Dissecting the genetic basis of wheat yellow rust resistance in the NIAB Elite MAGIC population

Laura Carmen Suzanne Bouvet

Lucy Cavendish College

National Institute of Agricultural Botany

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

November 2019

DECLARATION

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration except as declared in the preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed 60,000 words, excluding bibliography, figures and appendices.

Summary

Dissecting the genetic basis of wheat yellow rust resistance in the NIAB Elite MAGIC population

Laura Carmen Suzanne Bouvet

Yellow rust, caused by the biotrophic fungus *Puccinia striiformis* f. sp. *tritici* (*Pst*), poses a major challenge for wheat breeders and growers globally. The past two decades have seen the rise of *Pst* populations that are more genetically diverse, more aggressive and that have adapted to warmer temperatures. These features, likely further aided by increased international travel, have led to important epidemic outbreaks and have jeopardised wheat yellow rust resistance levels in the main wheat-producing regions globally. Widespread epidemics have been further facilitated by the deployment of genetically uniform material underpinned by major resistance over large areas. With such a rapidly changing *Pst* population landscape, disease resistance breeding strategies must adapt accordingly, and this starts with the continued characterisation of adequate yellow rust resistance loci and accompanying molecular and genomic tools. It is crucial that these loci are of direct relevance to breeding programmes for rapid varietal deployment. To this end, I used the NIAB Elite Multi-parent Advanced Generation Inter-Cross (MAGIC) population, a multi-founder population that captures 80 % of the genetic variation present in key representative varieties used in UK wheat breeding (1970-2010s), to identify and characterise genetic loci controlling yellow rust resistance in replicated multi-environmental field trials, in both leaves and ears. This approach has further opened up the avenue for dissecting disease resistance beyond the limited scope of single varieties. I found that nine Quantitative Trait Loci (QTLs) conferred resistance to yellow rust, with four consistently detected across environments and explaining nearly 50 % of the phenotypic variation, and the other five explaining 15-20 % with inconsistent detection across environments. There was a strong indication of additivity effects between the four strong-effect QTL. Furthermore, all founders but the most susceptible one contributed towards resistance, indirectly demonstrating that UK breeding germplasm has high resistance potential against yellow rust. In the second part of my thesis, I focus on

the physical interval of the eight most significant QTL previously identified, by examining gene annotations from the recently published IWGSC RefSeq v1.0 genome assembly. Five QTLs were characterised by NBS-LRR clusters. The presence of NBS-LRR-encoding genes with integrated domains revealed the potential for effector triggered immunity based on indirect recognition for a subset of those yellow rust resistance loci. The other three QTL were characterised by the absence of NBS-LRR-encoding genes in their physical interval, potentially indicating the role of non-race specific yellow rust resistance in the MAGIC population. Finally, I focus on glume infection, a phenotypic trait largely overlooked in QTL mapping studies, despite repeated reports of outbreaks. Despite high heritability (72 %), the five QTLs detected explained between 3 and 6 % of the phenotypic variation. Three QTLs co-located with QTLs for foliar resistance. The other two were associated with flowering time, suggesting that earlier ear emergence potentially leads to increased susceptibility to yellow rust in the glumes.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviation	Full name
ABC	ATP binding cassette
AIC	Akaike Information Criterion
AM	Association Mapping
AUDPC	Area Under the Disease Progress Curve
Avr	Avirulence
BAC	Bacterial Artificial Chromosome
BLUE	Best Linear Unbiased Estimator
CIM	Composite Interval Mapping
DH	Doubled Haploid
DS	Disease Severity
FDR	False Discovery Rate
FHP	Founder Haplotype Probabilities
GIF	Glume infection score
GM	Genetically Modified
GWAS	Genome Wide Association Scans
HA	Haplotype Analysis
HIF	Heterogeneous Inbred Families
IBD	Identity By Descent
IM	Interval Mapping
MAGIC	Multiparent Advanced Generation InterCross
MEL	MAGIC Elite
NIL	Near Isogenic Line(s)

NBS-LRR	Nucleotide-Binding Site Leucine-Rich Repeat
NLR	Nucleotide Binding Site Leucine-Rich Repeat
Pst	*Puccinia striiformis* f. sp. *tritici*
PVE	Phenotypic Variation Explained
QTLs	Quantitative Trait Loci
R	Resistance genes
REML	Restricted Maximum Likelihood
Rht	Reduced Height genes
RIL	Recombinant Inbred Line
START	Steroidogenic Acute Regulatory protein-related lipid Transfer domain
SMA	Single Marker Analysis
SNP	Single Nucleotide Polymorphism
SSD	Single Seed Descent
TPK	Tandem Kinase-Pseudokinase (TPK)
UK	United Kingdom
YR	Yellow Rust
WAGTAIL	Wheat Association Genetics for Trait Advancement and Improvement of Lineages

LIST OF APPENDICES

1 GENERAL INTRODUCTION

Wheat is one of the most important staple crops to the human diet and its production faces many threats. Yellow rust, caused by the fungal pathogen *Puccinia striiformis* f. sp. *tritici* (*Pst*), is one of them. Epidemics have been recurrent over the past 60 years and have caused significant yield losses. The past two decades, have seen the global emergence of *Pst* populations that have adapted to warmer temperatures, are more aggressive and have become more genetically diverse. Consequently, the disease resistance breeding strategies that have historically been used to control yellow rust, e.g. deploying major resistance genes singly, are no longer effective. Breeding strategies are starting to change, however. They are requiring additional sources of resistance which, together with adequate genomic and molecular tools, are aiding towards the development of resilient wheat varieties. My thesis sets out to address these issues.

In this introduction, I set the scene for the work undertaken herein, by pointing to key advances in wheat rust disease resistance and pathology up until 2018. I initially provide an overview of the importance of wheat as a staple crop and how its production is under threat. I continue by giving a similar introduction to the pathogen *Pst,* specifically focusing on its impact on wheat production, its evolution and adaptation as a population. I then move on to the genetic control of yellow rust resistance, outlining the two major types and their deployment. I finish with describing the discovery of yellow rust resistance loci and outlining the aims of my thesis.

1.1 Wheat: the rise of a food security staple

Since its domestication 8,000 to 10,000 years ago in the Fertile Crescent, wheat has accompanied and shaped human civilisation through centuries, across continents and across cultures. With 221.3 million hectares of land dedicated to its cultivation in 2016, wheat is the most widely grown crop in the world and the most important agricultural commodity traded globally (FAO, 2017). It is also the biggest contributor of calories in the human diet out of all cultivated crops. Wheat's popularity for human food and animal feed stems from its nutritional and processing properties, coupled with its adaptability to a wide range of climatic conditions.

1.1.1 Nutritious and versatile to cook with

Wheat grains are a rich source of starch, proteins and minerals and as a processed product, are one of the biggest contributors of calories and proteins in our diets (Chaves et al., 2013; Shiferaw et al., 2013). The gluten proteins present in wheat grains are essential for wheat's versatile uses as a cooking ingredient. These storage proteins are insoluble in water and form strong covalent and non-covalent bonds, providing the resulting doughs with the necessary structural properties to create a wide range of baked products (Day, 2013; Shewry, 2009). Wheat grains are also important natural sources of zinc, iron and selenium, micronutrients that cannot be synthesised by our bodies and that must be ingested instead (Shewry & Hey, 2015).

From a dietary perspective, wheat makes up 20 % of human daily caloric needs globally, with little difference between developing and developed countries (based on 2005-2009 average, Shiferaw et al., 2013). The importance of wheat as a cereal crop is emphasised when you consider its contribution amongst cereals only: as a source of protein, wheat outcompetes its cereal cousins, even in countries where wheat does not form part of the widely grown crops but is imported instead. Where it is widely grown, wheat accounts for over 70 % of cereal protein and protein intake, with Central Asia being the most dependant region on wheat as a source of protein (91 %).

These nutritional properties would be of little use if they could not be extracted efficiently and utilised in a versatile manner. The appealing processing properties of wheat flour have enabled this crop to establish itself in many different cultures and societies.

1.1.2 Adaptability to wide-ranging agricultural environments

Wheat's success as a staple crop is also due to its wide-ranging adaptability to different climates, made possible by centuries of human selection, starting with Neolithic farmers, and more recently, via the more targeted breeding approaches first adopted in the 19th Century. Very few crops can match wheat's adaptability to a different range of climatic and agronomic conditions.

The selection and cultivation of cereals made it possible for our society to shift from hunter gatherers to a more sedentary lifestyle. Initially, this was largely due to the selection for two important agronomic traits: non-shattering rachis, resulting in wheat ears that remain intact at maturity, and free-threshing, making the milling process less laborious. These enabled the development of a more controlled way of cultivating crops. However, it is the control and fine-tuning of flowering time that enabled wheat to be cultivated far from its centre of domestication. Flowering time is controlled by three different pathways: photoperiod, vernalisation and earliness per se (Cockram et al., 2007).

In the 1950s and 1960s, the Green Revolution further facilitated the up-scale of wheat cultivation and significantly contributed to yield improvements through the development of shorter and more input-receptive varieties. This was achieved by fine-tuning height, another important agronomic trait, controlled by the *Reduced Height* (*Rht*) genes (Peng et al., 1999). The tremendous advances from the Green Revolution enabled a doubling of wheat yields and alleviated millions of people from hunger.

1.2 Challenges facing wheat production

With demand for wheat predicted to increase to 324 kg/year (per capita) by 2050 (Alexandratos & Bruinsma, 2012), mainly driven by dietary changes in sub-Saharan Africa and Asia, its position as a staple food crop is set to remain but will require a

matching supply. Demand and supply effectively sit on either side of an equilibrium that can be regarded as a highly dynamic equation, shaped by a number of continuously changing forces. Providing a comprehensive overview of all factors affecting this equilibrium falls outside the remit of this work. This section will take a closer look at the supply side of this equation, specifically focusing on historical trends in wheat production and which factors have shaped it over the past four to five decades, followed by the contribution of pests and diseases to challenges faced by wheat production.

1.2.1   Wheat production trends

Historically, increased wheat productivity was achieved by allocating more land to cultivation. However, area expansion for arable crops has only increased by 9 % approximately over the past 60 years (Pretty, 2008). The expansion that has occurred has mainly been in the tropics, often resulting in the loss of tropical forest. Nevertheless, despite the relatively small increase in arable cropped area, wheat production has increased from 222 Mt in 1961 to 711 Mt in 2013 (FAO, 2017). When looking at wheat yields however, we have been seeing a plateauing trend in important wheat producing regions (Brisson et al., 2010; Calderini & Slafer, 1998; Ray et al., 2012).

As our understanding of wheat breeding developed and technological progress was made in the field of agronomy, the early Green Revolution of the 1950s and 1960s brought about a dramatic increase in wheat yields and thereby production, alleviating millions of individuals out of food poverty (Evenson & Gollin, 2003), largely thanks to the development and international deployment of shorter wheat varieties bred for ease of mechanical harvesting and for input responsiveness. Increased usage of crop protection products (15-20-fold increase worldwide over the period 1970-2010) was a significant contributor to increased productivity (Chakraborty & Newton, 2011). While the benefits of the Green Revolution are undisputed, they have relied on increased input use such as fertilisers, water and crop protection products, practices that have had a detrimental impact on the environment (Evenson & Gollin, 2003). The sustainability of such practices and

their negative impacts on the environment, in the context of the Green Revolution, have previously come under scrutiny (Pingali, 2012; Pingali & Rosegrant, 1994).

While wheat productivity in the past two decades has continued to increase, it is faced with several climatic, economical and societal challenges. Climate change will bring about higher temperatures, droughts and unpredictable rainfalls , directly impacting on global wheat production levels (Lobell, Schlenker, & Costa-Roberts, 2011). Such extreme weather patterns, in combination with other inter-connected factors, often lead to a spike in wheat prices and imports, leading to economic instability and political unrest. For example, recurrent droughts in 2006-2010, intensified the already unstable political climate that led to the Syrian uprising in 2010-2011 (De Châtel, 2014). In addition, as developing countries become wealthier, their diets change, typically resulting in consumption of a higher proportion of wheat-based products, resulting in increased demand.

In the future, these challenges will continue to destabilise wheat productivity. For each step-wise increase in temperature, several crop models have estimated yield reductions of 6-13% in wheat (Mondal et al., 2016). The total global cropping area is predicted to decrease by 8-20 % by 2050 (Nelleman et al., 2009).

1.2.2 Contribution of pests and diseases

In addition to the climatic, economic and societal factors outlined above, crop pests and diseases represent a recurrent and significant threat to wheat productivity: 10 to 16 % of wheat harvest globally is lost to pests and diseases (Oerke, 2006; Strange & Scott, 2005). Some of the most detrimental diseases of wheat are caused by biotrophic and hemibiotrophic fungi (Dean et al., 2012) and have been estimated to collectively cause annual yield losses of 12.4 % globally (Oerke et al., 1994).

The intensification of cultivation systems, with genetically uniform material now grown densely over larger and larger areas, means that more host material is available for infection and colonisation by pests and diseases, increasing their incidence and severity (Duveiller, Singh, & Nicol, 2007; Oerke, 2006). This is particularly the case for biotrophic fungi like rusts, which exhibit high evolutionary rates of adaptation (McDonald & Linde, 2002).

Rusts are among the most damaging pathogens to wheat production, causing recurrent epidemics and significant yield losses. These kinds of pathogens have a direct effect on grain-fill, particularly after anthesis. This is due to the damage to photosynthetic tissues leading to reduced light interception and radiation use efficiency, resulting in lower yields. Leaf (brown) rust caused by *Puccinia triticina*, stem (black) rust caused by *P. graminis* f. sp. *tritici* and stripe (yellow) rust caused by *P. striiformis* f. sp. *tritici* (*Pst*) have collectively been documented in all wheat growing regions worldwide, spanning more than 60 countries (Roelfs, Singh, & Saari, 1992; Singh et al., 2016; Wellings, 2007). Average global *Pst* mediated wheat yield losses have varied between 5 % and 50 %, but some localised severe epidemics have led to even higher losses (Wellings, 2011). In Australia, the incursion of an exotic race in 2003 and subsequent variants have resulted in A\$ 40-90 million in annual chemical control costs (Wellings, 2007).

1.3 Wheat yellow rust

Pst has a fairly complex lifecycle that goes through five different spore stages and requires two alternate hosts for completion of its life cycle. The two hosts broadly represent the two main stages in which the *Pst* life cycle can be divided into; the asexual and the sexual stages, which occur on wheat (the primary host) and *Berberis* species (the alternate host), respectively. The complete life cycle is outlined and described in Figure 1.1 (taken from Chen et al., 2014).

1.3.1 Pathogen lifecycle

Rust fungi have a heteroeceious macrocyclic life cycle, meaning that two different spore stages occur on two different host species. However, for *Pst* this could not be demonstrated for a long time, and there have been many historical attempts at characterising the alternate hosts for *Pst*. Although *Berberis* species were long speculated to support the lifecycle of the pathogen (Mains, 1933), it was not until 2010 that Jin et al. (2010) were able to demonstrate that Berberis species supported the development of *Pst* pycnia and aecia. Another host that supports the growth of the fungus during the sexual stage of its lifecycle is *Mahonia aquifolium*, commonly known as Oregon grape (Wang & Chen, 2013).

The relevance of the sexual stage of the lifecycle on the alternate host to *Pst* survival or epidemics is still largely unknown. However, current knowledge on *Pst* growing conditions through the different spore stages and recent evidence would suggest that it has little relevance. *Berberis* species infected with *Pst* are rarely observed in the wild, with only two accounts from China so far (Zhao et al., 2013, 2011). This could be explained by the difficulty in finding an environment that simultaneously accommodates teliospore (part of the asexual stage; enclosed in telia that form on wheat leaves towards the end of the infection season and produce basidiospores) and basidiospore (part of the sexual stage; form on barberry leaves and require dew for germination) germination, which both have short viability (Wang & Chen, 2015). The relevance of the sexual stage may lie in the ability of *Pst* infection on *Berberis* species to generate novel genetic variation (Jin et al., 2010).

