decrease for GP in V2. M7C-1 showed higher or equal PA to BRR in V1 and V3 for both logDON and GP, but reduced PA in V2 for both traits. M9D showed PA equal or higher than BRR for logDON and slightly lower for GP in all populations. M12D did not show higher PA than BRR in any population or trait. M18D-1 showed higher PA than BRR for GP in all population, and no notable difference for logDON. M5HT showed notably higher PA for logDON in V1 and for GP in V3, but notably lower for logDON in V2 and V3, and for GP in V2. M15HT showed notably large decrease in PA for logDON in all population, and in V2 and V3 for GP.

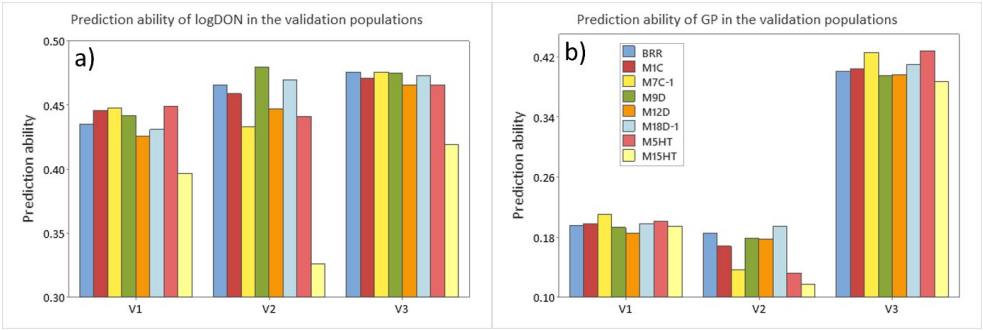



***Figure 7:*** *Prediction ability of log-transformed DON values (logDON) and Germination percentage (GP) in the three validation populations V1, V2 and V3 for the different models.*

## 3.3. Selections

For each model 1/3 of the lines with lowest predicted logDON values were selected for analysis of phenotypes and haplotype frequencies. Additionally, selections were done for the observed logDON values in order to compare genomic and phenotypic selection in terms of allele frequencies.

### 3.3.1. Phenotypes

The observed phenotypes of DON, GP, DTH and PH were averaged in each selection from the 30 replicates of the cross-validation for each model and analysed for significant differences (Figure 3). Similarly, the observed phenotypes of the selections done for each model in the validation populations were compared to each other (Figure 4).

#### *3.3.1.1. Training population*

Of the single haplotype models only M9D showed significantly lower observed DON values than BRR with 0.04 ppm mean difference (Figure 3a), but it was also the only one that showed significantly lower DTH than BRR with 0.04 days (Figure 3c). M1C, M7C-1 and M9D showed significantly higher GP than BRR with mean differences of 0.1, 0.1 and 0.09 % respectively (Figure 3b). M12D also showed significantly higher PH than BRR with a mean difference of 0.18 cm (Figure 3d). Both M5HT and M15HT showed significantly lower DON (Figure 3a) than the M9D with a mean difference of 0.03 and 0.04 ppm, respectively, but they were not significantly different from each other. M5HT showed significantly higher GP (Figure 3b) than any single haplotype model with 0.2 higher mean GP than M1C, while M15HT showed significantly higher GP than M5HT with a mean difference of 0.25. Neither M5HT nor M15HT were significantly different from BRR in terms of DTH (Figure 3c), but both showed significantly higher PH (Figure 3d) than BRR with a mean difference of 0.22 and 0.45 cm.

There was no significant difference between M5HT and M12D, but M15HT had significantly higher PH than both.

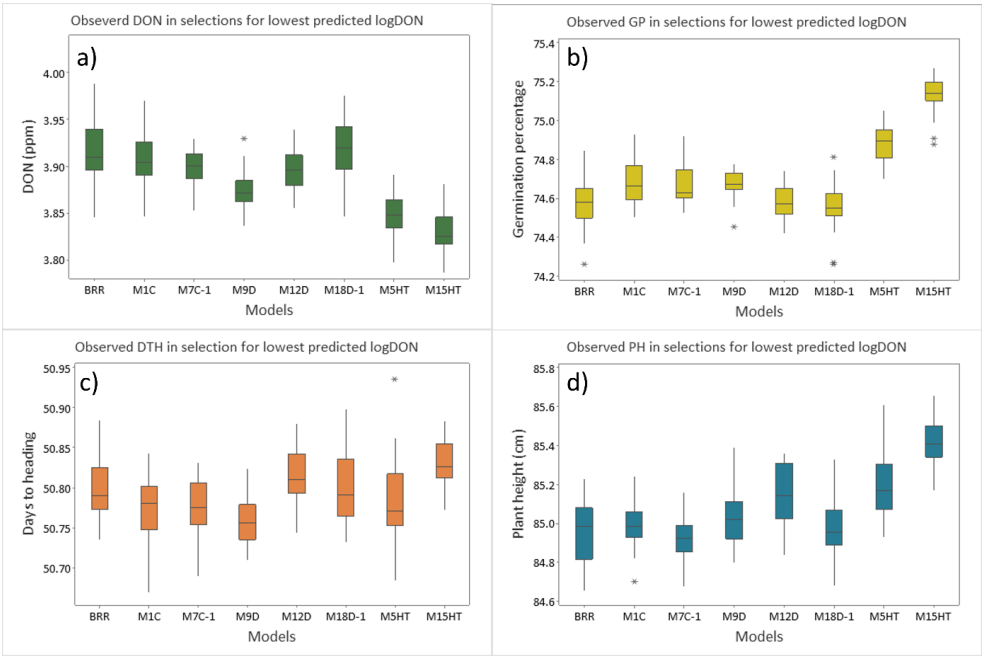



***Figure 3:*** *Average observed phenotypes of DON (a), Germination percentage (GP; b), Days to heading (DTH; c) and Plant height (PH; d) in selections based on the lines with 33% lowest predicted logDON with each model in the 30 cross-validation replicates.*

### *3.3.1.2. Validation populations*

For the single haplotype models in V1 there were notably higher DON for M12D than BRR (Figure 4a), while the rest were at the same level as BRR. For GP there was an increase with M9D, and a slight decrease with M12D (Figure 4b). However, V1 showed low variation in GP in selections with difference of 0.2 between the lowest and highest. In V2 there were notably lower DON (Figure 4c) and higher GP (Figure 4d) than BRR for M12D and M18D-1, and higher GP for M1C. In addition, M1C showed notably higher GP and M7C-1 showed lower DON. However, M7C-1 also showed notably lower GP. In V3 there were notably lower DON than BRR for M7C-1 and M18D-1, and higher GP for M1C and M9D. M18D-1 also showed notably lower GP.

M5HT showed notably higher DON values than BRR in all populations, and the lowest GP in V2 and V3 of all the models. M15HT showed the highest DON values in all populations, but notably higher GP than M5HT in V2 and V3. In V1, both M5HT and M15HT showed higher GP than BRR. Data for DTH

and PH are shown in Supplementary Figure S1 and shows that there were no notable differences for DTH between the models in any population. Of note for PH is that M15HT selections consistently were the shortest of the models. No models resulted in taller plants than with BRR.