The disease yellow rust occurs during the asexual stage of the *Pst* lifecycle and is caused by the multiple cycles of *Pst* urediniospores re-infecting the primary host. Urediniospores are airborne and disperse via the wind, with evidence of cross-continental dispersal (M S Hovmøller, Justesen, & Brown, 2002). Since this is the stage that is particularly detrimental to wheat production, the uredinial infection and colonisation processes have been studied extensively. Early phenotypic studies identified the environmental conditions conducive of infection and colonisation by uredinia: free moisture, a 9-13 °C temperature range for sporulation and low light levels. Urediniospores colonise wheat leaves by entering through the stomata. Disease symptoms can be observed 12 to 14 days after initial infection. During the initial penetration stage, urediniospores germinate on the leaf surface, with the resulting germ tubes entering the leaf tissue via the stomata. Growing hyphae develop into a dense network extending between and inside host mesophyll cells. Among this network, haustoria infection structures will form and specifically develop in host cell walls to extract nutrients from it (Szabo & Bushnell, 2001). Haustoria are also the site from which effectors are secreted and enable the pathogen to regulate host immunity (Garnica et al., 2013). The recent genome sequencing of *Pst* isolates and transcriptomics study enabled the mining of candidate effectors and identified
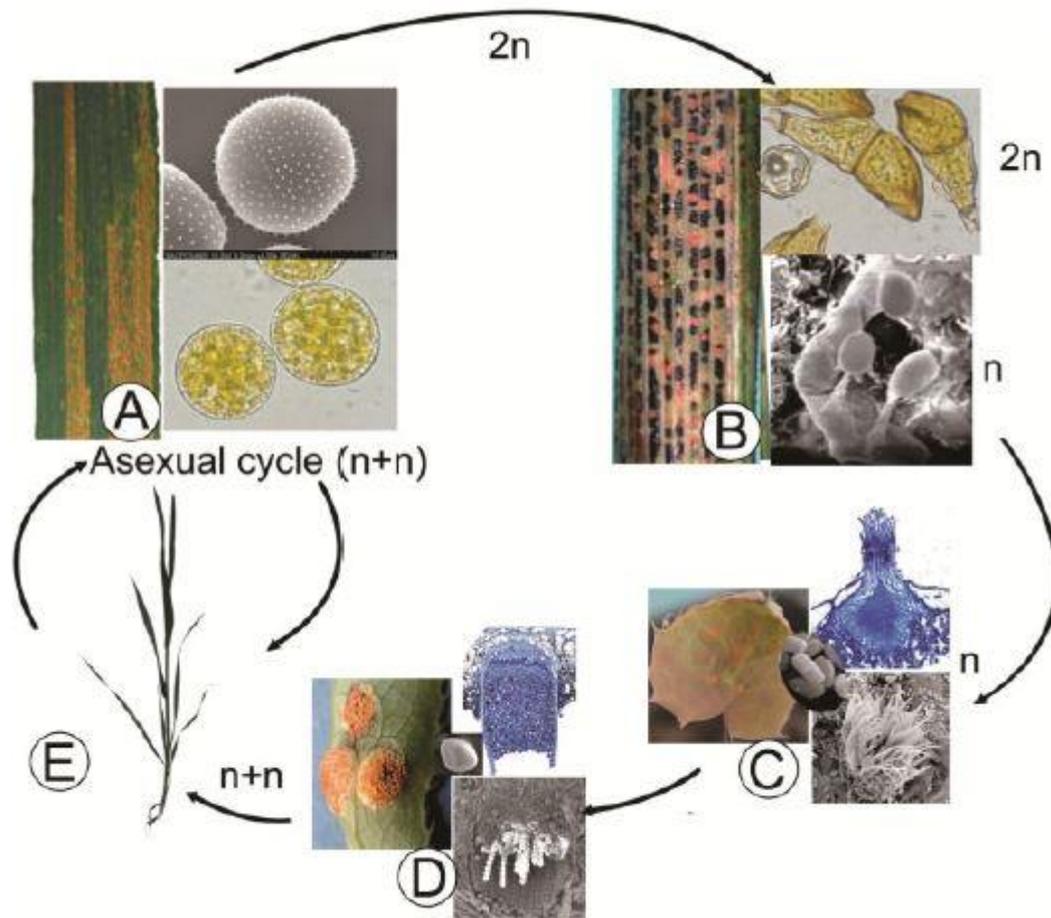
Figure 1.1 Lifecyle of *Puccinia striiformis* f. sp. *Tritici*

Image and figure legend taken from Chen et al. 2014. "(A) Uredinia on wheat leaf containing single-celled dikaryotic urediniospores (n + n) originating from aeciospores (n + n) or urediniospores. Top inset: echinulate surface of a urediniospore under a scanning electron microscope (SEM) (×4000). Bottom inset: broadly obovoid urediniospores (×1000). (B) Telia typically form beneath the leaf epidermis near the end of the growing season. Top inset: the two-celled, oblong-clavate teliospores (2n) (×1000). Bottom inset: the elliptoid basidiospores (n) from the germination of teliospores (×2500). (C) Pycnia produced by basidiospore infection on *Berberis chinensis* on upper leaf surfaces via inoculation with germinating teliospores of *P. striiformis*. Top inset: a magnified flask-shaped pycnia (×400). Middle inset: the oblong-shaped pycniospores (×4000). Bottom inset: magnified receptive hyphae (×900). (D) Cluster of sunflower-shaped aecia produced on the lower leaf surface of *Berberis shensiana*. Top inset: a campanulate aecium (×200). Middle inset: flat spherical-shaped aeciospores (×3300). Bottom inset: cluster of aeciospores (×250). (E) A wheat seedling that can be infected by aeciospores produced on barberry plants and can produce urediniospores. "

a subset for further validation (Cantu et al., 2013). This represents a significant first step towards effector-based breeding, an approach that has successfully been implemented for other patho-systems (Vleeshouwers & Oliver, 2014).

In addition to the micro-scale that is the uredinial infection process, *Pst* has extensively been studied at the population level. Global dispersal patterns are discussed further in section 1.4.

1.3.2 Disease impact on wheat production and control measures

Yellow rust causes yield losses and reduced grain quality. All wheat-producing regions have at some point suffered from epidemics that have led to yield losses, ranging in the extent of severity, magnitude and occurrence. Wellings' (2011) comprehensive review of the global status of yellow rust epidemics and consequential losses revealed that most of these regions have documented recurrent crop losses of 5 % to 10 %, with higher and much more severe losses of up to 25 % occurring less frequently. Beddow et al. (2015) provide a further perspective on these losses, highlighting an increase in (i) the number of countries being significantly hit by them and (ii) the extent of these losses, following the global spread of aggressive *Pst* races since 2000. The same study goes on to estimate the financial implications of such a change, estimating that a global average of US\$ 979 million was lost annually post 2000s, compared to US\$ 158 million pre 2000s. Epidemics can also occur following severe infection of heads and result in significant yield losses. Cultivars resistant in leaves have been reported to show significant infection levels in heads, resulting in yield and grain weight losses of 20 % and 11 %, respectively (Cromey, 1989b; Purdy & Allan, 1965). These concerns have been reported more recently by Wellings (2003, 2009).

To combat yellow rust epidemics and reduce yield losses as much as possible, wheat growers have two possible options at their disposal. They can (i) protect their crop with agro-chemicals that limit initial infection and progression of pathogen colonisation or/and (ii) grow genetically resistant varieties. Systemic fungicides became commercially available in the 1980s and have since then formed an important part of integrated control measures against yellow rust (Chen, 2005). Several fungicides with different modes of action are available to growers and these

have been extensively reviewed by Chen and Khang (2017). Timely application is a key aspect of an effective fungicide programme. Such an approach has, for example, prevented significant financial losses in periods of severe epidemics in the USA (Line, 2002).

While fungicide control provides an essential tool in combatting sudden yellow rust epidemics and in situations where growing resistant varieties is a limited option, they remain costly. In Australia for example, Murray and Brennan (2009) estimated that breeding remains the primary control measure for controlling YR (50 %), but that this is shortly followed by the use of chemical control (41 %), which is estimated to cost A$ 359 million per annum. Furthermore, the regular exposure of *Pst* to fungicides in the long term increases the risk that *Pst* populations could potentially develop resistance, as Oliver's proposed reassessment of rust fungi risk classification argues (2014).

Another way to control yellow rust population is to grow resistant varieties, which will be discussed in Sections 1.5 and 1.6 in more detail. Cultivation measures can also be taken to mitigate the spread of yellow rust epidemics. Varietal mixtures create a wide resistance spectrum by growing two or more varieties in a single field (M. R. Finckh et al., 2000; Mundt, 2002). This is an approach that has been frequently used in the US Pacific northwest (Chen, 2005). Such mixtures used to be recommended in the UK, via the UK Cereal Pathogen Virulence Survey (UKCPVS) Diversification Scheme (Priestley, Bayles, & Crofts, 1982). The scheme grouped wheat varieties according to how they responded to different *Pst* races. However, the genetic complexity that has arisen in the prevailing *Pst* lineages post-'Warrior' race incursion has made the grouping of varieties difficult, resulting in the halt of the Diversification Scheme in 2014. The practical aspect of deploying such mixtures is discussed further in section 1.6.

Chemical control and varietal resistance both have their advantages and should be considered as part of an integrated management approach towards durable control of stripe rust, something that the European Union is implementing as part of its Common Agricultural Policy under the European Directive 2009/128/EC. Ultimately, effective yellow rust control is a careful balancing act between both options. Control strategies favouring low fungicide input in combination with

varietal resistance have been shown to cut down on fungicide costs and result in similar yield levels to those observed under higher input (Loyce et al., 2008).

1.4 Status of *Pst* populations globally

Continuous monitoring of virulence changes in *Pst* populations across years and principle wheat producing regions has revealed major changes in pathogen movement and adaptation. These studies have traditionally been based on pathogenicity surveys, which utilise a set of differential wheat lines that carry known resistance genes (*R* genes), either Near Isogenic Lines or cultivars, for the characterisation of pathotypes at the seedling stage (Wellings et al., 2009). More recently, molecular techniques have been used to infer *Pst* population structure and genetic diversity, confirming patterns of adaptation hypothesised in earlier pathotype studies. Key findings and events from the past three decades are summarised here, specifically focusing on the evolution and adaptation of *Pst* and patterns of spore dispersal.

1.4.1 Pathogen evolution and adaptation

Molecular studies and pathogenicity surveys from across the main wheat-producing regions globally (USA, Europe, Australia, China) collectively report that *Pst* populations are clonal in nature, with pathotypes exhibiting close-relatedness and low genetic variation predominantly underpinned by single step-wise mutations (Ali, Gladieux, et al., 2014; Enjalbert et al., 2005; Hovmøller et al., 2002; Steele et al., 2001).

The only exception to this pattern is the Himalayan (Nepal and Pakistan) and near Himalayan (China) regions, which exhibit high levels of genetic recombination, high ability for sexual reproduction and high genetic diversity (Ali, Leconte, et al., 2014; Duan et al., 2010; Mboup et al., 2009). These regions have recently been suggested as the putative centres of origin and of genetic diversity for *Pst* (Ali, Gladieux, et al., 2014).

Recent single step-wise mutations have been responsible for a number of severe and costly epidemics, resulting from the 'breakdown' of specific *Yr* resistance genes present in large acreages in the environment. Notable examples include the

breakdown of *Yr17* in Northern Europe (Bayles et al., 2000), *Yr9* in America (Chen, Penman, Wan, & Cheng, 2010) and the Middle East and Indian sub-continent (Singh et al., 2004), and *Yr27* in Ethiopia (Solh et al., 2012).

1.4.2 The rise of atypical and aggressive *Pst* races

While most *Pst* populations have been shown to reproduce asexually and exhibit low genetic diversity, the last two decades have seen the emergence of unusual virulence profiles and aggressive strains.

The most noteworthy event is by far the rise of two strains, PstS1 and PstS2, across the USA (Chen, 2005; Markell & Milus, 2008), Europe (Hovmøller & Justesen, 2007a) and Australia (Wellings, 2007) in the space of three years in the 2000s. A global study of pre and post 2000s *Pst* races combining detailed virulence pathotyping and DNA fingerprinting (Hovmøller et al., 2008) revealed that these two strains were genetically similar but genetically highly divergent from previous races in each of the geographic regions. Despite their relatedness, PstS1 was exclusive to North America (presence since 2000) and Australia (present since 2002) while PstS2 was exclusive to Europe, Western and Central Asia and the Red Sea area. In addition, the two strains exhibited short special-temporal occurrence, as evidenced by their cross-continental presence and incidence in a short period of time. The rapid spread of PstS1 and PstS2 could be explained by their increased aggressiveness (ability to yield more spores and for disease symptoms to occur more quickly) and high temperature adaptation, which was later demonstrated in the detailed study by Milus et al. (2009).

In addition to the rise in dominance of *Pst*S1 and *Pst*S2, other atypical occurrences of *Pst* races have been reported since 2000. Enjalbert et al. (2005) demonstrated high levels of genetic divergence between the *Pst* population in northern France and a single clone specific to the South. What is atypical is that this single pathotype has maintained for a long time in this region, despite the presence of gene flow between Northern and Southern *Pst* populations. This isolate was later found to be more closely related to the *Pst* Mediterranean population (Ali, Gladieux, et al., 2014). Similarly, instances of strong genetic divergence have also been revealed in North Western Europe (Hovmøller & Justesen, 2007b). Two

groups of highly divergent pathotypes from the 'old' North-Western European population exhibited three to four times higher levels of genetic diversity (Hovmøller & Justesen, 2007b).

1.4.3 Patterns of *Pst* dispersal

Pst urediniospores are windborne and therefore have the potential to disperse over long distances. Coupled with the obligate nature of the pathogen (requiring living tissue to survive), this has led to different scenarios of seasonal and geographic patterns of dispersal including local extinction and re-colonisation, and continent-island migration. More recently, cases of incursion of foreign races leading to the rapid spread of *Pst* in the recipient country have been reported, demonstrating the potential of cross-continental dispersal and the role of anthropogenic activities (globalised trade and increased air travel) in enhancing the effects of foreign incursions.

One pattern of dispersal is that observed between the different wheat growing regions in China (Brown & Hovmøller, 2002; Zeng & Luo, 2006). The northern Sichuan and the southern Gansu provinces, where *Pst* prevails all year round, act as a source of inoculum to the northern provinces, where wheat is predominantly grown as a winter crop. *Pst* populations can therefore re-establish at the beginning of each wheat cropping season in those regions, where *Pst* spores are usually not able to over-winter. A similar pattern of spore movement (i.e. according to prevailing winds and the seasonality of the cropping seasons) has been speculated in North America, with spores migrating from southern central states of USA and Mexico to northern central states of USA and Canada (Chen, 2005).

In addition to these inter-regional and intra-continental modes of dispersal, *Pst* populations have been demonstrated to migrate over much longer distances. The 'continent-island' model first described by Hedrick (1985), has been the predominant model of dispersion of *Pst* spores in North Western Europe. In this region, urediniospores' mode of transport is the wind. They can travel up to 1,700 km, migrating between UK, France, Germany and Denmark, with Denmark generally acting as a recipient country (Hovmøller et al., 2002). In 2011, three novel *Pst* races disrupted the European *Pst* landscape (Figure 1.2). Termed after the host

varieties they were first detected on, two races were virulent on wheat ('Warrior' and 'Kranich') and the other was virulent on triticale ('Triticale aggressive'). Both the 'Warrior' and 'Kranich' races were detected simultaneously across Europe and were found to induce disease is varieties that had exhibited durable adult plant resistance (A. Hubbard, Lewis, et al., 2015). Both races were also distinct from the typical European isolates in that they produced an unusually high number of teliospores. Hubbard et al. (2015) used an RNA-seq approach termed field pathogenomics to examine the genetic structure of UK *Pst* races from 2013, demonstrating the substantial shift in population structure between pre and post-2011 *Pst* races, from a historically clonal to more genetically diverse population. Altogether, this evidence points to an aerial-induced foreign incursion across North Western Europe.

The emergence of yellow rust in countries where it was previously absent is another instance of rapid inter-continental foreign incursion. Australia has been subject to several incursions. Of these, two have been notably detrimental to the wheat industry, because of their rapid spread: (1) the first occurrence of the pathogen in 1979 (Wellings, 2007). (2) The 2002 incursion in Western Australia (Wellings et al., 2003), now known to have originated from the Middle East-East Africa (Ali, Gladieux, et al., 2014), and attributed to a single *Pst* isolate (Wellings et al., 2003), demonstrating the damage that a single pathotype can do. Yellow rust reached South Africa much later on, in 1996, from the Mediterranean and Central Asian populations (Ali, Gladieux, et al., 2014; Boshoff, Pretorius, & van Niekerk, 2002). How it got to South Africa remains unknown but is speculated to have resulted from wind dispersal or human intervention (Ali, Gladieux, et al., 2014). In all three cases, human activity, most likely through accidental transport on clothing, has been either demonstrated or strongly speculated, highlighting the increasing role of globalised trade and international air travel as a means for rapid and long-distance dispersal of *Pst* urediniospores.
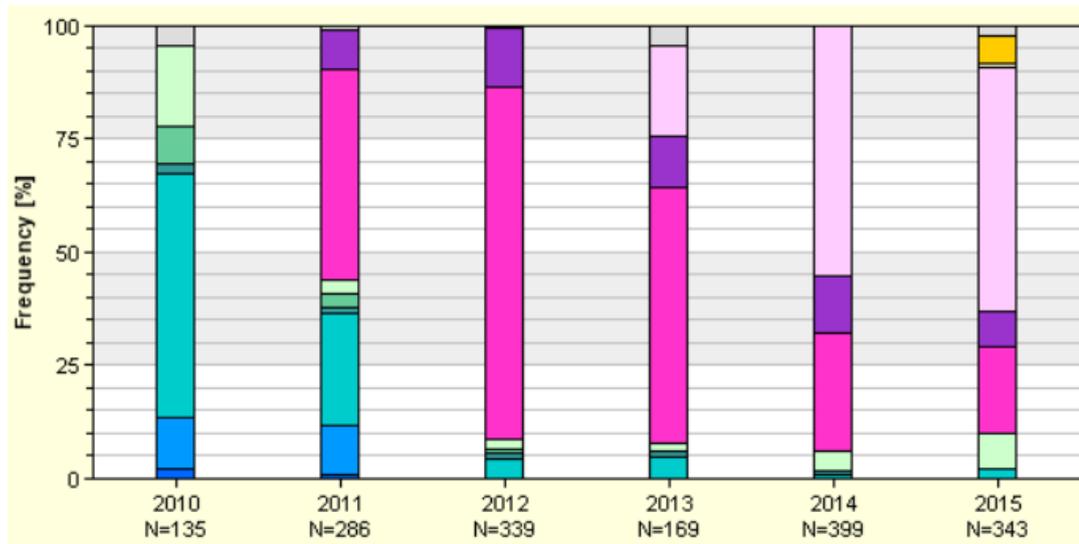
Figure 1.2 European race changes of *Pst* between 2010-2015.

Pst race changes expressed as relative % frequency based on overall number of races(N) from: Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Latvia, Lithuania, Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Ukraine, United Kingdom. Isolates in the blue shades prevailed in Europe pre-2011. Isolates in pink and purple shades have come to dominate the European *Pst* landscape after the Warrior race incursion. See Figure 8.1, Appendix A for detailed figure legend. © Global Rust Reference Centre 2019.

These well-documented historical accounts of dispersal are evidence of how *Pst* can disperse in different and sometimes unpredictable ways, constantly responding to their environmental cues. Movement of urediniospores only leads to epidemics provided the environmental conditions are suitable for the pathogen to establish itself and has enough host material to rely upon for survival. Some of these events have led to devastating epidemics, largely due to a combination of favourable climatic conditions for *Pst* development and the cultivation of susceptible material across a large acreage.

1.5   Mechanisms of yellow rust resistance

Sections 1.4 highlighted the potential for yellow rust to spread over long-distances through wind dispersal or accidentally aided by human intervention and the detrimental consequences this can have in wheat producing regions, given the right environmental conditions for pathogen development. One way in which this spread can be mitigated is through the cultivation of resistant cultivars. Here, I outline the two types of host resistance deployed in yellow rust resistance breeding and provide examples to illustrate their underlying genetics and molecular mechanisms. In addition to the host developmental stage criteria, other categories relating to aspects of pathogen development or underlying genetic mechanisms have been used to describe and refer to similar types of resistance. The use of these terminologies in the literature is also discussed.