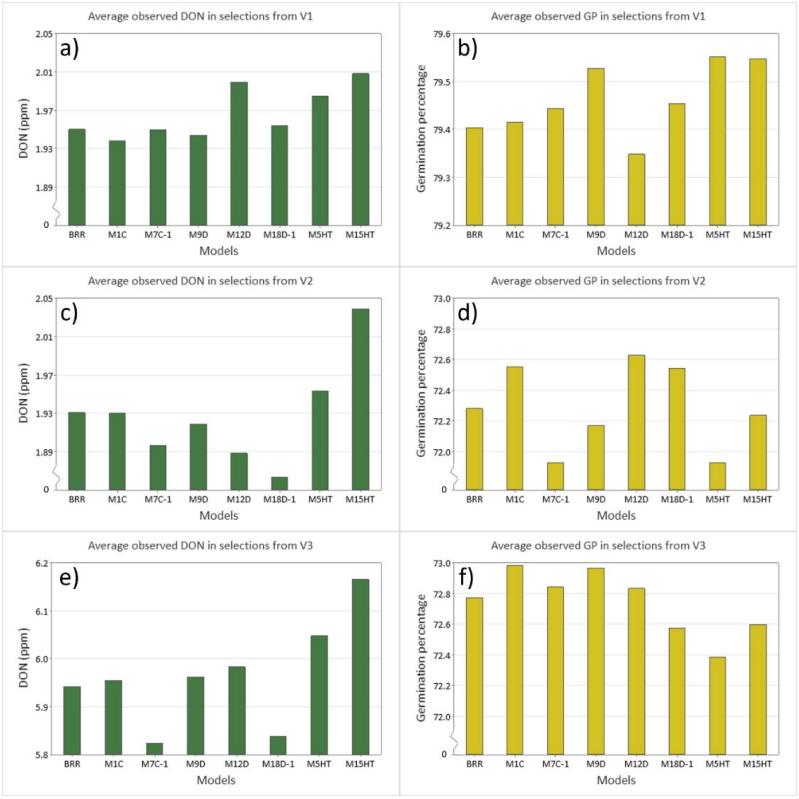



***Figure 4:*** *Average observed DON values (DON; a, c, e) and Germination percentage (GP; b, d, f) in selections made in Validation populations V1(a & b), V2(c & d) and V3 (e & f). Selections consist of 1/3 of the lines with lowest predicted logDON values with each model.*

### 3.3.2. Allele frequency

#### *3.3.2.1. Total population, phenotypic selection, and genomic selection*

The frequency of the five validated resistant alleles in each population (Training population, V1, V2 an V3) were summarized from the total population, from a selection based on observed phenotypes, and based on predicted phenotypes with BRR (Figure 5). There is a large difference in the initial frequencies of the resistant alleles, and the difference is consistent across the populations. Allele 1C had a low frequency (0.19-0.28), 7C-1, 9D and 12D had medium frequency (0.41-0.5, 0.46-0.57 and

0.47-0.68) and 18D-1 had high frequency (0.77-0.89). In the training population (Figure 5a) there was generally increased frequency of all alleles in phenotypic selections with the highest increase in 7C-1, 9D and 12D. BRR selections had slightly higher frequency than phenotypic selection for all alleles.

There were some major differences between the validation populations (Figure 5b, c and d) in terms of frequency of resistant allele in response to phenotypic and genomic selection. Most alleles increased in frequency with phenotypic selection, except for 7C-1 in V2 and 9D in V3. BRR selection generally increased the frequency more than phenotypic selection except for 1C in V1 and 9D in V1 and V2. There was notably higher increase of 1C in V1 and 9D in V2 with Phenotypic selection than the other alleles.

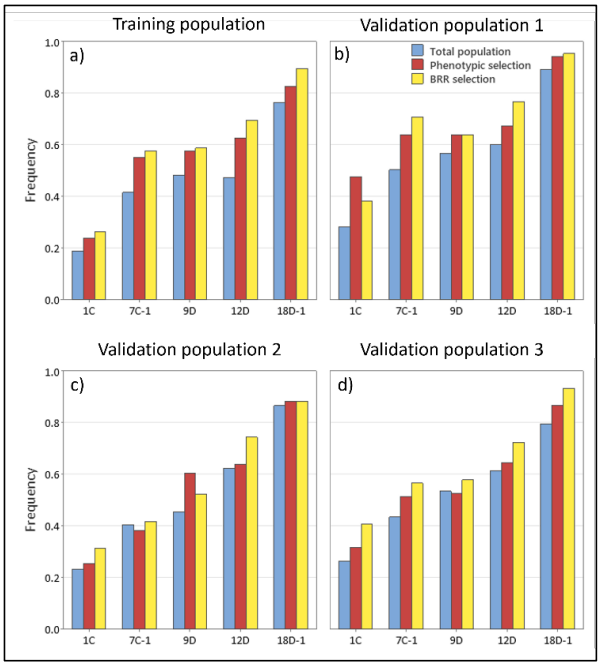



***Figure 5:*** *Frequency (y-axis) of resistant haplotype alleles of 1C, 7C-1, 9D, 12D and 18D-1 (x-axis) in the total population (Blue), Phenotypic selections (red) and BRR selections (yellow) in the training population (a), validation population V1 (b), V2 (c) and V3 (d). Selections were made on 1/3 of the lines with the lowest observed (Phenotypic selection) and predicted (BRR selections) DON values. The frequency of BRR selection for the training population was an average across the 30 replicates of the cross-validation.*

*3.3.2.2. Training population*

In order to gain more insight into how the different haplotype-based models affected the targeted loci, the frequency of the 15 resistant alleles were calculated in each selection and analysed to see if they differed significantly compared to BRR (Table 3). For the single haplotype models there were significant increases in their targeted resistant alleles. Other positive effects were a significant increase in 21D with M9D. The negative effects were significant reductions in 12D and 19A with M1C, in 20D with M9D, in 1C with M12D and in 7C-2 with M18D-1.

For the multiple haplotype models there were significant increases for all five resistant alleles included in M5HT. In addition, there were a significant increase in resistant allele of 21D. There were significant decreases in five other haplotypes (5C-1, 6C, 7C-2, 18D-2 and 19A). For M15HT that included all 15 haplotypes there were significant increases in 10 resistant alleles, and a significant decrease in two (5C-1 and 20D), while 6C, 18D-1 and 18D-2 were not significantly different from BRR.

| R-HT | M1C | M7C-1 | M9D | M12D | M18D-1 | M5HT | M15HT |
|---|---|---|---|---|---|---|---|
| 1C | *** | | | ** | | *** | *** |
| 2C | | | | | | | *** |
| 5C-1 | | | | | | *** | *** |
| 5C-2 | | | | | | | *** |
| 6C | | | | | | ** | |
| 7C-1 | | *** | | | | *** | *** |
| 7C-2 | | | | | ** | *** | *** |
| 9D | | | *** | | | *** | *** |
| 12D | ** | | | *** | | *** | * |
| 15A | | | | | | | *** |
| 18D-1 | | | | | *** | *** | |
| 18D-2 | | | | | | ** | |
| 19A | * | | | | | *** | *** |
| 20D | | | ** | | | | *** |
| 21D | | | *** | | | *** | *** |

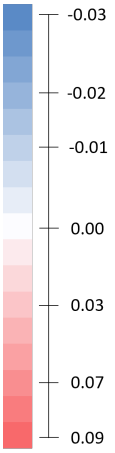



***Figure 6:*** *Difference between the fixed effect models and the base model BRR in frequencies of resistance haplotype alleles (R-HT) within populations selected based on the 1/3 lowest predicted DON values. The number are rounded up to the closest second decimal, and empty cells means that the difference was less than 0.005. p-values < 0.05 = *, 0.01 = ** and 0.001 = ***.*

### *3.3.2.3. Validation populations*

This section shows the differences between the fixed effect models and BRR for the validation populations in terms of frequencies of resistant alleles frequencies (Table 4). Haplotypes 2C, 5C-1 and 18D-2 were omitted from the results because the resistant allele of 2C did not show any notable difference from BRR with any model, 5C-1 had missing marker information in all populations, and the resistant allele of 18D-2 showed very large reduction for all models in V1 and V3, which obscured the results for the other alleles. The difference in allele frequencies between BRR and the fixed effect models were in general inconsistent across populations and haplotypes. The single haplotype models did increase the frequency of their target resistant allele in selection for all validation populations. In the M5HT the resistant alleles of 6C and 15A also increased consistently along with the five targeted resistant haplotype alleles, while the resistant allele of 21D decreased consistently across all population. With M15HT there were consistent increase in frequency of the seven resistant alleles of 1C, 5C-2, 6C, 7C-1, 7C-2, 9D and 15A. The alleles of 19A, 20D and 21D showed consistently reduced frequency or no difference with BRR. The alleles of 12D and 18D-1 increased in V1 and V2 but were reduced in V3. This coincided with the trend in V3 that 12D were reduced in all single haplotype models except M12D, and the same for 18D-1.

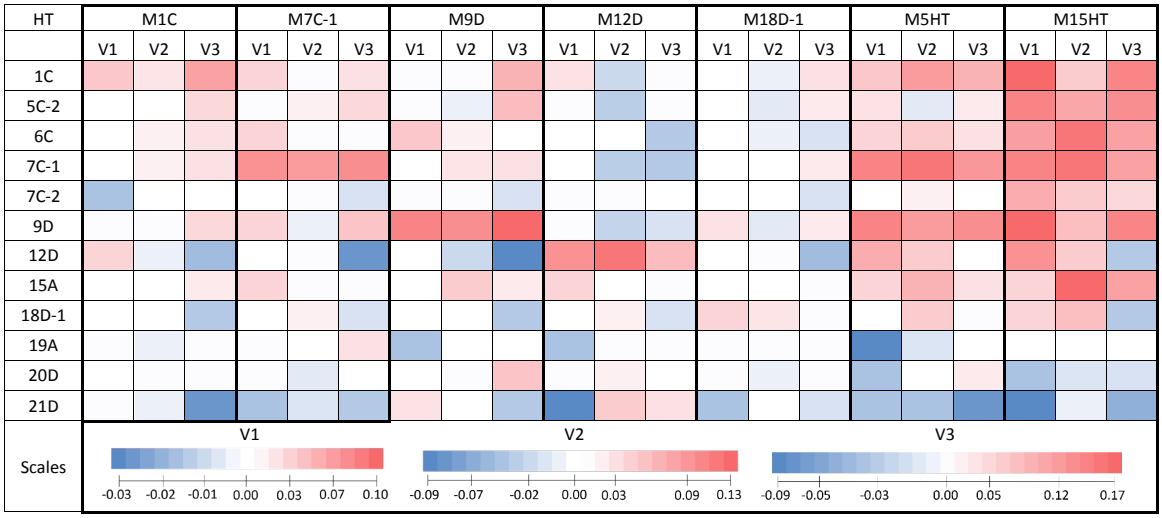



***Figure 7:*** *Differences in haplotype frequencies for all 12 resistant haplotype alleles (HT) between the fixed effect models (M1C, M7C-1, M9D, M12D, M18D-1, M5HT & M15HT) and BRR in selections based on the lowest 1/3 predicted DON values in Validation population 1(V1), 2(V2) and 3(V3). The colours represent the size of difference for each population shown in the scale at the bottom of the diagram.*

# 4. Discussion

The genomic selection models that included the fixed effects of haplotypes worked as intended in the training population and validation populations and increased and reduced the effects of resistant and susceptible alleles respectively in the models. This was demonstrated by the increased frequencies of targeted resistant alleles in selections of the lowest predicted logDON values (Table 3 and 4). The added effects resulted in increased PA for logDON and GP (Figure 1) as expected in the training population as this is the same population as was used for the GWAS. However, in the validation population PA varied depending on the haplotype, trait, and population. The following discussion will try to explain the potential reasons for the varying results.

## 4.1. Quality of data analysis

The trial statistics (Table 1) showed that the heritability of DON was relatively high for all population with lowest $H^2$ of 0.48 for V1, and highest of 0.81 for V3. Previous studies have shown large differences in heritability of DON between trials and years with the same material ranging from 0.28 to 0.81 (He *et al.,* 2013, Haikka *at al.,* 2020). For GP heritability was low in V1 (0.19) and V2 (0.18) with large environmental variance for V1 and large residual variance for V2. This indicates that there is a large difference in variance between the locations for V1, thus it would have been better to use phenotypes from single environments to estimate PA. For V2 the coefficient of variance (CV) for both DON and GP in V2 indicates some unknown error has occurred causing large differences between samples of the same varieties in the same experiments. The trial statistics of individual experiments indicate that the error has occurred in both locations as they both have relatively high coefficient of variance (CV) for both traits (Supplementary Table S1). As the error is present in both locations and traits it is most likely something that affects the general disease pressure like inoculation or environmental conditions. Additionally, the low heritability of GP in both locations indicate that there was something wrong with the post-harvest GP analysis as the analysis was carried out by Graminor for both locations. Or that some unknown factor unrelated to Fusarium has affected the GP as the statistics showed lower GP values in V2 than V1 even though the DON levels are almost identical.

## 4.2. Effect of added haplotypes on allele frequency

One of the main goals in genomic selection is to get accurate predictions, and this is achieved through a model that assigns effects of markers in relation to observed phenotypes of the training population. But each marker is not independent of the rest and can be inherited together due to physical linkage on the chromosome, family structure or due to selection in a breeding program (Slatkin, 2008). This makes the assigned effect of each marker connected to the effect of others. In this study we added

the effect of single QTL regions assuming they do not affect the other QTL regions. But as the selections showed (Table 3) that the effect of other QTL regions was relatively reduced, in some cases significantly compared to the basic BRR model, like with resistant allele of 12D with M1C, and 1C with M12D. Two potential reasons for this are that; (1) there is a linkage between the resistant allele of 1C and non-resistant alleles of 12D and vice versa, or (2) that the effect of 12D and 1C are overestimated in the BRR model, and that downregulating them with fixed effects of specific haplotypes contributes to improved accuracies. As PA is calculated as the correlation between the predicted and observed phenotypes it could be argued that to achieve 100% accuracy of the models, the allele frequency in selections should be the same with BRR as with phenotypic selection (Figure 5). But in selections made with BRR all resistant alleles have higher frequency than selection based on observed phenotypes. This indicates that the basic model overestimates the marker effects of these large effect QTL, increasing selection intensity for these regions, thus increasing genetic gain faster than phenotypic selection. A drawback of this is that alleles can get fixated faster in the breeding material, or that rare alleles are lost due to genetic drift, both of which reduce the genetic diversity of the breeding material making reducing potential genetic gain in the future with the same material (Wientjes *et al.,* 2023).

## 4.3. Types of resistance

Usually, the correlation between the DON and GP (Table 2) caused the reduced DON in selections to also show increased GP which was the case for the training population (Figure 3). But in some instances, it did not, like for M7C-1 in V2 (Figure 4c and d) and 18D-1 in V3 (Figure 4e and f). In the training population as well, there was a larger effect on GP with M1C than on DON, even though the selection was made for low DON (Figure 1 and 2). A last example is that M15HT showed consistently higher DON than M5HT in the validation populations, but also higher GP. One explanation to these inconsistencies is that there are genes that are specific to resistance to reduction in germination ability (Type 3), and resistance to DON accumulation (Type 5) which has been highlighted by several studies in wheat (Abate *et al.,* 2008, He *et al.,* 2019, Aviles *et al.,* 2020.). Similarly, studies in oat have identified QTL specific for FHB resistance (Types 1 and 2) and resistance to DON accumulation (He *et al.,* 2013, Bjørnstad *et al.,* 2017). It could be argued that a general FHB resistance would have similar effect on DON and GP which could be the case for M7C-1 as it has similar effect in the validation population for both DON and GP. For 9D however it has a clear effect on DON in the validation population, but no effect on GP indicating that it is Type 5 specific. In the GWAS study by Sørensen et al (2023) it was proposed that the QTL-region of 1C could be associated to the DON detoxification genes found by Khairullina *et al.* (2022). A possible explanation to the results for M1C in the training population is that the detoxification genes work within the grain itself, protecting the embryo from

damage. But if these genes only worked within the seeds, they would have limited effect on the total DON content as it has been shown that most of the fungus and DON is located in the husks (Brodal *et al.,* 2020), and our samples are not dehulled before DON analysis. An experiment on DON content of dehulled vs whole samples, or an analysis that shows where the detoxification genes are transcribed could potentially support this theory. Another important aspect to GP contra DON is that the reduction in GP can be affected by several other factors like heavy rainfall before harvest or other pathogens which could explain why there are differences between the results for DON and GP.