1.5.1   All-stage resistance

All-stage resistance is initially expressed at the seedling stage and extends throughout the entire growth of the wheat plant and is characterised by a hypersensitive response. It is generally effective against some but not all races of *Pst* and is therefore also referred to as race-specific resistance. Race non-specificity of all-stage resistance does occur for some pathogens: for example the brown rust resistance gene *Lr21* has so far shown no specificity to particular races of *P. graminis* (McIntosh, 2009) and it has been speculated that this could be due to the lack of deployment (Lagudah, 2011). So far, non-specificity of all-stage resistance has not yet been reported for yellow rust resistance.

All-stage resistance is underpinned by the gene-for-gene model, first explored by Flor (1956) in the flax-rust pathosystem, whereby the product of a resistance (*R*) gene must be recognised by the protein encoded by its corresponding avirulent (*Avr*) gene in the pathogen, with infection elicited by a lack of an incompatible *R-Avr* interaction. This type of interaction translates to a 'clear' resistant phenotype that can be easily assessed by breeders, making it a popular selection criterion in breeding programmes historically and one of the reasons this type of resistance is also called 'major' resistance. A number of yellow rust *R* genes catalogued exhibit

this type of phenotype and have become ineffective against post 2000 *Pst* races (Chen et al., 2002; Chen et al., 2010; Hovmøller et al., 2011).

Historical accounts of these instances tell us that this type of resistance has very often turned out to be a short-term strategy for controlling yellow rust, with host susceptibility emerging in as little as a year (Boyd, 2005). Indeed, the deployment of single small combinations of all-stage *Yr* genes over large acreages inevitably exerts high selective pressure on the pathogen, forcing it to evolve and mutate until resistance is 'broken down' in the host (McDonald & Linde, 2002). This cycle is typically referred to as 'boom and bust', where boom corresponds to the rapid deployment of a single *Yr* gene over a large geographical area and bust refers to the adaptation of the *Pst* population, resulting in host susceptibility. These have been recurrently observed in France (de Vallavieille-Pope et al., 2011) and other parts of Europe (Bayles et al., 2000; Hovmøller, 2001).

1.5.1.1   NBS-LRR-encoding *R* genes

Of the 18 cloned cereal rust *R* genes (up to 2018, Table 1.1), 14 encode Nucleotide-Binding Site Leucine-Rich Repeat (NBS-LRRs) proteins (*Yr5, Yr7, Yr10, YrSP, Lr1, Lr10, Lr21, Lr22a, Sr13, Sr22, Sr33, Sr35, Sr45, Sr50*). These are most common class of proteins encoded by plant *R* genes (Jones & Dangl, 2006). NBS-LRRs are intracellular immune receptors that can detect pathogen effectors encoded by *Avr* genes and elicit an immune response upon effector recognition.

The vast majority of cloned *R* genes encoding NBS-LRR proteins harbour a Coiled-Coil (CC) domain at the N terminal (Cloutier et al., 2007; Feuillet et al., 2003; Huang et al., 2003; Liu et al., 2014; Mago et al., 2015; Periyannan et al., 2013; Saintenac et al., 2013; Steuernagel et al., 2016; Thind et al., 2017; Zhang et al., 2017).

Functional studies of a subset of *R* genes have provided some initial insight into the different molecular mechanisms of all-stage resistance in cereal rusts. In most studied cases of effector triggered immunity in plants, effector recognition and elicitation of the immune response is conditioned by individual NBS-LRR proteins (Bernoux, Ellis, & Dodds, 2011; Jones & Dangl, 2006). Another effector triggered immunity model is emerging, whereby two NBS-LRR proteins are required for effector recognition (Cesari et al., 2014; Eitas & Dangl, 2010) and has been reported

for all-stage resistance against *P. graminis* (Loutre et al., 2009). *Lr10*-mediated resistance requires *Lr10* and *RGA2*, two CC-NBS-LRR encoding genes acting as a pair to elicit host immunity. Another leaf rust all-stage resistance gene, *Lr21*, requires functional chaperons for effective resistance (Scofield et al., 2005).

Table 1.1 Summary of cloned cereal rust resistance genes

Cloned resistance gene	Protein product	Reference
Lr1	CC-NBS-LRR	Cloutier et al., 2007
Lr10	CC-NBS-LRR	Feuillet et al., 2003
Lr21	NBS-LRR	Huang et al., 2003
Lr22a	CC-NBS-LRR	Thind et al., 2017
Sr13	CC-NBS-LRR	Zhang et al., 2017
Sr22	CC-NBS-LRR	Steuernagel et al., 2016
Sr33	CC-NBS-LRR	Periyannan et al., 2013
Sr35	CC-NBS-LRR	Saintenac et al., 2013
Sr45	CC-NBS-LRR	Steuernagel et al., 2016
Sr50	CC-NBS-LRR	Mago et al., 2015
Yr5	BED-NBS-LRR	Marchal et al., 2018
Yr7	BED-NBS-LRR	Marchal et al., 2018
YrSP	BED-NBS-LRR	Marchal et al., 2018
Yr10	CC-NBS-LRR	Liu et al., 2014
Yr15	Kinase-pseudokinase	Klymiuk et al.,2018
Yr46/Lr67	Hexose transporter	Moore et al., 2015
Yr18/Lr34	ABC transporter	Krattinger et al., 2009
Yr36	Kinase-START	Fu et al., 2009

Studies exploring the protein architecture of plant immune receptors have identified a wide variety of integrated domains associated with NBS-LRRs and are shedding some light on their specific role in pathogen recognition (Bailey et al., 2018; Kroj et al., 2016; Sarris et al., 2016). These domains often act as a bait for the pathogen effector. *Yr5, Yr7* and *YrSP,* three recently cloned *Yr* genes, encode NBS-LRR proteins with such integrated domains (Marchal et al., 2018). Unlike the majority of the NBS-LRRs-encoding cereal rust resistance genes, *Yr5, Yr7* and *YrSP* contain a zinc-finger BED domain in place of a CC domain at the N terminal. The functionality of these domains is not yet fully known but Marchal et al. (2018) have demonstrated that, for *Yr7*-mediated resistance to be successful, it requires a functional BED domain, whereas this is not the case for *Yr5* and *YrSP*-mediated resistance.

A more unique gene structure to that encoding BED-NBS-LRR, not previously observed for yellow rust *R* genes, was recently reported. The cloning of *Yr15* (Klymiuk et al., 2018) revealed that it encodes a putative Tandem Kinase-Pseudokinase (TPK) termed wheat tandem kinase 1 (WTK1). This tandem kinase structure has been observed in *RPG1*, which confers resistance to barley stem rust (Brueggeman et al., 2002).

Taken together, these findings suggest that the *R-Avr* recognition process in all-stage resistance can be more complex than initially envisaged in the gene-for-gene model, and that these are underpinned by different molecular strategies.

1.5.2 Adult Plant Resistance

Adult plant resistance (APR) is characterised by slow rusting (e.g. a long period of latent infection, small lesion size, Guo et al., 2008) or partial resistance at the adult plant stage and has long been established as a durable source of resistance against yellow rust. Two notable examples are *Yr18/Lr34/Sr67/Pm38*, extensively deployed in spring wheat cultivars through the international breeding programme at CIMMYT (Singh, Huerta-Espino, & William, 2005) and *Yr16*, an APR gene commonly used in European varieties such as Cappelle Desprez, which went on to be the progenitor of several other important cultivars in the wheat pedigree (Agenbag et al., 2012). Plants harbouring APR are susceptible at the seedling stage and become more and more resistant as the plant develops. At present, it still is not known whether there is a specific developmental stage for the onset of APR. This type of resistance is primarily non-race specific, although examples of APR specificity to *Pst* races do exist such as *Yr12* and *Yr13* (Johnson, 1992; McIntosh, Wellings, & Park, 1995). The issue around non-race specificity of APR is an interesting one, because *Yr* loci initially considered as non-race specific can actually turn out to be race-specific; and this cannot be demonstrated until the *Yr* genes in question have been exposed to different pathotypes for long enough. This phenomenon has recently been observed in Europe following the spread of atypical *Pst* races (Sørensen et al., 2014). Another example is that of *Yr49*, a resistance gene that was initially found to be non-race specific (against all Australian *Pst* isolates), but when tested against Chinese races, was found to show race-specificity (Ellis et

al., 2014). These occurrences undermine the durability of APR and puts into question whether this pathotype criteria should be used to describe APR.

Certain instances of resistance at the adult plant stage are more effective at high temperature (usually 25-30 °C) and are termed High Temperature Adult Plant (HTAP) resistance (see Chen (2013) a recent review). This type of resistance can be more difficult and time-consuming to characterise, because (i) their effect is often masked by major genes and (ii) some resistance genes can show temperature sensitivity without necessarily being HTAP (Chen, 2013). *Yr36* falls into the latter category. Initially characterised as conferring HTAP resistance (Uauy et al., 2005), Fu et al. (2009) went on to demonstrate that resistance is effective over 25 °C at all growth stages. Bryant et al. (2014) later showed that fluctuations in temperature rather than actual values are an important factor in characterising disease resistance. This is of relevance to HTAP resistance, particularly in the context of a changing climate and rapidly adapting *Pst* populations.

Another important characteristic of APR genes is that they can confer resistance against multiple biotrophic pathogens, an attractive selection criterion in wheat breeding programmes. Some well-characterised examples include *Yr18/Lr34/Sr67/Pm38* (Lillemo et al., 2008; Spielmeyer et al., 2005), *Yr29/Lr46/Sr58/Pm39* (Lagudah, 2011) and *Yr46/Lr67/Sr55/Pm46* (Herrera-Foessel et al., 2014). Interestingly, some of these genes are also associated with other traits such as leaf tip necrosis (*Yr18/Lr34/Sr67/Pm38,* Singh, 1992; *Yr29/Lr46/Sr58/Pm39*, Rosewarne et al., 2006; *Yr46/Lr67/Sr55/Pm46*, Herrera-Foessel et al., 2014). *Yr18/Lr34/Sr67/Pm38* was recently shown to provide resistance against a fifth disease, spot blotch, caused by the hemibiotroph *Bipolaris sorokiniana* (Lillemo et al., 2013).

Unlike for all-stage resistance, a defined picture of the underlying APR genetics and molecular mechanism is not yet clear. At the onset of this thesis, only three wheat rust APR genes had been cloned (Table 1.1). The cloning of *Yr36* was a first step towards a better understanding of temperature sensitive APR (Fu et al., 2009). It encodes a kinase START lipid-binding domain protein that has been suggested to play a role in regulating reactive oxygen species levels during immunity (Gou et al., 2015). *Lr34/Yr18* encodes an ATP-binding cassette (ABC) transporter

(Krattinger et al., 2009). Ten years after these findings, Krattinger et al. (2019) identified Abscicic Acid (ABA) as a substrate for the ABC transporter. ABA redistribution in leaves was hypothesised as a mechanism for disease resistance conferred by *Lr34*. The third APR gene to be cloned more recently was *Yr46/Lr67*, which encodes a hexose transporter (Moore et al., 2015).

Map-based cloning of *Yr36* and *Lr34/Yr18* was possible partly because the phenotypic response in mutant lines was strong enough to be detected and for fine-mapping to be carried out. Stark differences in the type of protein and their potential roles in regulatory pathways highlight a likely more complex molecular mechanisms underpinning APR compared to all-stage resistance. With only three genes cloned, there is therefore a large gap to fill in the identification of APR genes/quantitative trait loci (QTLs) effective against yellow rust, in the understanding of their underlying molecular mechanisms, and in the characterisation of interactions with other resistance genes when deployed together.

1.5.3 Use of yellow rust resistance terminology

Additional categories have been used to describe different types of resistances (Chen, 2013). These were initially based on phenotypic response (infection type, race specificity, resistance levels), temperature sensitivity (high vs low temperature), durability and developmental stage (seedling and adult plant). As the genetic basis of yellow rust resistance has now become better understood, criteria have expanded to include the number of genes involved (monogenic vs polygenic) and the effect of those genes (major vs minor).

Such an array of terminologies has inevitably led to inconsistencies in the description of yellow rust resistance in the literature, as researchers use these terms either interchangeably or simultaneously. Depending on the researcher's viewpoint or the nature of the study, emphasis tends to be placed on one criterion over the other. Different authors have used different terminologies. This is particularly evident from recent reviews on cereal rust resistance. Lowe et al. (2011) opted to describe resistance to cereal rusts as major vs partial resistance but recognise and argue for the need to integrate additional factors such as 'gene

functionality' in classification efforts. Lagudah (2011) on the other hand uses the growth stage and race specificity criteria simultaneously in reviewing the molecular genetics of resistance. Another review on yellow rust quantitative trait loci (QTLs) described fungal disease resistance as being conditioned by major and minor resistance genes (Rosewarne et al., 2013). In some cases, different terminologies have been used in direct comparison. For example, Ellis et al. (2014) compared *R* genes with APR genes, two terms belonging to two different categories, namely the growth stage criteria (APR) and the plant-pathogen interaction criteria (*R* genes). All-stage resistance has been described as a more accurate term to use by Lin and Chen (2007) and Chen (2013). The term seedling resistance, however, continues to be used.

More importantly, the issue that comes with defining yellow rust resistance with such a broad range of criteria is the assumptions that are associated with each of them. For example, APR is generally thought to be non-race specific, more durable than all-stage resistance and conditioned by genes with minor or partial effect. Nevertheless, several yellow rust APR genes have been shown to exhibit race specificity, such as *Yr11, Yr12, Yr13* and *Yr14* (Johnson, 1992; McIntosh et al., 1995). Similarly, race-specificity of adult plant resistance has also been observed following the emergence of atypical *Pst* races post-2000 in Europe (Sørensen et al., 2014) and the USA (Sthapit et al., 2012).

Finally, the term 'durable resistance', first defined by Johnson (1988) as resistance that remains effective over time and space and under environmental conditions conducive of disease infection, has raised some important questions. Firstly, how do you define durability beyond the time aspect? Ellis et al. (2014) argue for the use of multiple pathogen resistance as a classification category for durability, so that the race-specificity of APR genes argument as described above can be removed. Secondly, when does resistance start to become durable? This question still remains unanswered.

1.6 Deployment of YR resistance genes

Growing cultivar mixtures is one way in which resistance genes can be deployed in the field (Mundt, 2002). These mixtures provide a physical means of constraining

virulent pathotypes and spatially limit their spread. In addition, the presence of several resistance genes effectively reduces the selection pressure exerted on pathogen evolution as pathogen populations would have to overcomes all resistance genes in question, thereby extending the usability of resistance genes. The practicalities of such an approach on-farm must be taken into consideration. Cultivars with different drilling dates or end-markets for example are not likely to be taken up by farmers as a mixture. Wider landscape-level strategies such as deployment of resistance genes that are relevant to *Pst* populations at the regional level have been proposed for cereal rusts (Finckh & Wolfe, 2006) but have picked up little traction, probably for similar reasons of management challenges on-farm.

Nevertheless, this concept of mixtures can be taken one step further by pyramiding or stacking resistance genes into a single cultivar (Mundt, 2014; Wulff & Moscou, 2014). This approach has been a successful and durable control measure against some wheat rusts. Two 'stacks', also termed 'complexes' are notable. Global stem rust levels remained largely under control for over 40 years prior to identification and spread of Ug99, thanks to the deployment of the APR gene *Sr2* in combination with several different major resistance genes: with *Sr7b*, *Sr9b* and *Sr17* in cultivar Hope (reviewed by Ellis et al. 2014), and other sources of resistance in CIMMYT germplasm (Singh et al., 2014). Similarly, the CIMMYT breeding programme has made extensive use of the '*Yr18* complex' (*Yr18* and at least two other resistance genes, (Singh et al., 2005), which has provided durable resistance against yellow rust. This first iteration of resistance gene pyramiding was developed using conventional breeding techniques i.e. crossing and phenotypic selection, which are slow and time-consuming to introgress novel combinations into cultivars. Improvements in molecular marker developments and selection methods such as the 'single backcross approach' advocated by Singh et al. (2005) have accelerated this process. Much more recently, advances in genetic modifications (GM) and breeding are set to significantly fast track and refine the development of resistance gene pyramiding (Dangl, Horvath, & Staskawicz, 2013; Ellis et al., 2014; Wulff & Moscou, 2014). GM cassettes that integrate *R* and APR genes are advantageous for a number of reasons: (i) they are inherited as a single genetic block, preventing the segregation of its components during selection and avoiding linkage drag, (ii) they

enable the extension of the resistance gene pool, by making the use of genes from normally host incompatible species a possibility.

The concept of wheat rust resistance gene stacking as a means of deployment of sources of durable YR resistance has already shown potential. Nevertheless, little is still understood about which combination strategies offer the most durability, resilience and efficacy (Mundt, 2014). GM cassettes do not come without their challenges either; only a small fraction of *Yr* genes have been cloned and GM acceptance by the wider society remains an important barrier (Wulff & Moscou, 2014). Ultimately, effective *Yr* gene deployment strategies must also consider the bigger picture: breeding for disease resistance in conjunction with yield and quality, while mitigating fitness cost (Summers & Brown, 2013), and deployment of YR resistance genes/alleles into elite genetic backgrounds.

1.7 Yellow rust resistance loci and gene discovery

Over 140 QTLs conferring yellow rust resistance have been published since the early 2000s, yet only a small proportion has been fed into breeder selection pipelines (Rosewarne et al., 2013). In addition, a significant proportion of *Yr* genes catalogued to date are not effective against the emerging genetically diverse *Pst* races globally. Thirdly, very few resistance genes and molecular modes of action have been characterised for yellow rust resistance. We have now reached a critical point in yellow rust resistance research and breeding whereby the risks of future incursions and pathogen adaptation must be mitigated by smarter resistance breeding strategies, aided by increased knowledge of the genetic basis of yellow rust resistance and more suitable genetic and molecular resources.