## 4.4. Resistance vs susceptibility genes

In some instances, the models can result in increased PA for the traits, but not significantly reduced DON or increased GP in selected material. This suggests that the increased PA comes from higher predicted DON and lower GP in the non-selected material. This is the case for both M9D in V1 and V2 (Figure 2 and 4) and M12D in the training population (Figures 1 and 2). In both cases there is a clear increase in frequency of the resistant allele in selections without the expected effect on the phenotypes. This suggests that the causal QTL is not a resistance gene (R-gene), but a susceptibility gene (S-gene). R-genes are often dominant and mainly function to protect the plants from specific pathogens. S-genes on the other hand are often related to other functions in the plants and makes them more compatible with the pathogen which increase infection (van Schie & Takken, 2014). The difference is subtle and not easy to determine. Therefore, it is not possible to determine that based on this study alone.

Evidence to support this theory is found in Supplementary Figure S2 which shows that there was a relatively high frequency of susceptible alleles of 9D in the training population (0.48) and lower in V1 (0.09), V2 (0.2) and V3 (0.18). In both V1 and V3 BRR already reduces the frequency of these alleles by half in selections, while there was no reduction in V2. This suggests that there was limited number of susceptible alleles remaining to select against with the M9D model. Resulting in increased predicted DON in non-selected lines and increased PA in total, but no notable reduction in DON in selected elite genotypes. The training population had more susceptible alleles in the total population, thus it showed good effect of the M9D model for both PA and DON in selections. For M12D the frequency of the susceptible alleles was already very low in the training population (0.12) which explains its low effect in the training population. In the validation population it is even lower with frequencies of 0, 0.03 and 0.05 in V1, V2 and V3 respectively which explains why there is almost no difference in PA for M12D in the validations as no increase in the presumed resistant allele improves the results. S-genes are often targets for gene editing as a simple deletion or knock-out of these genes in any way would result in more resistant genotype. Editing a non-resistant allele of a R-gene

into a resistant allele would however require knowledge of the protein it is supposed to create and precise alteration of the DNA code. One way to test this hypothesis is to do a deletion of resistant and susceptible alleles of each QTL. If the gene is a R-gene deletion of resistant allele would cause the genotype to lose resistance, while deletion of the susceptible allele would cause no difference in resistance. Opposite would be true if the gene was an S-gene. It is important however to determine the function of the S-gene to avoid unintended consequences as these often serves as other functions in the plant (van Schie & Takken, 2014).

## 4.5. Non-random linkage

One explanation to why the effect of the models differed between populations is that there is a non-random linkage between the different QTL-regions due to family structure. This linkage differ between population as each population consists of several populations of half-siblings from different crosses. Adding any assumed effects of major QTL to the model would therefor positively or negatively affect other QTL they are in LD with. This linkage could result in worse predictions as demonstrated in the validation populations (Figure 2). Several studies demonstrate the benefits of including fixed effects of major QTL in the prediction models (Li *et al.,* 2019, Kim *et al.,* 2022) while others highlight that increasing effects of major QTL could come at the expense of other minor QTL important for durable resistance (Poland & Rutkoski 2016, Herter *et al.,* 2019).

Another aspect of the importance of linkage is the LD between marker and causal QTL detected in the GWAS could change over time due to recombination of the chromosomes. This is demonstrated by the apparent co-segregation of the resistant alleles used as fixed effect and the allele on 21D in the training population (Table 3) as it significantly increased in selections when all five were included in the model and when only 9D was included. Contrary, in the validation population it almost always decreased even when its effect was added in M15HT. This indicates that the LD between the five validated markers and the markers on 21D changed over time. It is also possible that the presumed effects from the GWAS results are not true in the breeding populations due to recombination within the QTL regions, or between the regions and the causal QTL.

## 4.6. Avoidance mechanisms

As described earlier there are several mechanisms that provide reduced symptoms of FHB through passive avoidance of infection. In this study we investigated the effect of DTH and PH in selected elite material to see if the assumed resistance affects these traits. Days to heading were not affected by the added effect of haplotypes, but PH were in some cases, especially for the M15HT. These effects were not consistent across populations and could therefore be attributed to non-random linkage

between the markers and PH. However, it is possible that the observed increased PH in selections from the training population contributed to increased PA and reduced DON. While the consistent short plants observed in selection from the validation populations could have contributed to the consistently low PA and increased DON in all three populations.

Anther extrusion (AE) is another avoidance mechanism that has been shown to heavily impact FHB resistance in both wheat and oat (Skinnes *et al.,* 2010; Lu *et al.,* 2013; Tekle *et al.,* 2020). This trait has not been analysed in this study but could be a causal trait for the resistance observed. Previous research has shown that AE genes affect resistance to initial infection (Type 1), but not spread of disease (Type 2). A possible way to test the hypothesis of AE genes is to use segregating populations for the target QTL-regions and analyse them for AE and confirm the effects for FHB resistance. An additional study of the material with point and spray inoculation would confirm if the resistance shown is Type 1 resistance only.

## 4.7. Gene expression

Some of the resistant alleles do not increase in frequency after phenotypic selection for every population (Figure 5), indicating that they do not affect the resistance in that specific population or environment. And the presence of a resistance allele does not mean that the presumed resistance is expressed as other factors could affect it like environmental factors or transcription factors. The environmental interaction has been shown to impact gene expression in wheat (Munkvold *et al.,* 2013, Li *et al.,* 2020) and specifically for R-genes in *Arabidopsis thaliana* (MacQueen and Bergelson. 2016). But environmental factors could also heavily impact the *F. graminearum* fungus creating favourable conditions for infection (Hjelkrem *et al.,* 2022). Further, the infection pressure itself could be a factor that affects the gene expression as shown by a study on *F. equiseti* in wheat (Manghwar *et al.,* 2021).

Gene to gene interactions could also affect the expression and thereby influence the genomic predictions. Some genes work as transcription factors for other genes to function, and it has been proven to be a factor in FHB resistance in wheat (Kage *et al.,* 2017). The reduced frequency of the QTL-region on 21D in the validation population, contrary to the training population could indicate that it is a transcription factor important for some of the other genes to function thus limiting their effect in the prediction models. In this study each of the validation populations were tested under different environmental conditions and disease pressures (Table 1) which limits the possibility of determining what factors affects the expression. Further evaluation under controlled environments and transcription analysis could shed light on this question.

## 4.8. Evaluation and potential improvement of genomic prediction

The accuracy of genomic prediction is often limited by the heritability of the trait. Accuracy of genomic selection is therefore often expressed as Pearson correlation divided by the square root of heritability (Estaghvirou *et al.,* 2013). PA above heritability could indicate that the model also predicts part of the environmental effects or non-additive effect. In a normal cross-validation setup the environmental effects are confounded in the training population giving potentially higher PA than expected. This is why we included breeding population as well to validate the models. However, as the validation populations are tested in the same years as the training, there is expected to be some environmental effects captured by the model as well. A better approach to evaluate the models would be to use new breeding material tested in a new environment.

In our trials there is a relatively high heritability for DON (0.79) and GP (0.62) in the training population. The initial PA of 0.49 for DON and 0.52 for GP with the base model showed both that the models worked well with PA close to what other studies have found (Haikka *et al.,* 2020), and that PA had potential to be improved. The validation populations also showed moderately high PA for DON close to the heritability in V1 and V2. For GP however there were low PA for V1 and V2, which matched exactly the heritability of GP measured in these populations (Table 1). Higher PA were obtained in V3 (0.4) with the base model when higher heritability is achieved. This indicates that the model itself works and that high prediction ability depends on the accuracy of phenotyping.

The fixed effects worked as intended and gave extra weight to the targeted QTL-region and improved PA in the training population but did not work that well in the validation population. We propose three ways to improve PA based the current approach of added effects of significant markers. One is to differentiate the effects of the regions as the GWAS showed that each region contributes different effect on resistance. A second is to use a model that includes epistatic effects which could capture the potential gene interaction for the trait (Raffo *et al.,* 2022). Third is to use only the QTL-regions that have the low LD with other known QTL regions.

Other aspects of improving PA include the following: (1) using models that include G x E interaction like “RKHS” (Hu *et al.,* 2023) as G x E is shown to have a significant effect on DON and GP (Table 1), (2) use multi-trait models that include effects of DTH and PH to improve PA (Gaire *et al.,* 2022, Zhang *et al.,* 2022) as they have shown to correlate with both DON and GP (Table 2), and (3) to increase the number of markers used, as the approximately 3000 markers used in this study do not necessarily capture all the small effect QTL in the large oat genome. All improvements mentioned above could be used together in a multi-trait-multi-environment model (Gill *et al.,* 2021) with fixed effect of significant markers.

## 4.9. Recommendations for implementation

From this study it has become evident that the effect of associated markers from a GWAS are not necessarily transferable into better prediction models in breeding populations (Figure 2). This is contrary to most other studies that report that adding prior QTL information improves PA for Fusarium head blight (Rutkoski *et al.,* 2012, Zhang *et al.,* 2021, Alemu *et al.,* 2023). The approaches to incorporating QTL information differs and is not always directly comparable. There are four main points to make about incorporation of significant QTL as fixed effect for genomic prediction. One is to analyse the known resistance QTL for non-random linkage and co-segregation and use markers that have low LD to others. This should reduce the risk of lower effects of non-targeted resistance QTL as has been shown in this paper (Table 3 and 4; Herter *et al.,* 2019). A second point is to use markers that has a stable effect across environments, as this study has shown that some QTL have very low effect in certain environments (Figure 5). A third point is that markers that increase PA but do not increase breeding values in predicted resistant genotypes (Figure 4) could be more useful for discarding susceptible genotypes as the increased PA likely comes from reduced breeding values of the most susceptible genotypes. Finally, the markers could be specifically associated to DON or GP, even if they are identified as associated to both in a GWAS. Thus, only markers associated to the specific trait should be added to the model. Further work is needed on identifying effective markers for improved PA in new breeding populations.

## 4.10. Conclusion

This study has shown that the resistant QTL-regions found in the training population can be used to improve prediction accuracy. Selections of elite material based on the models showed lower DON and higher GP than with the base model. However, the models showed inconsistent effect in the validation populations. Increasing the effects of specific QTL-regions have shown to negatively affect others, and the effect differs between populations likely due to difference in linkage between the regions. This linkage can be positive or negative depending on which alleles that are linked. Normal genomic prediction models show increased selection intensity for the five validated resistance alleles compared to phenotypic selection. The complexity of FHB resistance in oat with multiple small effect QTL, several types of resistance, high GxE interaction, correlated traits and potential epistatic effects emphasise the need to develop more complex models in order to capture these effects. The results from this study will be used for implementation of genomic selection in the Norwegian breeding program, and the relatively high prediction ability will make it possible to do selection in early generation with the developed models.

# References


Aamot, H. U., Ward, T. J., Brodal, G., Vrålstad, T., Larsen, G. B., Klemsdal, S. S., Uhlig, S. & Hofgaard, I. S. (2015) Genetic and phenotypic diversity within the *Fusarium graminearum* species complex in Norway. *European Journal of Plant Pathology, 142*, 501–519. https://doi.org/10.1007/s10658-015-0629-4.

Abate, Z.A., Liu, S. and McKendry, A.L. (2008) Quantitative Trait Loci Associated with Deoxynivalenol Content and Kernel Quality in the Soft Red Winter Wheat 'Ernie'. *Crop Science, 48*: 1408-1418. https://doi.org/10.2135/cropsci2007.07.0411.

Abrahamsen, U., Brodal, G. & Waalen, W. (2016) Virkning av ulike forgrøder på neste års avling av hvete. NIBIO BOK 2 (1) Jord- og Plantekultur 2016, 106-111. URL: https://nibio.brage.unit.no/nibio-xmlui/handle/11250/2577326.

Alemu, A., Batista, L., Singh, P. K., Ceplitis, A., & Chawade, A. (2023) Haplotype-tagged SNPs improve genomic prediction accuracy for Fusarium head blight resistance and yield-related traits in wheat. *Theoretical and applied genetics*, *136*(4), 92. https://doi.org/10.1007/s00122-023-04352-8.

Alvarado, G., López, M., Vargas, M., Pacheco, A., Rodríguez, F., Burgueño, J. & Crossa, J. (2016) META-R (Multi Environment Trial Analysis with R for Windows.) Version 6.0. International Maize and Wheat Improvement Center.

Aviles, A. C., Harrison, S. A., Arceneaux, K. J., Brown-Guidera, G., Esten Mason, R., & Baisakh, N. (2020) Identification of QTLs for Resistance to *Fusarium* Head Blight Using a Doubled Haploid Population Derived from Southeastern United States Soft Red Winter Wheat Varieties AGS 2060 and AGS 2035. *Genes*, *11*(6), 699. https://doi.org/10.3390/genes11060699.

Becher, R., Hettwer, U., Karlovsky, P., Deising, H. B., & Wirsel, S. G. (2010) Adaptation of Fusarium graminearum to tebuconazole yielded descendants diverging for levels of fitness, fungicide resistance, virulence, and mycotoxin production. *Phytopathology*, *100*(5), 444–453. https://doi.org/10.1094/PHYTO-100-5-0444.

Bernhoft, A. Eriksen, G. S., Sundheim, L., Berntssen, M., Brantsæter, A. L., Brodal, G., Fæste, C. K., Hofgaard, I. S., Rafoss, T., Sivertsen, T. & Tronsmo, A. M. (2013) Risk Assessment of Mycotoxins. in Cereal Grain in Norway. *Opinion of the Scientific Steering Committee of the Norwegian Scientific Committee for Food Safety; VKM (Norwegian Scientific Committee for Food Safety):* Oslo, Norway, 2013. URL: https://vkm.no/english/riskassessments/allpublications/riskassessmentofmycotoxinsincerealgraininnorway.4.27ef9ca915e07938c3b2bad8.html.

Brodal, G., Aamot, H. U., Almvik, M., & Hofgaard, I. S. (2020) Removal of Small Kernels Reduces the Content of *Fusarium* Mycotoxins in Oat Grain. *Toxins*, *12*(5), 346. https://doi.org/10.3390/toxins12050346.

de Chaves, M.A., Reginatto, P., da Costa, B.S., de Paschoal, R. I., Teixeira, M. L. & Fuentefria, A. M. (2022) Fungicide Resistance in *Fusarium graminearum* Species Complex. *Current Microbiology* 79, 62. https://doi.org/10.1007/s00284-021-02759-4.

Estaghvirou, S. B. O., Ogutu, J. O., Schulz-Streeck, T., Knaak, C., Ouzunova, M., Gordillo, A., & Piepho, H. P. (2013) Evaluation of approaches for estimating the accuracy of genomic prediction in plant breeding. *BMC genomics*, *14*, 860. https://doi.org/10.1186/1471-2164-14-860.

Felleskjøpet Agri SA. (2016) Prognose for tilgang og forbruk av norsk korn for sesongen 2016/2017. URL: https://www.fk.no/Media/Files/prognose-for-tilgang-og-forbruk-av-norsk-korn-for-sesongen-2016-2017.

Felleskjøpet Agri SA. (2023), Plantevernkatalogen 2023. Page 26. https://ipaper.ipapercms.dk/TidRom/felleskjopet/plantevern/.

Gaire, R., de Arruda, M. P., Mohammadi, M., Brown-Guedira, G., Kolb, F. L., & Rutkoski, J. (2022) Multi-trait genomic selection can increase selection accuracy for deoxynivalenol accumulation resulting from fusarium head blight in wheat. *The plant genome*, *15*(1), e20188. https://doi.org/10.1002/tpg2.20188.