1.7.1 QTL mapping with experimental populations

Yellow rust resistance loci, together with their genetic mode of inheritance, have for the most part been discovered and studied via QTL mapping studies. These require phenotypic and genotypic information from an experimental population. With the development of molecular markers, high-density genetic maps, the availability of a reference genome for wheat, and powerful statistical tools to establish associations between phenotype and genotype, studying and mapping

quantitative traits at a greater resolution has become more amenable. In this section, I give a brief overview of the different types of populations that have been used in the characterisation of Yr genes and QTL conferring yellow rust resistance. I also introduce a novel type of experimental population that has emerged in recent years, termed multi-parental populations.

In the last two decades of rust research, most of the published *Yr* genes and yellow rust resistance QTL have been characterised in biparental populations, consisting for example of Recombinant Inbred Lines (RILs) or Double Haploids (DH). Some parents have been used more frequently than others, such as the French variety Camp Rémy (Boukhatem et al., 2002; Mallard et al., 2005) and the UK variety Claire (Chng, Cromey, & Shorter, 2011; Powell et al., 2013). Commercial varieties and elite breeding material exhibiting durable resistance have been extensively used in bi-parental crosses, because these types of population are relatively cheap and quick to develop (Lin & Chen, 2007; Mallard et al., 2005; Powell et al., 2013). Multi-environmental trials with larger populations have become more frequent in the last decade (Dolores Vazquez et al., 2012; Rosewarne, Singh, Huerta-Espino, & Rebetzke, 2008; Singh et al., 2014; Singh et al., 2012). Coupled with advancements in high throughput genotyping and increased marker density, the power of QTL mapping with bi-parental populations has improved over the years. However, with only two parents making up a cross and for example just a single round of recombination exploited in the generation of F_1-derived Doubled Haploid (DH) progeny, opportunities to simultaneously characterise a large number of QTLs (>6) and explore their interactions in bi-parental populations remain limited.

As an alternative to bi-parental populations, association panels have recently emerged as a powerful alternative to characterising yellow rust resistance QTL in Genome Wide Association Studies (GWAS). They rely on linkage disequilibrium and involves identification of marker-trait associations using large germplasm collections typically consisting of landraces, commercial varieties, or elite breeding material. Such panels capture more genetic diversity and mine the historical recombination present in each line, thereby avoiding the need to generate experimental populations. This has largely been possible thanks to advances in high throughout genotyping and the application of suitable statistical approaches

for taking into consideration population structure, which can lead to false positives if not accounted for properly, and QTL characterisation. A landmark study in the use of LD to detect wheat rust resistance QTL is that of Crossa et al. (2007), which highlighted the usefulness in mining historical phenotypic data from breeding programmes and exploiting LD for the detection of marker-trait associations. For yellow rust resistance specifically, there are two most recent examples of note. The first is a coordinated approach between academia and industry for the rapid deployment of yellow rust resistance with a panel of European (mostly UK) wheat varieties (Gardner and Cockram, unpublished). The second example lies in the application of GWAS to a core collection of spring wheats from the USDA-ARS National Small Grains Collection (Maccaferri et al., 2015).

In response to the limitations of bi-parental populations and associations panels, a novel type of mapping population has been emerging for investigating complex quantitative traits in crops: Multiparent Advanced Generation Inter Cross (MAGIC) populations (Cavanagh et al., 2008). Following the footsteps of the mouse Collaborative Cross for animal genetics (Churchill et al., 2004), wheat MAGIC populations have recently been constructed by the structured crossing of multiple founders (or parents) of breeding relevance over several generations. Constructing experimental populations in this way has two advantages when it comes to dissecting quantitative traits. Firstly, the use of multiple founders increases the genetic diversity captured. Secondly, the intercrossing over several generations results in a higher number of recombination events. These advantages were recently demonstrated in two proof-of-concept studies exploring height and awning in the Australian 4 founder MAGIC population (Huang et al., 2012) and the UK 8 founder NIAB Elite MAGIC population (Mackay et al., 2014), respectively. Both studies inspected levels of LD and found high levels of recombination in the respective populations under study. Since these initial demonstrations, the Australian MAGIC population has been used to dissect a diverse range of traits including seed dormancy (Barrero et al., 2015), coleoptile length and seedling growth (Rebetzke et al., 2014), and rhizosheaths size (Delhaize, Rathjen, & Cavanagh, 2015). The UK MAGIC population has been used to examine senescence (Camargo et al., 2016) and resistance against the necrotrophic fungi

Parastagonospora nodorum (Downie et al., 2018). Further wheat MAGIC resources have now been developed, including an 8 founder population relevant to Germany (Stadlmeier, Hartl, & Mohler, 2018), and another more broadly representative of European (Sannemann et al., 2018).

1.7.2 Resistance gene cloning

A step further to the genetic dissection of yellow rust resistance is gaining a functional understanding of the gene(s) and genetic variants underlying QTL, and how this translates into different modes of action in the plant during pathogenicity. To delve at this deeper molecular level, the nucleotide sequence of the gene loci in question must first be characterised. Traditionally, this has involved map-based cloning approaches, also known as positional cloning. Much more recently, rapid cloning techniques have started to emerge, providing a quick and cheaper alternative. Nevertheless, despite the large number of *Yr* genes (72) and yellow rust resistance QTL (>140) characterised in the past couple of decades, at the start of this thesis just four genes had been cloned (Fu et al., 2009; Krattinger et al., 2009; Liu et al., 2014; Moore et al., 2015), highlighting a bottleneck in the map-based cloning process. In this section I provide an overview of the approaches used to clone these genes, examine where the bottlenecks might lie and how emerging approaches have the potential to overcome some of the drawbacks of map-based cloning.

There are three main steps to map-based cloning: (i) Fine-mapping of a QTL or named gene and high-density genetic map development, (ii) anchoring of this genetic interval onto a physical map (cM to bp), (iii) identification of candidate gene(s) and confirmation of biological function.

In the first step, the genetic interval initially characterised is narrowed down to <1 cM, which usually requires the generation of additional germplasm. For example, *Yr18/Lr34* was initially mapped to a 3.6 cM interval on the short arm of chromosome 7D (Spielmeyer et al., 2005). Before a candidate gene could be identified, the region was narrowed down to a smaller target interval of 0.15 cM via genetic analysis of three different bi-parental populations, representing over 5,000 individual lines (S. G. Krattinger et al., 2009). Similar numbers have been reported

in other positional cloning studies for wheat rust resistance, for example 4,500 and 3,120 lines in the cloning of *Yr36* and *Lr10*, respectively (Fu et al., 2009; Stein et al., 2000). The development of such large populations therefore takes considerable time and investment. Secondly, additional markers are developed to increase marker density within the genetic interval, a process in wheat that has commonly utilised synteny between grass species, but more recently relies on the availability of the wheat reference genome sequence (International Wheat Genome Sequencing Consortium, 2014) and associated genomic resources.

Following fine-mapping, the narrowed loci must be anchored onto a physical sequence of the wheat genome. For wheat rust resistance genes, constructing such maps has historically been done by screening large Bacterial Artificial Chromosome (BAC) libraries (Krattinger et al., 2009: *Yr18/Lr34;* Huang et al., 2003: *Lr21*; Periyannan et al., 2013: *Sr33*). The relatively small size of each clone (100 kb) meant that they were manageable to sequence, essentially circumventing the drawbacks of dealing with the large 17 Gb hexaploid wheat genome. However, these libraries take a significantly long time to develop due to the large size of the wheat genome, its highly repetitive nucleotide sequence and three homologous genomes. This has historically hindered the speed at which gene cloning can be undertaken.

Once candidate genes have been identified within the physical intervals defined, their biological function must be confirmed to prove causality. This final stage can be performed using several approaches. *Yr18/Lr34* was found to encode an ABC transporter through mutant and haplotype analysis (S. G. Krattinger et al., 2009). For *Yr36*, the presence of several paralogous genes at the locus was narrowed down to the causal Kinase Start encoding gene through mutation and transformation evidence (Fu et al., 2009).

Map-based cloning has proven useful in the absence of a wheat reference genome for the cloning of *Yr* resistance genes. Nevertheless, the relatively small number of cloned *Yr* genes so far highlights this approach takes a considerable amount of time and financial investment for research groups. As QTL mapping increases in resolution, high-throughput genotyping becomes more affordable and whole reference genome assemblies become more readily available for wheat, the

historical bottlenecks to fine-mapping and map-based cloning are beginning to be alleviated.

It is highly likely that more rapid cloning approaches will be readily available over the course of this study. Evidence seems to point towards the application of complexity reduction technologies, such as that recently reported by Jupe et al. (2013). They developed a pipeline for *R* gene Enrichment Sequencing (RenSeq) in potato, that uses RNA probes to specifically capture NBS-LRR encoding genes. When used in conjunction with suitable germplasm resources, RenSeq enables the rapid mapping of *R* genes.

1.8 Yellow rust resistance breeding and research in the UK

In the UK, yellow rust is one of the major foliar diseases of winter wheat, where the disease has led to up to 50 % yield losses in susceptible varieties in some years. However, it is recurrent pathogen population changes in the causal agent *Pst* and resulting outbreaks of yellow rust that continue to impact UK disease resistance breeding programmes.

1.8.1 Advances in breeding for yellow rust resistance

Breeding for yellow rust resistance in the UK has historically relied on race-specific all-stage resistance, which typically leads to regular occurrences of boom and bust cycles. *Yr* resistance genes were usually deployed in a single fashion in varieties. As the acreage of these resistant varieties increased, the corresponding virulence would build up in the *Pst* population as the pathogen gradually adapted, until the variety was deemed completely susceptible. Notable examples of this rise and fall in resistance in the 1960s-90s include *Yr6, Yr9 and Yr17* (Boyd, 2005). The significant yellow rust outbreaks that followed the release of Rothwell Perdix in 1964, a variety carrying *Yr6*, led to the creation of the UK Cereal Rust Virulence Pathogen Survey in 1967 (Hubbard, Wilderspin, & Holdgate, 2017). Boom and bust cycles can be drastically short, with pathogen adaptation arising in under a year. This was the case for the *Yr9*-carrying varieties Clement (on the National List 1975 but not recommended) and Stetson (released in 1983, resistance broke down same year). In the early 1990s, it took four years for *Yr17* virulence to be detected in *Pst* isolates, following the deployment of *Yr17* in the variety Brigadier four years earlier (Bayles et al., 2000). In the 2000s, race changes in the UK *Pst* population led to the breakdown of resistance in varieties Robigus (2006) and Solstice (2009).

Since these recurrent outbreaks, there has been a much stronger drive to move away from race-specific resistance breeding, and to instead incorporate more durable resistance in UK wheat breeding programmes. Winter wheat varieties such as Cappelle Desprez and Alcedo are all varieties that were once commonly grown in France, Germany and the UK, and that remained resistant against yellow rust for considerable periods of time. These lines have gone on to become an important source of parental material for several of the modern UK elite cultivars, which have

displayed effective levels of yellow rust resistance (Angus, 2001). One prominent example of durable resistance against yellow rust in the UK is that of the Claire cultivar, which remained highly resistant against the disease for over 10 years (AHDB Recommended Lists Archive https://ahdb.org.uk/rlarchive).

More recently, it is the incursion of novel and genetically diverse *Pst* races into the UK that has significantly disrupted the resistance levels of UK wheat germplasm (Hovmøller et al., 2016; Hubbard, Lewis, et al., 2015). Since 2011, breeders in the UK and across Europe have had to develop varieties resistant against a completely novel *Pst* population, following the total replacement of the historically clonal population by one made up of multiple lineages of *Pst* that are genetically diverse. A year after the incursion of the Warrior race in 2011, varieties like Claire and those related to it, once showing effective and often long-term levels of yellow rust resistance, were becoming susceptible (Hubbard & Bayles, 2013). In 2016, another significant race change occurred in the UK *Pst* population. The novel virulence combination these races displayed was effective against a number of resistant varieties on the Recommended List. Reflection, Zulu and Myriad all saw their yellow rust ratings significantly fall by ≥ 3 as a result of this in the following year (Hubbard et al., 2017; AHDB Recommended Lists Archive https://ahdb.org.uk/rlarchive).

1.8.2 The genetic basis of yellow rust resistance in UK wheat germplasm

Significant advancements in genetic mapping of yellow rust resistance loci and molecular marker development have aided the characterisation of known sources of resistance. This has been of particular benefit to our understanding of APR, which is quantitative in nature and often the result of the expression of several resistance genes each conferring partial resistance.

The genetic basis of some of the yellow rust resistance deployed in the UK, both all-stage and APR, has been studied and reported for several varieties, including Alcedo (Jagger et al., 2011), Claire (Powell et al., 2013) and Guardian (Melichar et al., 2008). Yellow rust resistance of these varieties was investigated in bi-parental populations at the seedling and adult plant stages. These studies have enabled the characterisation of over ten major and small-effect QTLs conferring mostly APR

but also all-stage resistance, located on chromosomes 1B, 2B, 2D, 3B, 4B, 5A and 7B. Between three and five QTLs were detected in each study, comprising one to two major effect QTLs and one to three QTLs of minor effect. Additivity effects between QTLs conferring APR was observed in the Alcedo x Brigadier population (Jagger et al., 2011). These genetic mapping studies have been valuable in increasing our knowledge of the location of yellow rust resistance loci and have yielded molecular markers for use in breeding programmes. But because they were carried out in bi-parental populations genotyped with a relatively low number of markers (as little as 279 in Melichar et al., (2008)), they ultimately only provided a granular and transient snapshot into the resistance levels of UK wheat germplasm.

Beyond the single variety lens, genome-wide association studies have allowed for a much more detailed insight. Bansept-Basler (2013) harnessed the UKCPVS seedling and adult plant data gathered over three decades for 327 winter wheat varieties listed on the Recommended List and used genome-wide associations mapping to reveal 23 marker-trait associations. This association mapping panel was also used for in-field testing against *Pst* races of relevance at the time of the study. Field-based resistance was governed by 23 loci representing all-stage resistance and APR. Bansept-Basler's (2013) study demonstrated the use of historical data as a cost-effective starting point for assessing yellow rust resistance levels in UK wheat germplasm. Reassuringly, it also showed that there were high levels of field-based resistance against relevant *Pst* isolates present in the panel.

The association mapping panel created by Bansept-Basler (2013) to specifically examine yellow rust resistance formed the basis of another association mapping panel. The Wheat Association Genetics for Trait Advancement and Improvement of Lineages (WAGTAIL) panel was developed as a collaboration between academia and industry to provide the first comprehensive overview of the genetic factors underlying resistance against the four major foliar diseases of wheat in the UK and North Western Europe i.e. Septoria tritici blotch, yellow rust, brown rust and powdery mildew (Gardner and Cockram, unpublished; BBSRC reference BB/J002607/1). Field trials for this project spanned the crucial period of 2012-2015, which saw dramatic shifts occur in the genetic make-up and structure of *Pst*

population across Europe (Hovmøller et al., 2016; Hubbard, Lewis, et al., 2015). One of the key outcomes of genome-wide associations in WAGTAIL has been the deployment of independently validated molecular markers for yellow rust resistance loci effective against the post-2011 highly diverse *Pst* population.

These recent genetic mapping studies altogether represent a substantial body of work that has been crucial in informing breeders of the range of resistance loci available in UK wheat germplasm to date.

1.9 Aims

The study of wheat yellow rust resistance is of a dynamic nature and should always progress alongside changes observed in *Pst* populations and advances in phenotypic, genetic, and genomic resources.

My thesis therefore sets out to examine the genetic basis of yellow rust resistance in UK breeding germplasm and build on the knowledge gathered so far, by exploiting the high-resolution potential of the eight-founder NIAB Elite MAGIC population. Specifically, the aims are to:

1. Identify QTLs conferring field-based yellow rust resistance at the adult plant stage, in both the leaves and glumes.
2. Compare and contrast the resistance mechanisms that operate in these two tissues.
3. Identify candidate genes located within the physical intervals of identified QTLs and infer potential disease resistance mechanisms.
4. Generate resources with which to underpin future map-based cloning of YR QTLs.

2 GENERAL MATERIALS AND METHODS

2.1 Plant material and experimental design

The development of the NIAB eight-parent winter wheat MAGIC population is described in Mackay et al. (2014). Table 8.1 in Appendix B provide an overview of the postulated and/or reported all-stage resistance and APR.

2.1.1 2015 pathology trials

F_7 MAGIC lines were grown across three untreated pathology trials in 2015. The seed for each trial originated from the NIAB 2015 multiplication nursery. 2015 MAGIC pathology trials were grown at three different locations: Osgodby (OSG15) and Rothwell (ROTH15) in Lincolnshire and NIAB's experimental field trials station in Cambridge, Cambridgeshire (NIAB15). OSG15 and ROTH15 are owned by the wheat breeding company Limagrain UK Ltd.

Table 2.1 provides a summary of the plant material, trial layout and experimental design for each trial in 2015. All trials were designed with the software DEW (www.expdesigns.co.uk).

2.1.1.1 NIAB 2015 MAGIC pathology trial

The 1,085 F_7 MAGIC lines and eight MAGIC founders were replicated twice in the NIAB15 trial. Controls were also replicated twice and included a positive (Vuka) and a negative control (Cougar, disease rating of 8 on AHDB Recommended Lists 2013-2016), together with several different varieties routinely used in YR differential tests followed an incomplete randomisation block design consisting of 92 blocks. Each block consisted of 24 treatments (MAGIC line or control). Each treatment was randomly allocated to a block. Figure 8.5 (Appendix A) illustrates the field layout of NIAB15.

In NIAB15, a traverse consisted of several adjacent 1x1m plots (Figure 8.2, Appendix B). Double rows of varieties highly susceptible to yellow rust, referred to as 'spreader rows', were planted in the middle of each plot. Treatments were sown in double rows on either side of the spreaders.

Table 2.1 2015 pathology trials summary

	NIAB15	OSG15	ROTH15
Location	Cambridge Cambridgeshire	Osgodby Lincolnshire	Rothwell Lincolnshire
MAGIC lines	1,085	1,060[a]	1,060[a]
MAGIC founders	8	8	8
Varietal controls	9	2	2
MAGIC controls	-	20	20
Additional checks	-	Vuka[b]	Vuka[b]
Spreader positions	Every traverse	Every 3 traverses	Every 4 traverses
Trial design	Unbalanced incomplete random block design	Unbalanced incomplete random block design	Unbalanced incomplete random block design

All entries were replicated twice apart from a and b. a: Unreplicated MAGIC lines; b: Vuka replicated 80 times. A traverse consists of nth plots, with each plot made up of six one metre long rows.

Table 2.2 NIAB15 YR control varieties

Variety	Type	YR resistance genes
Ambition	Differential	*Am*
Cadenza	Differential	*Ca*
Cougar	Negative Control	*Co*
KWS Sterling	Differential	*St*
Rendezvous	Differential	*Re*
Solstice	Differential	*So*
Spladings Prolific	Differential	*Sp*
Timber	Differential	*Ti*
Vuka	Positive control	-
Warrior	Differential	*Wa*

Varieties listed as differentials are routinely used in pathotyping tests because of their known YR resistance genes.

2.1.1.2   Limagrain 2015 MAGIC pathology trials

To test all MAGIC lines at each site, unreplicated trials consisting of 1,080 F_7 MAGIC lines were planted at OSG15 and ROTH15. A single fully replicated pathology trial was not possible at the two Lincolnshire sites due to shortage of field space at each site. The positive and negative controls used at NIAB15 were also included. In addition, a subset of 20 MAGIC lines was replicated to provide additional positive and negative controls. These MAGIC lines were selected based on their response to YR infection in previous trials, as indicated in Table 8.2 (Appendix B). In each trial, the varietal and MAGIC controls, as well as the eight founders, were replicated twice. OSG15 and ROTH15 followed an incomplete randomisation block design. All controls were randomly allocated to blocks so that each block contained two to three controls. Each trial consisted of 28 blocks, each comprised of 40 treatments. In addition to these controls, the variety Vuka was sown every 20 treatments across each trial, as a check for drilling errors. A summary of the field layout for the Limagrain trials, using ROTH15 as an example, is given in Figure 8.6 (Appendix B).

At OSG15 and ROTH15, MAGIC lines were not individually exposed to a spreader, unlike NIAB15. Figure 8.6 illustrates how spreaders were sown as a separate traverse every three to four treatment traverses. Spreader traverses consisted of double rows of Armada, Victo and Cerco, varieties highly susceptible to brown rust, yellow rust and mildew respectively. Each of the treatment traverses has an empty double row in the middle of it. This follows the standard pathology field trial set up used by Limagrain UK Ltd.

2.1.2   2016 pathology trials

The 2016 trials focused largely on assessing the subset of genotyped MAGIC lines in partially replicated trials at two locations: Osgodby (OSG16) and NIAB, Cambridge (NIAB16). The seed used for these trials originated from the NIAB 2016 multiplication nursery.

Of the 707 genotyped MAGIC lines, there was sufficient seed for 678 F_8 MAGIC lines only. In addition, 48 ungenotyped F_8 MAGIC lines were selected based on 2015 infection severity data: 24 highly resistant and 24 highly susceptible lines as

specified in Table 8.3 (Appendix B). In total, 726 F_8 MAGIC lines were grown in two untreated pathology trials in 2016.

In both trials, 65 % of genotyped MAGIC lines were replicated while the remaining 35 % were present only once. This represented 444 and 234 replicated and unreplicated MAGIC lines respectively. MAGIC lines to be replicated were selected at random using a true random numbers generator (www.random.org). Controls in each trial included the eight founders replicated three times, together with Oakley and Cougar as positive and negative controls respectively. Treatments were randomly allocated to 12 sub-blocks, each nested within a block, nested within a replicate block, as illustrated in Figure 8.7 (Appendix B). Each sub-block consisted of 100 treatments. The number of controls per sub-block varied between one and four treatments. Table 2.3 provides a summary of the plant material, trial layout and experimental design for each pathology trial in 2016.

All treatments in NIAB16 were adjacent to a spreader double-row while at OSG16, spreaders were sown every three traverses.

Table 2.3 2016 pathology trials summary

	NIAB16	OSG16
Location	Cambridge, Cambridgeshire	Osgodby, Lincolnshire
Genotyped MAGIC linesa	678	678
Ungenotyped MAGIC lines	48	48
MAGIC founders[b]	8	8
Positive control[b]	1	1
Negative control[b]	1	1
Spreader positions	Every traverse	Every 3 traverses
Trial design	Unbalanced incomplete random block design	Unbalanced incomplete random block design

a: 444 MAGIC lines were replicated twice while the remaining 234 were present once in each trial; b: MAGIC founders and varietal controls were replicated three times. A traverse consists of nth plots, with each plot made up of six one metre long rows.

2.2 Genotypic data and MAGIC genetic map

643 F_5 MAGIC lines were genotyped with the Illumina Infinitum iSelect 90,000 SNP wheat array (Wang et al., 2014), using a single seed from each individual MAGIC line. The resulting genotypic data was reported in Mackay et al. (2014) and Gardner et al. (2016). Of the 20,639 polymorphic markers obtained, 18,601 were mapped to 4,578 unique locations across the wheat genome, resulting linkage map of 5,305 cM in total (Gardner et al., 2016), as illustrated in Figure 2.1.
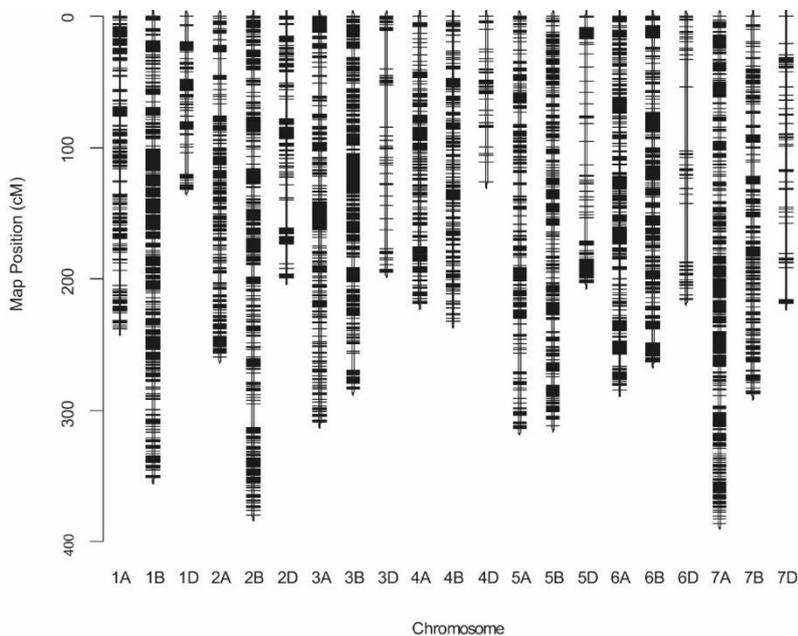
Figure 2.1 Genetic map of the NIAB Elite MAGIC population.

Chromosomes short arms on top, at 0 cM. Taken from Gardner et al. (2016).

2.3 *P. striiformis* inoculum: inoculation and natural infection

OSG15, ROTH15, NIAB15 and NIAB16 were all artificially inoculated with a mixture of *Pst* races Solstice (isolate 08/21 virulent on *Yr 1, 2, 3, 4, 6, 9, 17, 25, 32*) and Warrior (isolate 11/08 virulent on *Yr 1, 2, 3, 4, 6, 7, 9, 17, 25, 32, Sp*). Victo seedlings, one infected with the Solstice race and the other with the Warrior race were transplanted in pairs at regular intervals in spreader rows in all four trials. Transplanting of infected seedlings was undertaken in spring, when all lines in the trials had reached the tillering stage (Zadoks Growth Stage (GS) 23-25, (Zadoks, Chang, & Konzak, 1974). Levels of *Pst* infection were particularly high in the spreader rows of OSG16, and so no artificial inoculation was undertaken. At the time of transplanting, symptoms of yellow rust were already visible in trials so it is highly likely that the infection occurring during phenotypic assessment of the MAGIC population did not originate from the artificial inoculation.

Sufficient inoculum was generated by multiplying *Pst* spores for each of the isolates described above. Victo seedlings in 96-well trays (4-5 seedlings per well) at the one leaf stage were inoculated with a dry mixture of talc and spores following a 1:16 ratio (spores:talc). The mixture was evenly applied to seedlings in trays using a rotary inoculator. Trays were incubated in plastic bags in the dark at 4°C for 48 hours. Following the incubation period, bags were removed, and seedlings were grown under controlled conditions (16/8 hours, at 18/11°C, day/night cycle). Spores were harvested for two to three weeks, 14 days post-inoculation onwards.

2.4 Phenotypic assessments

2.4.1 Yellow rust on leaves

YR infection severity of the leaves was assessed as the percentage of total leaf tissue infected by YR. This was estimated using the modified Cobb's scale (Peterson, Campbell, & Hannah, 1948) ranging from 0 to 100 %. Traces of yellow rust observed in the plot were indicated as 0.01 %. Leaf infection severity was assessed on three to four occasions for each trial, from the end of booting (Zadoks Growth Stage (GS) 45-49) until the mid-to hard-dough stage (GS 85-87), at 12 to 18 days intervals (Zadoks et al., 1974).

2.4.2 Agronomic traits

Three agronomic traits were assessed and used to identify seed preparation and drilling errors in each trial: heading time (HT), height (H) and awn presence/absence (A). Further details on the phenotypic assessment of agronomic traits can be found in Table 2.4.

Table 2.4 Summary of measured agronomic traits in pathology trials

	NIAB15	OSG15	ROTH15	NIAB16	OSG16
Heading time[a]	yes	no	no	no	no
Height[b]	yes	no	no	yes	yes
Awns[c]	yes	yes	yes	yes	yes

a: Heading time was estimated as the number of days from the sowing date, when at least 50% of the plot had reached GS55. b: Height was estimated by planting a ruler in the middle of the plot and taking a measurement from the base of the stem to the top of the head, excluding awns. c: Plots were given overall scores, with 0 = not awned and 1 = awned.

3 CHARACTERISATION OF QTLS CONFERRING YELLOW RUST RESISTANCE IN LEAVES

The last two decades have seen the emergence of rapidly adapting *Puccinia striiformis* f. sp. *Tritici* (*Pst*) populations with high virulence potential globally. These populations have had a significant detrimental effect on the resistance levels of varieties that had previously shown durable resistance to yellow rust (YR). These changes in the *Pst* population landscape require adequate resistance breeding strategies, starting with the continued characterisation of adequate YR resistance loci that can be rapidly deployed in current elite germplasm. With this purpose, I used the eight-founder NIAB Elite MAGIC population, which is representative of key varieties used in UK wheat breeding, to characterise YR resistance loci at the adult plant stage and explore the phenotypic effect of their combinations. Replicated and partially replicated trials across three sites in 2015 and 2016 revealed the consistent detection across environments of four major QTLs on chromosomes 1A (*QYr.niab.1A.1*), 2A (*QYr.niab-2A.1*), 2B (*QYr.niab-2B.1*) and 2D (*QYr.niab-2D.1*). Altogether, these QTLs explained nearly 50 % of the phenotypic variation observed. Five small-effect QTLs were also detected, *QYr.niab-2A.2, QYr.niab-3A.1, QYr.niab-6A.1, QYr.niab-6A.2* and *QYr.niab.6B.1*, cumulatively explaining 15-20 % of the phenotypic variation but with inconstant detection between environments. The phenotypic effect of QTL combinations, both 2-way and 3-way, showed that combinations were more effective in conferring resistance than QTL in isolation. In addition, a further five small-effect QTLs, *QYr.niab-3D.1, QYr.niab-4B.1, QYr.niab-4D.1, QYr.niab-5A.1, QYr.niab-6A.3*, were detected as borderline QTLs, falling just below the 5 % significance threshold.

3.1 Introduction

Yellow rust, caused by the biotrophic fungus *Puccinia striiformis* f. sp. *tritici* (*Pst*), is a substantial threat to wheat production globally. Since the 2000s, a subset of genetically diverse and divergent lineages of *Pst* have been responsible for recurrent epidemics in a number of wheat-producing regions (Ali, Gladieux, et al., 2014; Ali et al., 2017). The rapid adaptation and subsequent spread of these lineages in previously hostile environments has given rise to more aggressive pathotypes generally better adapted to higher temperatures (Chen et al., 2002; Milus et al., 2009; Wellings, 2007). In the UK and North-Western Europe, the historically clonal *Pst* population has been completely displaced by a genetically diverse group of lineages (Hovmøller et al., 2016; Hubbard et al., 2015). First detected in 2011 simultaneously across several European countries, the 'Warrior' race and the 'Kranich' race likely originated from the near Himalayan region and rapidly spread throughout the continent as a group of genetically distinct lineages. *Pst*'s ability to migrate over long distances, to locally adapt to new environments and to displace established populations, has had a significant impact on the resistant levels of wheat varieties in the UK (Hubbard, Pritchard, Coventry, & Holdgate, 2015) and beyond (Wellings, 2011), prompting breeders to review their YR resistance breeding and deployment strategies.

To tackle yellow rust, resistance breeding strategies have historically focused on the utilisation of major yellow rust resistance (*Yr*) genes in isolation. *Yr17*, for example, was a popular source of such resistance in North-Western Europe. Its effectiveness stopped soon after its deployment as a single resistance gene over a large wheat acreage (Bayles et al., 2000). Over-reliance on major resistance genes has largely been left behind, in favour of the more durable approach of combining resistance gene underpinned by different resistance mechanisms e.g. Adult Plant Resistance (APR) and all-stage resistance (Chen, 2013; Singh et al., 2005). Utilisation of genetic resistance in this way therefore requires the characterisation of *Yr* genes and QTLs, a detailed understanding of how they interact with each other, and molecular markers to track them in breeding programmes.

At present, over 300 genomic regions conferring yellow rust resistance and spanning all wheat chromosomes have been reported in the scientific literature

(Chen & Kang, 2017) and approximately 80 *Yr* genes have been catalogued (McIntosh et al., 2017, 2003). QTLs and *Yr* genes have been characterised in experimental mapping populations, mostly biparental (reviewed in Rosewarne et al., 2013) and more recently in Association Mapping (AM) panels (Jighly et al., 2015; Kertho et al., 2015; Maccaferri et al., 2015; Zegeye et al., 2014). Crossing just two parents in a biparental population invariably limits the number of resistance genes that can be investigated in a single population (Mackay, 2001). AM panels overcome this limitation by exploiting the resulting Linkage Disequilibrium (LD) from historical recombination events present in each accession of a germplasm collection composed, for example, of varieties and/or breeders' lines (Mackay & Powell, 2007). However, population structure in such collections can lead to false marker-trait associations if not carefully taken into consideration. While both experimental populations have proven complementary in the characterisation of *Yr* resistance genes, their limitations have hindered the discovery of new sources of resistance and its genetic dissection at high-resolution.

In the last 10 to 15 years, multi-founder experimental populations have been developed as an alternative to the populations described above (Cockram & Mackay, 2018). Of these, one of the most popular is the Multi-Parent Advanced Generation Inter-Cross (MAGIC) population (Cavanagh et al., 2008). First conceived and implemented in mouse research (Churchill et al., 2004), the concept of inter-crossing inbred lines over several generations to generate a multiparent population and its potential for QTL mapping was eventually extended to the plant kingdom (e.g. Huang et al., 2015; Kover et al., 2009). MAGIC populations have the ability to capture increased genetic variation and a greater number of recombination events, thus providing power and precision when mapping QTLs. The UK wheat MAGIC population, termed the 'NIAB Elite MAGIC' population (also referred as just the MAGIC population in subsequent sections), has eight elite founders and was estimated to capture over 80 % of the genetic variation observed in UK wheat at the time of its development (Gardner et al., 2016). Founders were selected in close collaboration with breeding companies to ensure the population has direct relevance to the wheat breeding industry. A high-

density MAGIC genetic map has recently been developed for high-resolution dissection of quantitative traits (Gardner et al., 2016).

3.2 Aims and overview

To date, QTL mapping of yellow rust resistance in wheat MAGIC populations has not yet been explored and a deeper understanding of this complex trait at the adult plant stage is required to inform rust resistance breeding efforts. The objective of this chapter was therefore to characterise the genetic basis of yellow rust in the NIAB Elite MAGIC population at the adult plant stage. This comprehensive study explores the use of different QTL mapping methods for characterisation of genomic regions of interest for yellow rust resistance. The study is divided in three parts. I first start by describing the phenotypic variation in yellow rust resistance observed in five independent trials conducted over two seasons and at three contrasting UK locations. This is followed by a detailed QTL mapping study based on several different statistical methods exploiting founder haplotype probabilities as well as marker allelic effects. I finally reflect on the results of this study in the discussion, focusing on what they mean for yellow rust resistance breeding in the UK and beyond, and providing an outlook on future work.

3.3 Materials and methods

This chapter focuses on the phenotypic data collected from the 2015 and 2016 MAGIC pathology trials, as outlined in Table 3.1. For further details on trial layout and phenotypic assessment scales, refer to Section 2.4 and Appendix B.

Table 3.1 Overview of 2015 and 2016 MAGIC pathology trials

	NIAB15	OSG15	ROTH15	NIAB16	OSG16
Location	Cambridge Cambs	Osgodby Lincs	Rothwell Lincs	Cambridge Cambs	Osgodby Lincs
MAGIC lines replicated	1,085	-	-	444	444
MAGIC lines unreplicated	-	1,060	1,060	234	234
Varietal controls	9	2	2	2	2
MAGIC controls	-	20	20	-	-
Additional checks	-	Vuka	Vuka	-	-
YR scores (S)	S2-3	S2-3	S2-3	S2-3	S2-3

3.3.1 Phenotypic data quality control and transformation

Prior to undertaking the statistical analyses, the raw phenotypic data from each pathology trial was checked for errors in trial sowing, extreme phenotypic values and data entry errors. Errors were identified and rectified by:

- Comparing the present agronomic data with historical height and awn data from previous MAGIC nurseries to identify discrepancies for each entry.