Gill, H. S., Halder, J., Zhang, J., Brar, N. K., Rai, T. S., Hall, C., Bernardo, A., St.Amand, P., Bai, G., Olson, E., Ali, S., Turnipseed B. & Sehgal, S. K. (2021) Multi-Trait Multi-Environment Genomic Prediction of Agronomic Traits in Advanced Breeding Lines of Winter Wheat. *Frontiers in Plant Science, 12*, https://doi.org/10.3389/fpls.2021.709545.

Haikka, H., Manninen, O., Hautsalo, J., Pietilä, L., Jalli, M. & Veteläinen M. (2020) Genome-wide Association Study and Genomic Prediction for *Fusarium graminearum* Resistance Traits in Nordic Oat (*Avena sativa* L.). *Agronomy, 10*(2), 174. https://doi.org/10.3390/agronomy10020174.

Hautsalo, J., Jalli, M., Manninen, O. (2018) Evaluation of resistance to *Fusarium graminearum* in oats. *Euphytica 214,* 139. https://doi.org/10.1007/s10681-018-2222-3.

Hautsalo, J., Jauhiainen, L., Hannukkala, A., Manninen, O., Veteläinen, M., Pietilä, L., Peltoniemi, K. & Jalli, M. (2020) Resistance to Fusarium head blight in oats based on analyses of multiple field and greenhouse studies. *European Journal of Plant Pathology, 158*, 15–33. https://doi.org/10.1007/s10658-020-02039-0.

He, X. Skinnes, H. Oliver, R. E., Jackson, E. W. & Bjørnstad, Å. (2013) Linkage mapping and identification of QTL affecting deoxynivalenol (DON) content (*Fusarium* resistance) in oats (*Avena sativa L.*). *Theoretical and Applied Genetics, 126*, 2655-2670. https://doi.org/10.1007/s00122-013-2163-0.

He, X., Dreisigacker, S., Singh, R. P., & Singh, P. K. (2019) Genetics for low correlation between Fusarium head blight disease and deoxynivalenol (DON) content in a bread wheat mapping population. *Theoretical and applied genetics, 132*(8), 2401–2411. https://doi.org/10.1007/s00122-019-03362-9.

Herter, C. P., Ebmeyer, E., Kollers, S., Korzun, V., Würschum, T. & Miedaner, T. (2019) Accuracy of within- and among-family genomic prediction for Fusarium head blight and Septoria tritici blotch in winter wheat. *Theoretical Applied Genetics* 132, 1121–1135. https://doi.org/10.1007/s00122-018-3264-6.

Hjelkrem, A-G. R., Aamot, H.U., Lillemo, M., Sørensen, E.S., Brodal, G., Russenes, A.L., Edwards, S.G. & Hofgaard, I.S. (2022) Weather Patterns Associated with DON Levels in Norwegian Spring Oat Grain: A Functional Data Approach. *Plants*, 11(1):73. https://doi.org/10.3390/plants11010073.

Hu, X., Carver, B.F., El-Kassaby, Y.A., Zhu, L. & Chen, C. (2023) Weighted kernels improve multi-environment genomic prediction. *Heredity* 130, 82–91. https://doi.org/10.1038/s41437-022-00582-6.

International Rules for Seed Testing (ISTA), Chapter 5, (2021) i–5-56 (64)

Kage, U., Yogendra, K. & Kushalappa, A. (2017) TaWRKY70 transcription factor in wheat *QTL-2DL* regulates downstream metabolite biosynthetic genes to resist *Fusarium graminearum* infection spread within spike. *Scientific Reports, 7*, 42596. https://doi.org/10.1038/srep42596.

Khairullina, A., Tsardakas Renhuldt, N., Wiesenberger, G., Bentzer, J., Collinge, D. B., Adam, G., & Bülow, L. (2022) Identification and Functional Characterisation of Two Oat UDP-Glucosyltransferases

Involved in Deoxynivalenol Detoxification. *Toxins*, *14*(7), 446. https://doi.org/10.3390/toxins14070446.

Kim, G. W., Hong, J-P., Lee, H-Y., Kwon, J-K., Kim, D-A. & Kang, B-C. (2022) Genomic selection with fixed-effect markers improves the prediction accuracy for Capsaicinoid contents in *Capsicum annuum*, *Horticulture Research, 9*, uhac204, https://doi.org/10.1093/hr/uhac204.

Li, D., Xu, Z., Gu, R., Wang, P., Lyle, D., Xu, J., Zhang, H., & Wang, G. (2019) Enhancing genomic selection by fitting large-effect SNPs as fixed effects and a genotype-by-environment effect using a maize BC1F3:4 population. *PloS one*, *14*(10), e0223898. https://doi.org/10.1371/journal.pone.0223898.

Li, G. Z., Li, H. X., Xu, M. J., Wang, P. F., Xiao, X. H., & Kang, G. Z. (2020) Functional characterization and regulatory mechanism of wheat CPK34 kinase in response to drought stress. *BMC genomics*, *21*(1), 577. https://doi.org/10.1186/s12864-020-06985-1.

Lu, Q.X., Lillemo, M., Skinnes, H., He, X.Y., Shi, J.R., Ji, F., Dong, Y., Bjørnstad, Å. (2013) Anther extrusion and plant height are associated with Type I resistance to Fusarium head blight resistance in bread wheat line 'Shanghai-3/Catbird'. Theoretical and Applied Genetics 126, 317-334

MacQueen, A., & Bergelson, J. (2016) Modulation of R-gene expression across environments. *Journal of experimental botany*, *67*(7), 2093–2105. https://doi.org/10.1093/jxb/erv530.

Manghwar, H., Hussain, A., Ali, Q., Saleem, M. H., Abualreesh, M. H., Alatawi, A., Ali, S., Munis, M. F. H. (2021) Disease Severity, Resistance Analysis, and Expression Profiling of Pathogenesis-Related Protein Genes after the Inoculation of *Fusarium equiseti* in Wheat. *Agronomy, 11*(11):2124. https://doi.org/10.3390/agronomy11112124.

Mesterházy, A. (1995) Types and components of resistance to *Fusarium* head blight of wheat. *Plant Breeding, 114*: 377-386. https://doi.org/10.1111/j.1439-0523.1995.tb00816.x.

Mesterházy, Á., Bartók, T., Mirocha, C.G. and Komoróczy, R. (1999) Nature of wheat resistance to *Fusarium* head blight and the role of deoxynivalenol for breeding. *Plant Breeding, 118*: 97-110. https://doi.org/10.1046/j.1439-0523.1999.118002097.x.

Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819.

Miller, J. D., Young, J. C., & Sampson, D. R. (1985) Deoxynivalenol and Fusarium Head Blight Resistance in Spring Cereals. Journal of Phytopathology, 113: 359-367. https://doi.org/10.1111/j.1439-0434.1985.tb04837.x.

Minitab, LLC. (2021). *Minitab*. Retrieved from https://www.minitab.com

Munkvold, J. D., Laudencia-Chingcuanco, D., & Sorrells, M. E. (2013) Systems genetics of environmental response in the mature wheat embryo. *Genetics*, *194*(1), 265–277. https://doi.org/10.1534/genetics.113.150052.

Poland, J., & Rutkoski, J. (2016) Advances and Challenges in Genomic Selection for Disease Resistance. *Annual review of phytopathology*, *54*, 79–98. https://doi.org/10.1146/annurev-phyto-080615-100056.

Polley, A., Plieske, J., Grafarend-Belau, E., Manninen, O., Vetelainen, M., Ceplitis, A., Vallenback, P., Alsheikh, M., Gnad, H. & Ganal, M. W. (2023) A new 7K SNP genotyping array for oat. *Oat newsletter, 60:4.* URL: https://oatnews.org/oatnews_pdfs/2020/oatnews_2023_Ganal.pdf.

Raffo, M. A., Sarup, P., Guo, X., Liu, H., Andersen, J. R., Orabi, J., Jahoor, A., & Jensen, J. (2022) Improvement of genomic prediction in advanced wheat breeding lines by including additive-by-additive epistasis. *Theoretical and applied genetics,135*(3), 965–978. https://doi.org/10.1007/s00122-021-04009-4.

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J-L. and Sorrells, M. (2012) Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. The Plant Genome, 5. https://doi.org/10.3835/plantgenome2012.02.0001.

van Schie, C. C., & Takken, F. L. (2014) Susceptibility genes 101: how to be a good host. *Annual review of phytopathology*, *52*, 551–581. https://doi.org/10.1146/annurev-phyto-102313-045854.

Skinnes, H., Semagn, K., Tarkegne, Y., Maroy, A.G., Bjornstad, A. (2010) The inheritance of anther extrusion in hexaploid wheat and its relationship to Fusarium head blight resistance and deoxynivalenol content. Plant Breeding 129, 149-155. https://doi.org/10.1111/j.1439-0523.2009.01731.x.

Slatkin M. (2008) Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nature reviews. Genetics*, *9*(6), 477–485. https://doi.org/10.1038/nrg2361.

Statistisk sentralbyrå, Statistics Norway (2023, February 21). *04607: Areal av korn- og oljevekster (F) 2001 - 2021*. ssb.no. Retrieved October 7, 2023, from https://www.ssb.no/statbank/table/04607/.

Sørensen, E. S., Jansen, C., Windju, S., Crossa, J., Sonesson, A. K., Lillemo, M., & Alsheikh, M. (2023) Evaluation of strategies to optimize training populations for genomic prediction in oat (*Avena sativa*). *Plant Breeding, 142*(1), 41–53. https://doi.org/10.1111/pbr.13061.

Tekle, S., Lillemo, M., Skinnes, H., Reitan, L., Buraas, T. & Bjørnstad, A. (2018) Screening of oat accessions for fusarium head blight resistance using spawn-inoculated field experiments. *Crop Science 58*, 143-151. https://doi.org/10.2135/cropsci2017.04.0264.

Tekle, S., Stråbø Schofer, S., He, X., Dong, Y., Bjørnstad, Å. (2020) Variation in anther extrusion and its impact on fusarium head blight and deoxynivalenol content in oat (*Avena sativa* L.). Agronomy 10, 354. https://doi.org/10.3390/agronomy10030354.

Tinker, N. A., Chao, S., Lazo, G.R., Oliver, R. E., Huang, Y-F., Poland, J. A., Jellen, E. N., Maughan, P. J., Kilian, A. & Jackson, E. W. (2014) A SNP Genotyping Array for Hexaploid Oat. *The Plant Genome, 7*: plantgenome2014.03.0010. https://doi.org/10.3835/plantgenome2014.03.0010.

Wientjes, Y. C. J., Bijma, P., van den Heuvel, J., Zwaan, B. J., Vitezica, Z. G., & Calus, M. P. L. (2023) The long-term effects of genomic selection: 2. Changes in allele frequencies of causal loci and new mutations. *Genetics*, *225*(1), iyad141. https://doi.org/10.1093/genetics/iyad141.

Zaïm, M., Kabbaj, H., Kehel, Z., Gorjanc, G., Filali-Maltouf, A., Belkadi, B., Nachit, M. M., & Bassi, F. M. (2020) Combining QTL Analysis and Genomic Predictions for Four Durum Wheat Populations Under Drought Conditions. *Frontiers in genetics*, *11*, 316. https://doi.org/10.3389/fgene.2020.00316.

Zhang, W., Boyle, K., Brule-Babel, A., Fedak, G., Gao, P., Djama, Z. R., Polley, B., Cuthbert, R., Randhawa, H., Graf, R., Jiang, F., Eudes, F., & Fobert, P. R. (2021) Evaluation of Genomic Prediction for Fusarium Head Blight Resistance with a Multi-Parental Population. *Biology*, *10*(8), 756. https://doi.org/10.3390/biology10080756.

Zhang, J., Gill, H. S., Brar, N. K., Halder, J., Ali, S., Liu, X., Bernardo, A., St. Amand, P., Bai, G., Gill, U. S., Turnipseed, B. & Sehgal, S. K. (2022) Genomic prediction of Fusarium head blight resistance in early stages using advanced breeding lines in hard winter wheat. *The Crop Journal*, 10(6), 1695-1704. https://doi.org/10.1016/j.cj.2022.03.010.

Supplementary material

# Evaluation of genomic prediction for Fusarium resistance in the Norwegian oat breeding program

Manuscript

Espen Sannes Sørensen[1,2], Susanne Windju[1], Constantin Jansen[1], Anna Kristina Sonesson[3], Morten Lillemo[2], Muath Alsheikh[1,2]

[1]Graminor AS, Ridabu, Norway.
[2]Department of Plant Science, Norwegian University of Life Sciences, Ås, Norway.
[3]Nofima AS, Ås, Norway.

Corresponding author:

Espen Sannes Sørensen
E-mail: espen.sorensen@nmbu.no.
Orcid-ID: 0000-0003-4646-8250

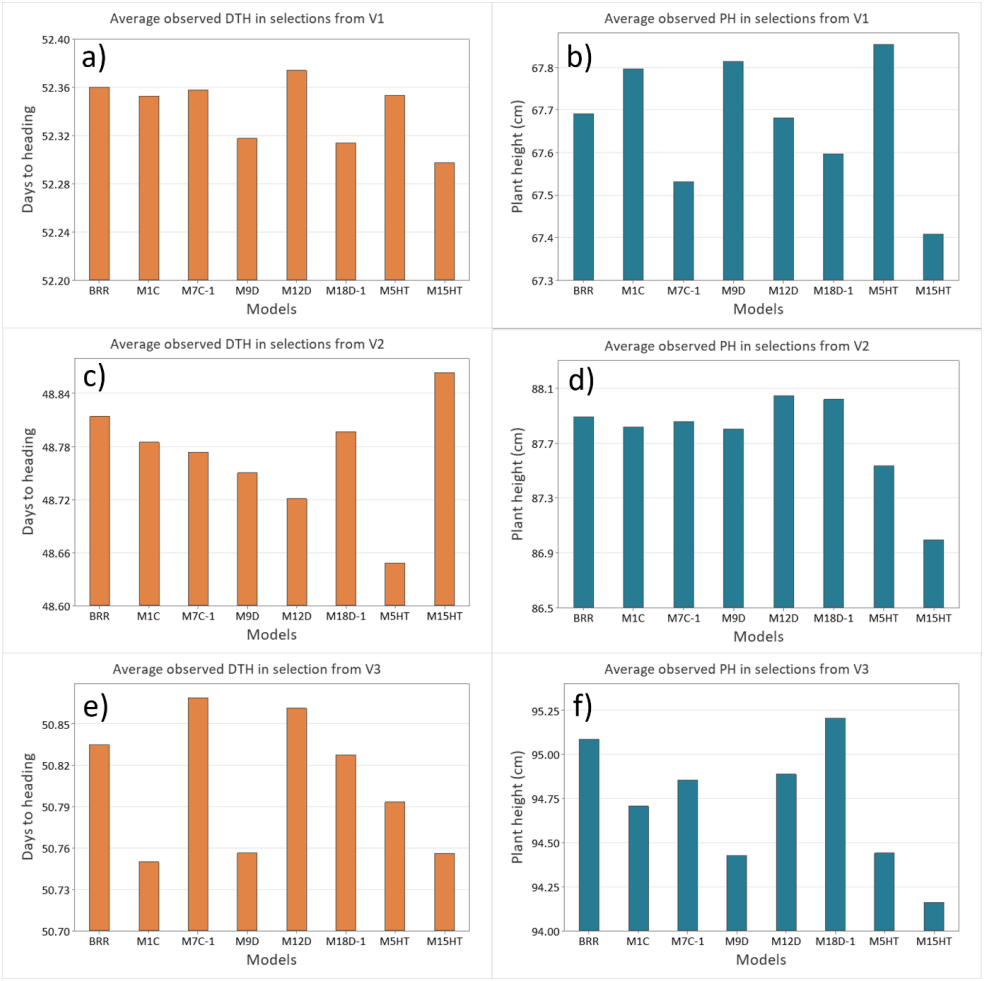



***Figure S8****: Average observed days to heading (DTH; a, c & e) and plant height (PH; b, d & f) in selecting 1/3 of the lines with lowest predicted logDON values in each model for validation population 1 (V1; a & b), 2 (V2; c & d) and 3 (V3; e & f)*