- Sorting YR infection scores in ascending order to identify values out of range.

- Plotting residuals for replicated trials to identify extremes.

- Checking field book notes.

When an error was identified, the original data was checked against any notes that may help in identifying the origin of the discrepancy. When no obvious

explanation was found for the identified error, the data point was kept. In addition, MAGIC lines for which no scores were available due to senescence for example, were removed from further analysis.

Infection severity scores were transformed in cases where residuals were not normally distributed. Normality was assessed by plotting a histogram of the residuals. Transformation was performed by applying the natural log to the original data:

$$y = \ln{(x + 1)}, \tag{1}$$

where x is the original score and y is the transformed score. The leaf infection severity scale starts at zero, a value not taken by the natural log equation. A value of one was thus added to all original scores.

3.3.2 Statistical analysis of the 2015 and 2016 pathology trials

A step-wise model selection approach was used to estimate the Best Linear Unbiased Estimators (BLUEs) for MAGIC lines, integrating spatial and non-spatial mixed linear methods based on Restricted Maximum Likelihood (REML). Three models were considered:

- Model 1 - Blocking. Genetic effects are estimated based only on the inter- and intra-block variation recovered from the model.

- Model 2 - Spatial. Only considers global and/or local field trends.

- Model 3 - Spatial + blocking. Combination of the above models.

Initially, each model was optimised by including field trends running in either row or column direction or accounting for the walking route taken when scoring, termed 'Scoring Order' (ScO). Refer to Appendix C for details on ScO for each trial. This was followed by between-model comparison to select the one that best fits the data. Several diagnostic tools and statistical tests were used for model optimisation and comparison, for which a more detailed description is found in section 3.3.2.1 below. Models were fitted using the 18th edition of the statistical software Genstat (VSN International 2015).

3.3.2.1 Baseline statistical models

Model 1: Blocking

Model 1 has a baseline model consisting of a fixed genotypic effect for MAGIC lines, a random block effect and an error term. The block effect consisted of different terms depending on the trial: 'block' for trial NIAB15, 'rep/block' for OSG15 and ROTH15 and 'rep/block/sub_block' for NIAB16 and OSG16. The '/' operator denotes nested terms. Model 1 can be formulated as:

$$y = \mathbf{X}_g g + \mathbf{Z}u + e \tag{2}$$

where y is the vector of YR infection in a column x row matrix, $\mathbf{X}_g g$ contains the fixed effect for MAGIC lines, $\mathbf{Z}u$ contains the random block effects and e is the residual error term. Model 1 was optimised by independently fitting random terms for column and ScO.

Model 2: Spatial

In the spatial model, MAGIC lines were also included as fixed effects. Global trends were incorporated as trends aligned with rows and/or columns. Local trends were incorporated in the model with correlated residuals from the first-order autoregressive processes across rows and columns, now referred to as the AR1 x AR1 model. Model 2 can be formulated as:

$$y = \mathbf{X}_g g + \mathbf{X}_\beta \beta + \xi + e \tag{3}$$

where $\mathbf{X}_g g$ contains the same fixed terms as Model 1 and vector β and ξ contain the spatial trends. Vector β may include linear global trends while ξ contains the local trends as the spatially correlated residuals from the AR1 x AR1 model. e is the residual error term. Model 2 was optimised by fitting global trends and/or an additional random term for ScO.

Model 3: Blocking + Spatial

In the last model, the non-spatial model serves as a baseline to the addition of spatial trend terms, following the sequential approach used by Gilmour et al (1997). This combined model can be formulated as:

$$y = \mathbf{X}_g g + \mathbf{X}_\beta \beta + \mathbf{Z}u + \xi + e \tag{4}$$

where $\mathbf{X}_g g$, $\mathbf{X}_\beta \beta$, $\mathbf{Z}u$, ξ and e are defined as in model formulations (4) and (5).

3.3.2.2 Best model selection

The Akaike Information Coefficient (AIC) was used as a measure for model selection: the better the model, the lower the AIC (Akaike, 1974). Each model was optimised using the AIC as a measure of model fit improvement. The resulting values were compared and the model with the lowest AIC values selected. Although (Gilmour, Cullis, & Verbyla, 1997) used the sample variogram as a selection tool for spatial models, it was not used here because of the large number of models to process.

3.3.2.3 Broad sense heritability

Heritability was used as a measure of total phenotypic variation attributable to the genotypic effect. The VHERITABILITY function in Genstat (18th edition, VSN International 2015) was used to calculate broad sense heritability for each pathology trial. The function is based on the definition of heritability given by Cullis et al. (2006) and Piepho and Möhring (2007).

3.3.3 Genetic analysis pipeline for QTL mapping

Once YR scores were adjusted for field trends, BLUEs were incorporated in the genetic analysis pipeline to identify YR resistance QTLs in the MAGIC population.

3.3.3.1 QTL mapping approaches

Analyses of marker allelic state and founder haplotypes were the two approaches used for identifying QTLs conferring YR resistance in the MAGIC population. Founder haplotype probabilities were calculated with the mpprob function in R/mpMap (Huang & George, 2011). QTL mapping analyses undertaken in this study are listed below:

- Single Marker Analysis (SMA): Regression analysis on allelic state of 7,369 mapped SNP markers from the MAGIC genetic map (Gardner et al., 2016).

- Haplotype analysis (HA): Regression analysis on founder haplotype probabilities.

- Interval mapping (IM): IM with founder probabilities and no covariates.

- Composite Interval Mapping (CIM): CIM with founder probabilities and 10 marker covariates.

SMA and HA analyses were performed with R/lme4 in the R environment (R Core Team, 2017). IM and CIM were carried out in R/mpMap.

SMA: Regression on allelic state

The adjusted YR scores for each MAGIC line were regressed on the 7,367 SNP markers from the MAGIC map. For each regression, the following mixed model was applied:

$$Y = \mu_x + G_m + \beta + e \tag{5}$$

where Y is the YR resistance value, μ is the adjusted YR score for MAGIC line x, G_m is the fixed SNP marker effect, β is the population structure consisting of 'funnels' and 'plants within funnels' effects, and e as the residual error term. Model (7) has one degree of freedom, since regression is carried out on allelic state i.e. 0 or 1. The p values were adjusted for false positives by estimating q values with the R package R/qvalue (Storey, 2015). A q value significance threshold was set empirically at 0.05.

HA: Regression on founder probabilities

Here, adjusted YR scores were regressed on founder probabilities. The following mixed model was used:

$$Y = \mu + G_p + \beta + e \tag{6}$$

where Y, μ, β and e are as for model (7) and G_p is the fixed term for founder probabilities. Here, the statistical model has seven degrees of freedom. The p values were adjusted for false positives and the empirical significance threshold computed, as in the SMA.

IM and CIM approaches

Founder probabilities were used to perform IM with 0 marker covariates and CIM with 10 marker covariates in R/mpMap. Within mpMap, an automated forward selection process based on AIC values selects the best ten marker covariates for each MAGIC line. Significant QTLs are then selected in two stages. First, mpMap

scans the 100 markers surrounding a particular marker location and selects QTLs based on a threshold of $-\log10(p)>3$. The number of significant QTLs is then reduced by fitting a model with $p<0.05$ and with percentage variation explained (PVE)>0.5%. Additionally, R/mpMap outputs founder contributions for each significant QTL, computed using a regression approach at each marker location. The empirical $p=0.05$ significance threshold was computed in R/mpMap using the sim.sig.thr function.

Cut-off p values for each score and mapping method are summarised in Table 3.2 below.

Table 3.2 p value thresholds (expressed as $-\log10(p)$, 5 % significance) at all five tested environments, for Interval and Composite Interval Mapping

	Interval Mapping		Composite Interval Mapping	
	S2	S3	S2	S3
NIAB15	4.11	4.58	4.57	4.81
OSG15	4.26	4.84	4.32	4.67
ROTH15	4.25	4.12	4.37	4.48
NIAB16	4.71	4.59	4.68	4.88
OSG16	4.56	4.44	4.49	4.53

3.3.3.2   QTL peaks: identification and selection

For each SNP marker, the p and q values from the four different mapping methods for all adjusted and log transformed YR scores, in all environments, were aligned to the MAGIC genetic map, and compiled into a single table. Firstly, the most significant marker at each QTL interval was manually selected for SMA and HA approaches, using q value thresholds as calculated in section 3.3.3.1. These locations were then compared to that automatically identified in the IM and CIM approaches and consensus peak markers were selected using the following criteria:

▪ Consensus between the different QTL mapping approaches – For each, the peak marker at a given QTL interval was the marker with the most significant q values. This value was compared across all mapping

approaches. Markers found significant only with CIM 10 covariates were not considered.

- Consensus between the different YR scores across 2015 and 2016 – A marker was selected as the overall consensus peak marker at a given QTL interval when it was the most significant in at least two YR scores in a single trial.

In addition, the genomic context (centromere, introgressions, translocations) and marker density were also taken into consideration when selecting an overall peak marker.

A single peak marker was then identified by examining the physical map position of each peak marker candidates. The physical map is based on the IWGSC RefSeq v1.0 wheat genome assembly (Appels et al., 2018). The locations of the genetically mapped MAGIC SNPs on the physical map were determined by Santos and Gardner (2017, unpublished).

Nomenclature for the QTL discovered in this study follows that recommended by the Catalogue of Gene Symbols for Wheat (McIntosh et al., 2008).

3.3.4 QTL effects

MAGIC lines were divided into 16 different genotypes based on the presence of the four most significant ($p<0.0001$, PVE>8 %) YR QTLs *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B.1* and *QYr.niab-2D.1*. For each QTL, the resistant allele was based on the allelic state of the consensus peak markers for each MAGIC founder (Table 3.3). A one-way ANOVA and a *t* test comparison was performed to identify any significant differences in % disease severity (% DS) means between all the possible QTL combinations (2-way and 3-way combinations). Both statistical analyses were undertaken with predicted means from the third YR disease assessment.

Table 3.3 Allelic state of MAGIC founders for the four main YR QTL

	QYr.niab-1A.1	*QYr.niab-2A.1*	*QYr.niab-2B.1*	*QYr.niab-2D.1*
Al	0	0	2	2
Br	0	0	2	0
Cl	0	0	2	2
He	2	2	2	0
Ri	0	2	2	0
Ro	0	0	2	0
So	0	2	0	0
Xi	0	2	2	0

Allele presented here correspond to that of the peak SNP marker for each QTL (*QYr.niab-1A.1, QYrniab-2A.1, QYr.niab-2B.1, QYr.niab-2D.1*). Cells highlighted in grey correspond to the resistant alleles. MAGIC founders are abbreviated as Al: Alchemy, Br: Brompton, Cl: Claire, He: Hereward, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19.

3.4 Results

3.4.1 MAGIC population response to *Pst* leaf infection

3.4.1.1 Statistical analysis of individual trials

Three baseline mixed linear model approaches were applied to the 2015 original data and the log transformed data for the 2016 pathology trials data: Model 1 – blocking, Model 2 – spatial, Model 3 – blocking + spatial. The AIC value was used as an indicator of model fitness in the selection process. Each baseline model was optimised by accounting for blocking structure and/or field trends and the model with the lowest AIC value was selected as the best mixed linear model (Table 3.4).

Overall, model optimisation showed that field trends in the column direction were present at all locations, except for NIAB15. Therefore, the simplest model i.e. blocking, could be improved by taking global and/or local field trends into account. Model 1 included global trends in the form of a random term for column or scoring order. Model 2 included local trends in the spatial analysis and local trends with the addition of random terms for rows and/or columns and fixed terms for linear row and column trends. Model 3 was a combination of the above.

Due to the unreplicated nature of trials OSG15 and ROTH15 and the presence of highly replicated checks, spatial analysis (model 2) was the most adequate model to capture the presence of field trends. Models 1 and 3 failed to fit the data for those trials.

The normality tests undertaken on all YR scores revealed different patterns of distribution for residuals. 2015 residuals were normally distributed, regardless of location and therefore required no further transformation (Section 8.1, Appendix C). In 2016 however, the residuals diagnostic plots confirmed the skewness observed in the raw data distribution. This was strongly observed in the fitted-value and normal plots: data points form clusters in the fitted-value plots while, in the normal plots, they do not follow a straight line through the middle (Section 8.1, Appendix C). Normality of the 2016 YR scores was improved with the log transformation (Section 8.1, Appendix C). Transformed values were used in subsequent analyses. For NIAB16, transforming YR scores highlighted outliers for

S2 and S3 previously not evident from the raw data residuals. Going back to the original data and comparisons with historical agronomic data for MAGIC lines did not reveal any obvious error. The original values were thus maintained for subsequent analyses.

Table 3.4 Summary of best models per score for 2015 and 2016 pathology trials

Trial	Score	Best model
NIAB15	S2	Blocking + Column
	S3	Blocking
OSG15	S2	Spatial
	S3	Spatial
ROTH15	S2	Spatial
	S3	Spatial
NIAB16	S2	Blocking + Scoring Order
	S3	Blocking + Scoring Order
OSG16	S2	Spatial + Scoring Order
	S3	Blocking + Spatial + Column

3.4.1.2   Assessment of yellow rust resistance

YR scores were similarly distributed at the different sites within each year but differed considerably between 2015 (Figure 3.1) and 2016 (Figure 3.2).

In 2015, as YR progressed through the season and susceptibility increased among the MAGIC population, a subset of lines remained highly resistant to YR (YR % Infection (IF) <1) at all three NIAB15, OSG15 and ROTH15 sites. Additionally, intermediate YR scores (between 10 and 80 % DS) were normally distributed, as illustrated in Figure 3.1. This trend was more evident at OSG15 and ROTH15, where *Pst* developed more gradually than at NIAB15. In the latter trial, a higher number of highly resistant MAGIC lines (% DS<1) was observed in the last score (S3, 84 lines) compared to the previous (S2, 60 lines). This is explained by a sub-set of 28 lines that had a higher % DS score at S2 compared to S3. The % DS scores of these 28 MAGIC lines at the second score ranged between 1-5 %, these lines can thus still be considered as resistant. The only exceptions were MEL_032-3 (mean % DS =

12.5) and MEL_062-2 (mean % DS = 30.5), which were not included in QTL mapping. Overall, 142 MAGIC lines exhibited a resistant response (% DS<10) to *Pst* throughout the 2015 scoring season, at all three locations.

In 2016, YR scores were skewed towards the resistant end of the % DS scale (Figure 3.2 and Figure 3.4). It is worth noting that the peak of resistant MAGIC lines observed across trials in the previous year was less prominent in 2016 and in particular for NIAB16 S3.

Phenotypic variation in the MAGIC parents spanned the entire YR % infection scale, as depicted in Figure 3.3 and Figure 3.4. Soissons (0-3 % DS) and Robigus (70-100 % DS) were the most resistant and susceptible parents respectively, regardless of year, site and scoring stage. The remaining parents varied in ranking depending on year and location, with variation generally occurring within two main clusters. A lower relatively resistant cluster, cluster 1, varying between 1 and 55 % DS, consisted of Alchemy, Hereward and Xi19. A higher relatively susceptible cluster, cluster 2, varying between 13 and 85 % DS, consisted of Brompton, Claire and Rialto. This was most evident from S2 onwards. Phenotypic variation was noticeable for a subset of parents, as illustrated by the large error bars in Figure 3.3 and Figure 3.4. This was the case for Alchemy, Brompton and to a lesser extent Rialto in 2015. A different subset of parents was affected in 2016, namely Robigus and Xi19. Except for Robigus, the most resistant and susceptible parents were unaffected by phenotypic variation.
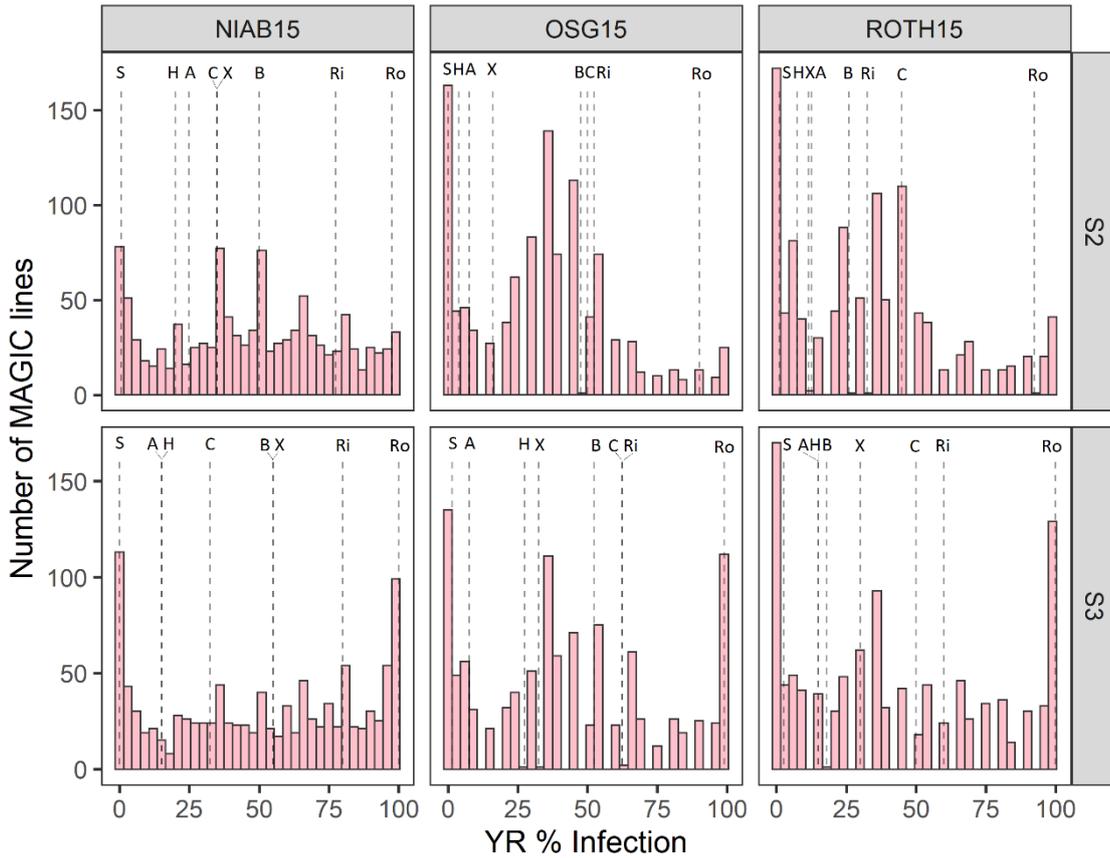
Figure 3.1 YR score distribution for the 2015 pathology trials.