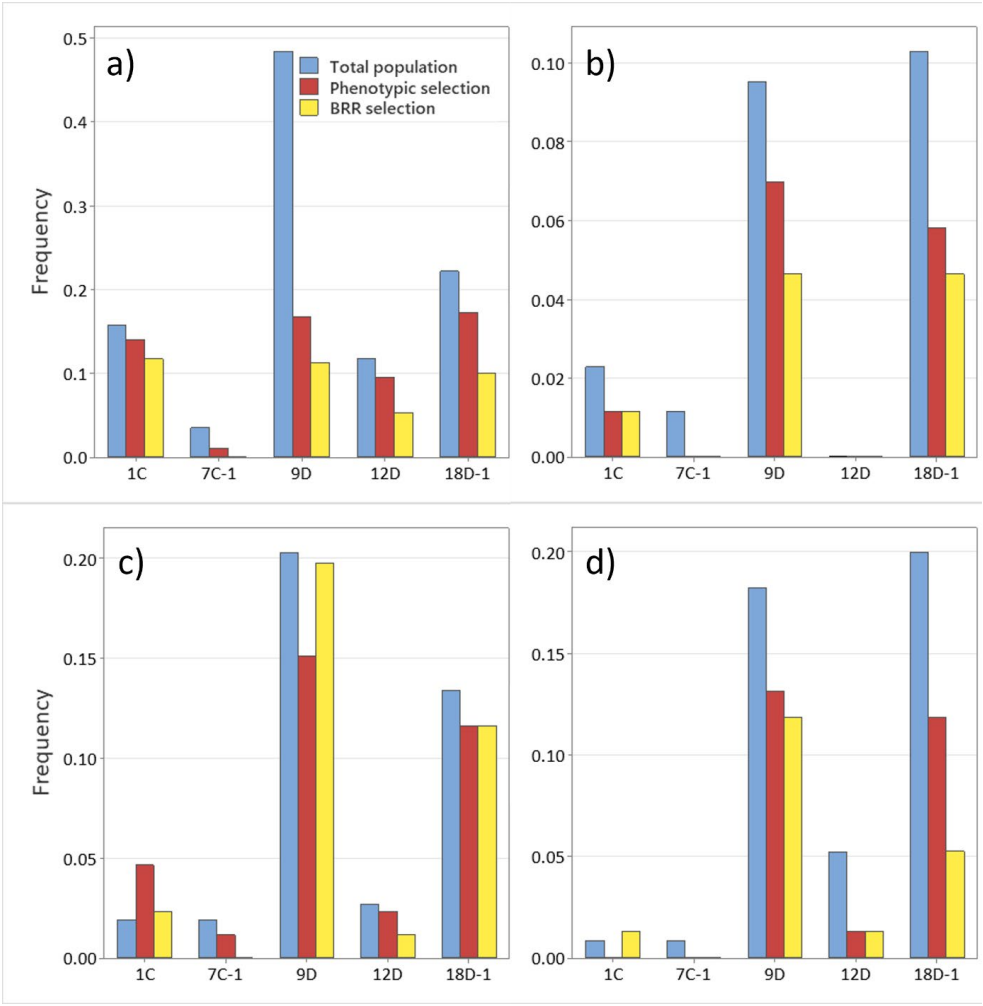



***Figure S9****: Frequency of susceptible haplotype alleles in the training population (a) validation population 1 (V1; b), 2 (V2; c), and 3 (V3; d) for the total population (blue), a phenotypic selection of the lines with the lowest 1/3 observed DON values (red) and genomic selection of the lines with lowest 1/3 prediceted DON values (yellow) for the five resistant alleles of 1C, 7C-1, 9D, 12D and 18D-1.*

***Table S6:*** Trial statistics of the individual experiments (Staur 2020, Vollebekk 2020, Staur 2021 and Vollebekk 2021) for the traits Days to heading (DTH) DON accumulation (DON) and Germination percentage (GP) and plant height (PH) for Validation population 1 (V1) and 2 (V2) including broad sense Heritability ($H^2$), Genotypic Variance, Residual Variance, Grand Mean, Least significant difference (LSD) and Coefficient of variance (CV), number of replicates (n) and Genotype significance.

| Statistic | DTH | DON | GP | PH |
|---|---|---|---|---|
| **V1 Staur 2020** | | | | |
| **Heritability** | 0.88267822 | 0.50827632 | 0.36145419 | 0.69059985 |
| **Genotype Variance** | 1.07389784 | 0.39752954 | 0.00053866 | 11.9985724 |
| **Residual Variance** | 0.28154133 | 0.735283 | 0.00186599 | 10.7724898 |
| **Grand Mean** | 43.4029165 | 1.97071173 | 0.87254165 | 78.2529086 |
| **LSD** | 1.17779078 | 1.83692011 | 0.0984082 | 7.6228242 |
| **CV** | 1.22250928 | 43.5115106 | 4.95072505 | 4.19428132 |
| **n Replicates** | 2 | 2 | 2 | 2 |
| **Genotype significance** | 0 | 3.0629E-07 | 0.00147838 | 1.4433E-13 |
| **V1 Vollebekk 2020** | | | | |
| **Heritability** | 0.83478904 | 0.50380608 | 0.51518223 | 0.50259917 |
| **Genotype Variance** | 0.73632351 | 0.32634027 | 0.00210664 | 7.34235168 |
| **Residual Variance** | 0.28926834 | 0.70580681 | 0.00398922 | 14.3515937 |
| **Grand Mean** | 61.200823 | 2.58288767 | 0.71393834 | 56.2798121 |
| **LSD** | 1.21185468 | 1.73257943 | 0.14525571 | 8.59680728 |
| **CV** | 0.87880636 | 32.5265052 | 8.84674193 | 6.73127645 |
| **n Replicates** | 2 | 2 | 2 | 2 |
| **Genotype significance** | 0 | 5.2971E-06 | 3.1774E-06 | 3.9163E-06 |
| **V2 Staur 2021** | | | | |
| **Heritability** | 0.52439771 | 0.42859541 | 0.13977773 | 0.26766417 |
| **Genotype Variance** | 1.73349733 | 0.50595833 | 8.21035216 | 6.78220311 |
| **Residual Variance** | 3.12443783 | 1.34565949 | 101.353436 | 37.489635 |
| **Grand Mean** | 41.8720787 | 2.8746064 | 70.4774912 | 92.4843173 |
| **LSD** | 3.45849862 | 2.26970415 | 19.8517857 | 11.9800124 |
| **CV** | 4.22144779 | 40.3542426 | 14.2846236 | 6.62045001 |
| **n Replicates** | 2 | 2 | 2 | 2 |
| **Genotype significance** | 1.4963E-10 | 8.4002E-07 | 0.10231555 | 0.00547889 |
| **V2 Vollebekk 2021** | | | | |
| **Heritability** | 0.88691044 | 0.43930933 | 0.21907203 | 0.6286097 |
| **Genotype Variance** | 1.60524013 | 0.25968082 | 12.3839436 | 10.5761668 |
| **Residual Variance** | 0.40911872 | 0.66565368 | 87.968802 | 12.5479745 |
| **Grand Mean** | 55.9726322 | 1.68684502 | 73.9916372 | 81.5083026 |
| **LSD** | 1.25148698 | 1.5963416 | 18.3535635 | 6.93088156 |
| **CV** | 1.14274398 | 48.3669816 | 12.6759846 | 4.3459524 |
| **n Replicates** | 2 | 2 | 2 | 2 |
| **Genotype significance** | 1.8738E-65 | 4.1535E-07 | 0.01597067 | 2.7267E-17 |

ISBN: 978-82-575-2126-4
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

119587 / MEDIEHUSET-ANDVORD.NO