NIAB15 histograms for scores 2 and 3 are based on averages (of two replicates), whereas for OSG15 and ROTH15, they are based on single replicates. NIAB15, OSG15 and ROTH15 represent the three trial sites from 2015. Dashed lines correspond to the average disease severity phenotypes of each of the MAGIC founders Alchemy (A), Brompton (B), Claire (C), Hereward (He), Rialto (Ri), Robigus (Ro), Soissons (So) and Xil9 (Xi).
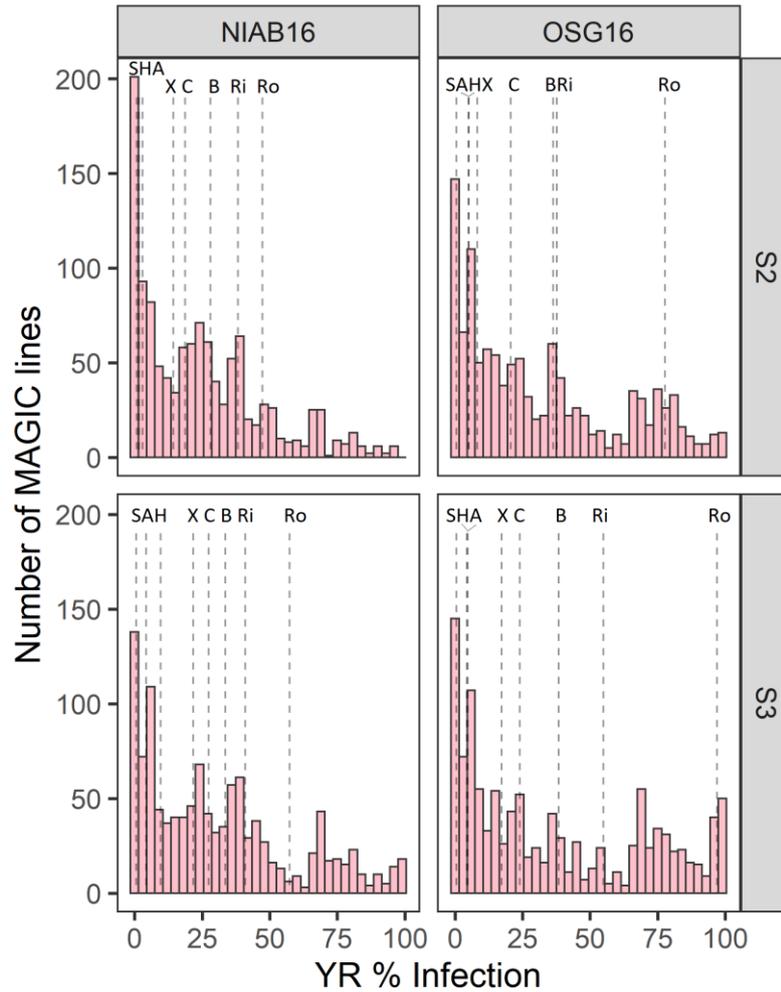
Figure 3.2 YR score distribution for the 2016 pathology trials.

The number of MAGIC lines includes single counts for unreplicated lines and total counts for replicated lines. NIAB16 and OSG16 represent the two trial sites from 2016. S2 and S3 correspond to the second and third YR assessment scores. Dashed lines correspond to the average disease severity phenotypes of each of the MAGIC founders Alchemy (A), Brompton (B), Claire (C), Hereward (He), Rialto (Ri), Robigus (Ro), Soissons (So) and Xil9 (Xi).
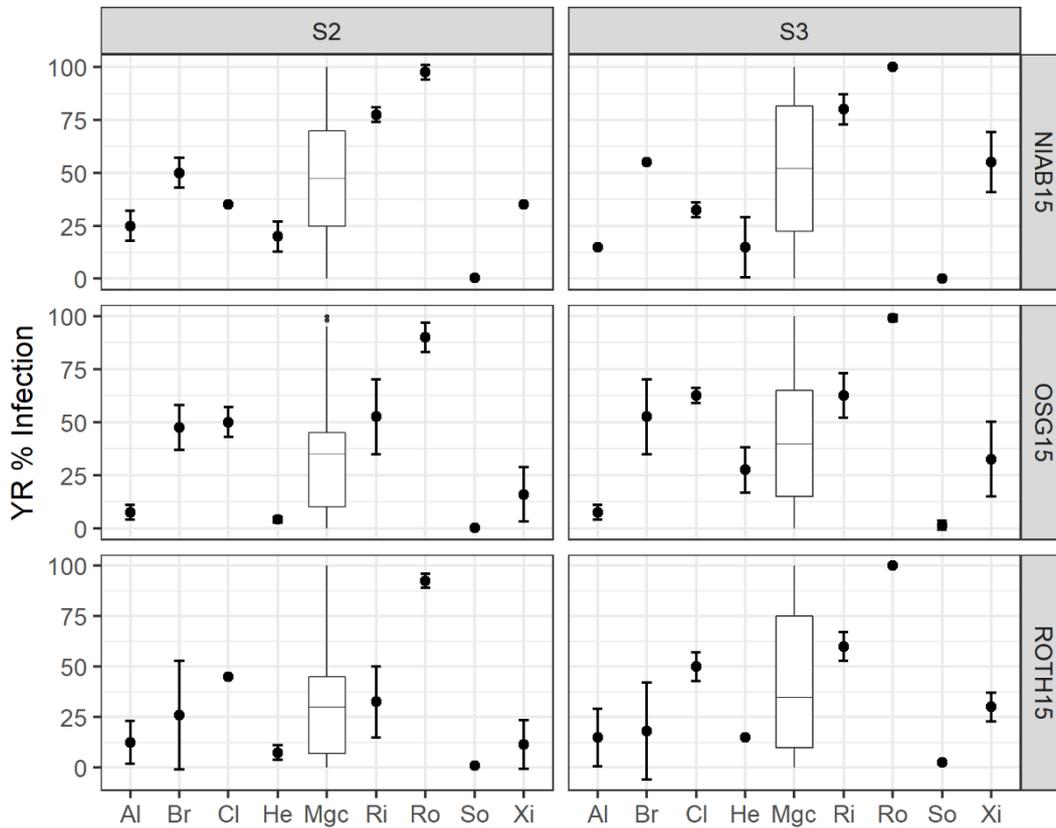
Figure 3.3 2015 Distribution of YR % infection for the MAGIC parents relative to the MAGIC population.

Mean % infection is represented by black dots for all parents, with standard deviation plotted as error bars. The central boxplot represents the distribution of the MAGIC population (mean values of 2 reps). Al: Alchemy, Br: Brompton, Cl: Claire, He: Hereward, Mgc: MAGIC lines, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19. S2 and S3 correspond to the second and third 2015 YR assessment scores. NIAB15, OSG15 and ROTH15 represent the three trial sites from 2015.
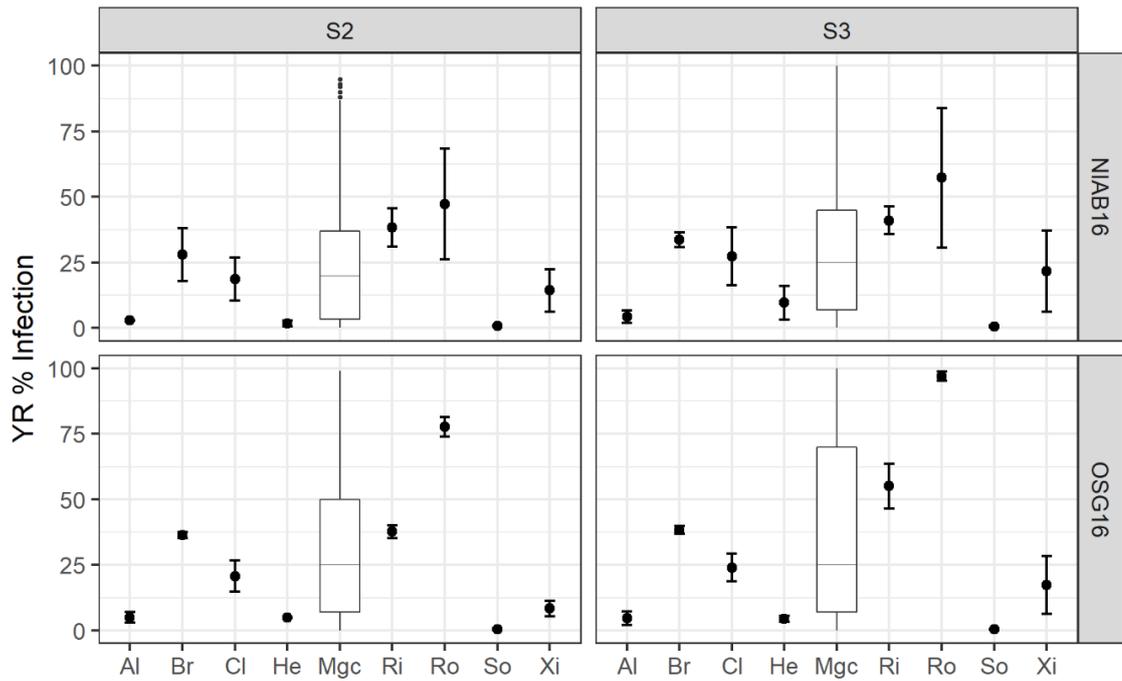
Figure 3.4 2016 Distribution of YR % infection for the MAGIC parents relative to the MAGIC population.

Mean % infection is represented by black dots for all parents, with standard deviation plotted as error bars. The central boxplot represents the distribution of the MAGIC population (total counts). Al: Alchemy, Br: Brompton, Cl: Claire, He: Hereward, Mgc: MAGIC lines, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19. S2 and S3 correspond to the second and third 2016 YR assessment scores. NIAB16 and OSG16 represent the two trial sites from 2016.

3.4.2 QTL analysis of yellow rust resistance in the MAGIC population

3.4.2.1 Comparison of the different QTL mapping methods

QTL mapping methods were compared to identify differences and similarities in the way they detect significant QTLs. The cluster of $-\log10(p)$ values observed when comparing Method 1 (Single Marker Analysis, SMA) vs Method 2 (Haplotype Analysis, HA), 3 (Interval Mapping, IM) and 4 (Composite Interval Mapping, CIM) indicates that in most cases, regression on allelic state is not as precise as the other methods in locating significant regions of interest (Figure 3.5). Methods 3 and 4, which were both based on interval mapping of FHPs, mostly correlated. A subset of $-\log10(p)$ values however, were significant with Method 3 but not with Method 4, suggesting that the addition of 10 covariate for interval mapping considerably reduces the number of significant markers in an interval. A similar pattern was observed for Method 2 vs Method 1.

Inspection of the underlying p values (adjusted for FDR) showed that composite interval mapping with 10 marker covariates (method 4) narrowed down a particular QTL interval to a small number of significant SNP markers (data not shown). While often narrowing down the QTL interval significantly, Method 4 appeared to generate a number of QTLs not detected by other QTL mapping methods. This was also occasionally the case for Method 3.
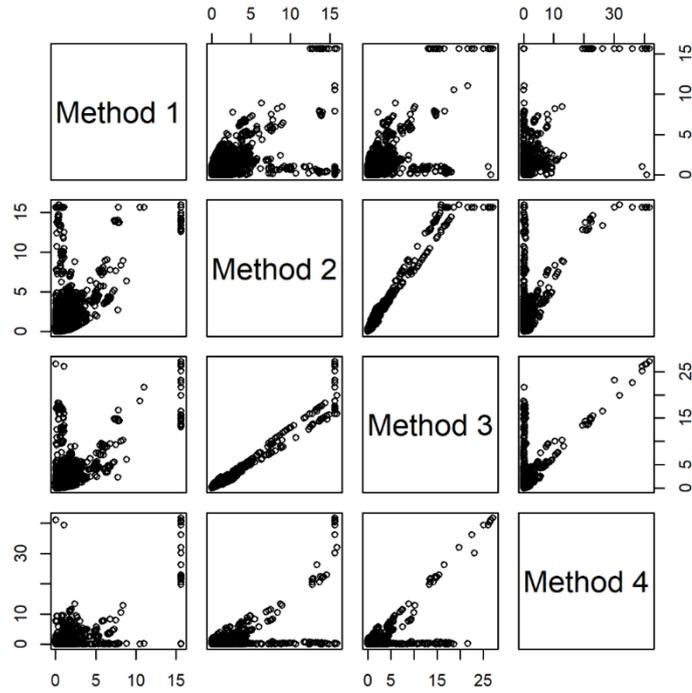
Figure 3.5 QTL mapping methods comparison for the final score (S3) of NIAB15.

$-\log10(p)$ values of the four different mapping methods used were plotted against each other. Method 1: Single Marker Analysis; Method 2: Haplotype Analysis; Method 3: Interval mapping (0 covariates); Method 4: Composite Interval Mapping (10 covariates).

3.4.2.2   Identification of QTL intervals and peak marker locations

Overall, nine QTLs were identified as significantly contributing to field YR resistance in the MAGIC population in 2015 and 2016 (Table 3.5, Table 3.6). Manhattan plots for each environment can be found in Appendix C (section 8.3). Five QTLs, *QYr.niab-1A.1*, *QYr.niab-2A.1*, *QYr.niab-2B.1*, *QYr.niab-2D.1* and *QYr.niab-3A.1* consistently and significantly contributed across all years, locations and disease scores, with intervals ranging between 10.8 and 20.8 cM (Table 3.6). Taken together, these five QTL explained approximately half of the overall phenotypic variation (Table 3.6). *QYr.niab-1A.1*, *QYr.niab-2A.1*, *QYr.niab-2B.1*, and *QYr.niab-2D.1* were identified with all four mapping methods, except for *QYr.niab-1A.1* at ROTH15 (Table 3.5, Table 3.6). QTLs on chromosomes 2B and 2D

represented the two major QTLs conferring resistance to YR in the MAGIC population, explaining 17.72 and 18.38 % of the phenotypic variation respectively (Table 3.5, Table 3.6). The corresponding peak markers identified in all environments and YR scores were Kukri_c9118_1774 and Ra_c21099_1781, located on the long arms of chromosomes 2B (271.92 cM) and 2D (197.36 cM) respectively. These two QTL were shortly followed in significance by *QYr.niab-1A.1* and *QYr.niab-2A.1*, (approximately maximum of 8 % PVE for both). The last of the five major QTL identified in this study was located on the short arm of 3A. Somewhat less consistent in its detection than the other QTLs, *QYr.niab-3A.1* was identified with all methods at NIAB15, OSG15 and OSG16, with SMA and HA at ROTH15 and with HA, IM and CIM at NIAB16 (Table 3.5).

In addition to the five QTL described above, a series of less significant QTL were also identified, with less consistency in detection across environments and mapping methods (*QYr.niab-2A.2, QYr.niab-6A.1, QYr.niab-6A.2, QYr.niab-6B.1*, $4.52E-10 < p < 4.41E-05$, 2.5<% PVE<4.0). Five additional QTLs fell just below the 5 % significance threshold: *QYr.niab-3D.1, QYr.niab-4B.1, QYr.niab-4D.1, QYr.niab-5A.1, QYr.niab-6A.3* ($9.19E-05 < p < 8.61E-04$, Table 3.7 and Table 3.8). Similar to *QYr.niab-2A.2, QYr.niab-6A.1, QYr.niab-6A.2* and *QYr.niab-6B.1, QYr.niab-3D.1, QYr.niab-4B.1, QYr.niab-4D.1* explained 2.9-4 % of PVE (Table 3.8). *QYr.niab-2A.2* and *QYr.niab-6A.2* were detected by SMA only for example. In other cases, QTL were detected in specific environments, such as *QYr.niab-3D.1* mostly only detected in OSG15 and *QYr.niab-6A.3* only detected in OSG16. There was an instance in which QTL identification differed depending on whether the analysis was based on non-transformed or log-transformed YR scores (Table 3.7). At NIAB16, *QYr.niab-5A.1* could only be detected when log-transformed values were used as phenotypic data. The interval size for these QTL was on average much smaller (<10 cM) than the major QTL described above. In the case of *QYr.niab-3D.1*, QTL interval size was reported as 0 because markers flanking the interval are co-segregating to the same locus (Table 3.8).

Table 3.5 Nine QTL conferred field resistance against YR in the MAGIC population in 2015-2016

QTL	Chr	Environments									
		NIAB15		OSG15		ROTH15		NIAB16		OSG16	
		S2	S3	S2	S3	S2	S3	S2	S3	S2	S3
QYr.niab-1A.1	1A					1	1				
QYr.niab-2A.1	2A										
QYr.niab-2A.2	2A	1	1	1	1	1	1	1	1		
QYr.niab-2B.1	2B										
QYr.niab-2D.1	2D										
QYr.niab-3A.1	3A					1-2	1-2	2-4	3		
QYr.niab-6A.1	6A	2-4	2-4	3-4	2-4			3-4	2-4	2-4	2-4
QYr.niab-6A.2	6A	1	1		1				1	2	1
QYr.niab-6B.1	6B		4			2-3	2-3				

QTLs were detected at 5 % significance threshold. See Appendix C for corresponding *p* values for all mapping methods and environments. S2-S3: number of disease assessments. QTL mapping methods are numbered 1-4: Single marker analysis (1), Haplotype Analysis (2), Interval mapping with 0 covariates (3), Composite Interval Mapping with 10 covariates (4). For 2016 trials, detected QTL are based on log-transformed data, unless indicated as *, in which case original disease scores were used. Empty grey cells: QTL was detected with all four mapping methods. Empty white cells: no QTL detected. QTL detected by method 4 only across all environments and with peak markers away from marker consensus by >10 cM were not included.

Table 3.6 Peak markers for the nine QTL conferring field resistance against YR in the MAGIC population in 2015-2016

QTL name	Chr	Genetic map position, cM	Peak marker	p values		% var explained		QTL interval, cM
				min	max	min	max	
QYr.niab-1A.1	1A	185.80	RAC875_rep_c71093_1070	5.72E-10	1.23E-05	7.46	9.26	17.17
QYr.niab-2A.1	2A	140.26	BS00022903_51	3.7E-13	7.29E-08	6.8	9.37	19.29
QYr.niab-2A.2	2A	259.39	BS00011599_51	1.75E-08	2.38E-06	-	-	2.56
QYr.niab-2B.1	2B	271.92	Kukri_c9118_1774	0*	1.58E-11	7.38	17.72	18.62
QYr.niab-2D.1	2D	197.36	Ra_c21099_1781	0*	8.62E-12	10.83	18.38	10.85
QYr.niab-3A.1	3A	3.02	Kukri_c28650_111	1.41E-07	5.58E-04†	3.48†	6.02	16.64
QYr.niab-6A.1	6A	55.51	BS00011010_51	4.52E-10†	1.22E-05	5.5	6.57†	18.24
QYr.niab-6A.2	6A	75.69	Kukri_c21743_269	5.15E-10	6.24E-06	-	-	2.05
QYr.niab-6B.1	6B	60.56	BS00068615_51	1.44E-07	4.41E-05†	2.49†	4.08	13.37

Genetic map position and QTL interval are from the MAGIC genetic map (Gardner et al. 2016). Minimum (min) and maximum (max) values for p value (shown as uncorrected for False Discovery Rates) and percentage variation explained (% var explained) independently selected amongst all environments and QTL mapping methods. † % var corresponds to indicated p value. * <2.22e-16. QTL interval based on QTL mapping method 4 (CIM, 10 covariates). A summary of the flanking markers and their genetic map position is summarised in Table 8.9 (Appendix C).

Table 3.7 Five borderline QTLs (< 5 % significance threshold) conferred field resistance against YR in the MAGIC population in 2015-16

QTL	Chr	Environments									
		NIAB15		OSG15		ROTH15		NIAB16		OSG16	
		S2	S3	S2	S3	S2	S3	S2	S3	S2	S3
QYr.niab-3D.1	3D			2	4	2					
QYr.niab-4B.1	4B							3-4		4	
QYr.niab-4D.1	4D									*	3*
QYr.niab-5A.1	5A	2*-4*	2*,4*					3*-4*	3*-4*		
QYr.niab-6A.3	6A									3	4

S2-S3: number of disease assessments. QTL mapping methods are numbered 1-4: Single marker analysis (1), Haplotype Analysis (2), Interval mapping 0 covariates (3), Composite Interval Mapping 10 covariates (4). For 2016 trials, detected QTL are based on log-transformed data, unless indicated as *, in which case untransformed disease scores were used. Empty grey cells: QTL was detected with all four mapping methods. Empty white cells: no QTL detected. QTL detected by method 4 only across all environments and with peak markers away from marker consensus by >10 cM were not included.

Table 3.8 Peak markers for the five borderline QTL conferring field resistance against YR in the MAGIC population in 2015-2016

QTL name	Chr	Genetic map position, cM	Peak marker	p values		% var explained		QTL interval, cM
				min	max	min	max	
QYr.niab-3D.1	3D	162.20	BS00004334_51	3.30E-05	8.61E-04†	3.35†	-	0
QYr.niab-4B.1	4B	50.66	Ra_c26080_461	3.72E-05†	-	-	2.97†	17.73
QYr.niab-4D.1	4D	125.78	D_GDRF1KQ02H66WD_341	1.07E-05†	5.58E-04	3.48	4.03†	26.56
QYr.niab-5A.1	5A	301.25	IAAV3916	9.19E-05	6.53E-04†	3.48	3.6†	13.32
QYr.niab-6A.3	6A	220.32	wsnp_Ex_rep_c101766_87073440	1.49E-04	1.37E-04†	3.03†	3.49	0.50

Genetic map position and QTL interval are from the MAGIC genetic map (Gardner et al. 2016). Minimum (min) and maximum (max) values for p value (shown as uncorrected for False Discovery Rates) and percentage variation explained (% var explained) independently selected amongst all environments and QTL mapping methods. † % var corresponds to indicated p value. * <2.22e-16. QTL interval based on QTL mapping method 4 (CIM, 10 covariates). A summary of the flanking markers and their genetic map position is summarised in Appendix C.

3.4.2.3   MAGIC founder contributions

The nine QTL characterised were further investigated to identify which MAGIC founders contributed to YR resistance for each. Founder estimates are outlined in

Table 3.9. Negative values represent a contribution towards YR resistance while positive values indicate a contribution towards YR susceptibility.

For the five major QTL identified in this study, *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B.1, QYr.niab-2D.1* and *QYr.niab-3A.1,* all parents except the most susceptible variety Robigus, contributed towards YR resistance. Founders were found to contribute either in combination or on their own, resulting in a variety of patterns to the contributions. Contribution in pairs: Claire and Hereward for *QYr.niab-1A.1,* Alchemy and Claire for *QYr.niab-2D.1,* Hereward and Rialto for *QYr.niab-3A.1,* Soissons and Xi19 for *QYr.niab-2A.1* and *QYr.niab-6A.1.* The 2D combination was expected, since Alchemy is the direct descendent of Claire in its pedigree so is highly likely to carry the 2D locus. Xi19 appears to be a contributor of resistance for a large number of QTLs: *QYr.niab-2A.1, QYr.niab-6A.1, QYr.niab-6A.2,* and *QYr.niab-6B.* Other examples of single contributors: Soissons for *QYr.niab-2B.1* and Brompton for *QYr.niab-6B.1.*

Examining YR scores individually across all environments provides additional insight into parent contribution (Section 8.4, Appendix C). This was particularly the case for QTLs on 2D. Although it is evident from all founder effect estimations that Alchemy and Claire are the overall contributing parents to the 2D QTL, differences were observed between different environments and between mapping methods.

Table 3.9 MAGIC founder contributions for seven of the nine QTL conferring field resistance against yellow rust in the MAGIC population in 2015-2016

| QTL | Chr | Founder effects | | | | | | | | Origin |
		Al	Br	Cl	He	Ri	Ro	So	Xi	env, YR score
QYr.niab-1A.1	1A	1	0.85	-0.5	-0.43	0.92	0.999	0.28	-0.2	OSG16, S2
QYr.niab-2A.1	2A	13.19	19.73	25.82	11.65	3.46	19.59	9.29	1	OSG15, S3
QYr.niab-2B.1	2B	1	0.12	2.05	1.16	1.16	0.87	-0.71	0.69	OSG16, S2
QYr.niab-2D.1	2D	-21.77	13.46	-16.7	-5.12	33.03	21.44	12.13	1	NIAB15, S3
QYr.niab-3A.1	3A	0.27	-0.22	-3	-16.13	-19.82	-5.2	-3.17	1	OSG16, S3
QYr.niab-6A.1	6A	1.93	1.73	1.67	1.52	1.98	1.93	1.47	1	OSG16, S2
QYr.niab-6B.1	6B	15.24	0.91	14.72	10.36	9.72	13.43	4.01	1	OSG15, S2

Founder effects presented here are for the consensus marker (outlined in Table 3.5) with the most significant *p* value identified with QTL mapping methods 2, (Haplotype Analysis), 3 (IM, cov=0) and/or 4 (CIM, cov=10). For QTL mapping method 2, contributions are relative to Alchemy whereas for methods 3 and 4, they are relative to Xi19. This was calibrated by adding an arbitrary value of 1 to all founder effects. For *QYr.niab-2B.1* and *QYr.niab-2D.1*, founder effects represented are from the peak marker with the highest % variation explained. For 2016 trials, detected QTL are based on log-transformed data, unless indicated as *, in which case original disease scores were used. QTL identified with SMA only (*QYr.niab-2A.2, QYr.niab-6A.2*) are not included.

3.4.3 Characterisation of pairwise QTL interactions

MAGIC lines were divided into 16 different genotypes based on the presence of the four most significant ($p<0.0001$, PVE>8 %) YR QTLs *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B.1* and *QYr.niab-2D.1* (Figure 3.6). For each QTL, the resistant allele was based on the allelic state of the consensus peak markers for each MAGIC founder. A one-way ANOVA identified that there were significant differences in the response to YR disease severity between all the QTL combinations (p <2.2E-16, Table 3.10). This analysis was followed by a pairwise *t* test for the comparison of % DS means (Section 8.5, Appendix C).

Table 3.10 Analysis of Variance in QTL interactions in response to YR rust disease severity

	SS	df	MS	*F*	*p*
Interactions	27177.9	15	1811.86	47.245	2.20E-16
Residuals	2377.7	62	38.35		

SS: Sum of Squares; df: degrees of freedom; MS: Mean Squares

All three-way QTL combinations were among the most effective in significantly reducing YR infection in MAGIC lines and displayed high resistance responses to YR infection (*t* test values<2E-16, Section 8.5, Appendix C). Of note was the *QYr.niab-1A.1, QYr.niab-2B.1, QYr.niab-2D.1* combination, which conferred near immunity across all trial sites. *QYr.niab-2B.1* and *QYr.niab-2D.1* showed significant levels in fungal growth reduction as effective as the three-way combinations when present in combination with *QYr.niab-1A.1* and *QYr.niab.2A.1* respectively (*t* test values<2E-16). The only QTL to confer resistance at this level of significance and in a single fashion was *QYr.niab.2B.1* (*t* test values<2E-16). All other combinations were also significant in reducing disease severity (4.8E-16<*t* test values>2.4E-15). Single QTL still provided significant levels of resistance against YR but were less effective in doing so when compared to combinations (*QYr.niab-2D.1: t* test

value=1.3E-13, *QYr.1A-1*: *t* test values=4.3E-10 and *QYr.niab-2A.1: t* test values=1.8E-05). The combination of *QYr.niab-1A.1* and *QYr.niab-2B.1* was the only one that appeared to have no significant effect in reducing disease severity (*t* test values=0.04).

Year and site effects observed for the QTL combinations *QYr.niab-1A.1* & *QYr.niab-2A.1* and *QYr.niab-1A.1* & *QYr.niab-2B.1*. In both cases, the 2016 trial sites exhibited a narrower range of disease severity response compared to the 2015 trial sites. Similarly, the *QYr.niab-2A.1* & *QYr.niab-2B.1* combination appeared to be less effective in conferring YR resistance in OSG15 compared to the other trial sites. Furthermore, most QTL genotypes displayed YR susceptible outliers. This likely indicated that the resistant allele allocation for some QTL genotypes may not have been as effective at capturing all MAGIC lines with that particular genotype.

The 82 resistant MAGIC lines (% DS<10) resistant across all trial sites exhibited all QTL genotypes apart from the four-way combination (1A, 2A, 2B, 2D) and *QYr.niab-1A.1* present in isolation. YR resistance in over half of the MAGIC lines (51.2%) was underpinned by the two-way combinations. Nearly a quarter of lines (24.4%) exhibited *QYr.niab-2A.1, QYr.niab-2B.1* and *QYr.niab-2D.1* in isolation. The three-way combinations conferred resistant in the least percentage of MAGIC lines (18.3%).
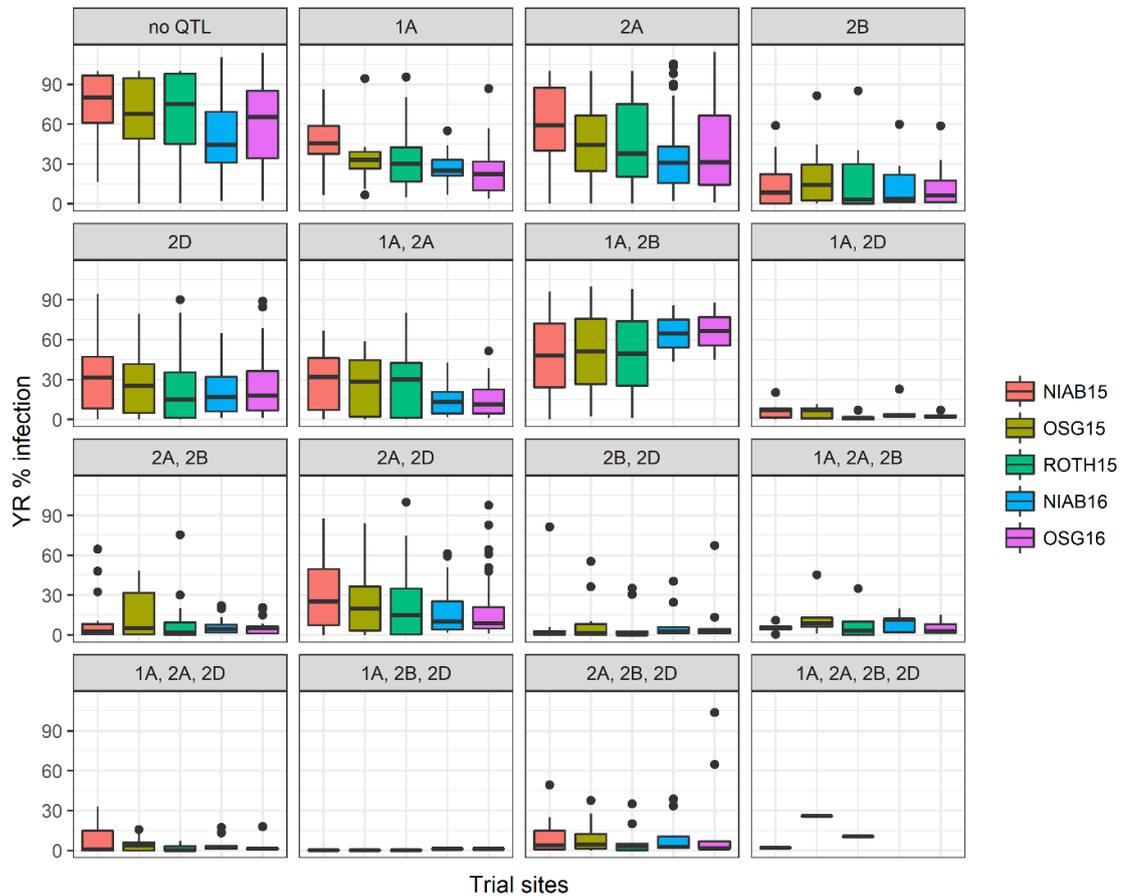
Figure 3.6 Distribution of YR % infection in MAGIC lines according to the four most main YR resistance QTLs.

Each panel represents a different QTL genotype, of which there are 16 in total. 1A: *QYr.niab-1A.1*, 2A: *QYr.niab-2A.1*, 2B: *QYr.niab-2B.1*, 2D: *QYr.niab.2D.1*. For each box plot, the box is outlined by the lower and upper quartiles and the thick line inside the box corresponds to the median. Lines on either side of the box represent the first and fourth quartiles. Outliers are presented by black dots.

3.5 Discussion

The NIAB Elite MAGIC population has eight founders, each with varying levels of YR resistance. Pathology trials undertaken in 2015 and 2016 provided an opportunity to investigate the response of the MAGIC population to *Pst* in detail, and thus accurately characterise QTLs conferring YR resistance in the field.

The MAGIC population exhibited a range of responses to *Pst* infection that were highly heritable (broad sense heritability = 0.77-0.94 %) in all trials and locations. Transgressive segregation was detected, with a subset of 20 MAGIC lines showing no signs of YR infection across environments (site x year). Approximately 10-13% of the MAGIC population remained highly resistant (<10 % DS) throughout the infection season, in both 2015 and 2016.

3.5.1 Phenotypic distribution of YR resistance in MAGIC pathology trials in 2015-2016

The distribution of phenotypic scores followed a bell-shaped curve in 2015 whereas in 2016, the phenotypic scores distribution was skewed towards resistance. While his skewness may in part be due to the slight earliness of the scoring in 2016, it is also likely that differences in disease pressure played a role. Indeed, studies have shown YR resistance at the adult stage to be influenced by varying levels of inoculum pressure (Chen, 2005, 2013; McIntosh et al., 1995). The disease severity scores of Robigus, the most susceptible founder, were higher on average in the 2015 trials than in the 2016 trials. In addition, of the genotyped lines present across all environments, 92 were highly resistant to YR in 2015 (% DS <10), a number that increased to 134 in 2016. Taken together, these observations suggest the presence of variable levels of disease pressure within and across years, which in turn, could have influenced the response of the MAGIC population to *Pst*.

Interestingly, these findings do not match the yellow rust pressure levels observed throughout the UK in 2015 and 2016. The number of *Pst*-infected leaf samples received by the UKCPVS increased by 81% between 2015 and 2016 (A. Hubbard, Pritchard, & Holdgate, 2016; A. Hubbard et al., 2017). In 2016, growers and agronomists reported early signs of YR, which is reflected by the reception of leaf-

infected samples as early as January by the UKCPVS (A. Hubbard et al., 2017). Furthermore, it is also worth noting that novel virulence combinations were detected in the same period, which led to a reduction in YR disease resistance ratings for some of the varieties on the Recommended List (A. Hubbard et al., 2017). Novel races included race 'Red 24' (virulent on *Yr1, 2, 3, 4, 6, 7, 9, 17, 2 5, 32, Re, Sp, Ro, So, Wa, Ca, St, Ap*) and Pink 14 (virulent on *Yr1, 2, 3, 4, 6, 9, 17, 25, 32, Re, Sp, Ro, So, Wa, Ca, St, Kr, Cr*). These novel virulence combinations are an additional factor that may have affected how the MAGIC population responded to YR infection.

Two scenarios are possible here. On the one hand, it is possible that the 2016 race change was not prevalent in this study's pathology trials and hence the MAGIC population would not have been exposed to it. On the other hand, if a race change was prevalent, the YR QTLs present in the MAGIC population were effective against those novel virulence combinations. Since pathology trials were very likely exposed to *Pst* by natural infection, inferring which of these hypothesis holds true is difficult to do without the virulence profile of the *Pst* in question. Nevertheless, the fact that nearly 100 lines remained highly resistant to yellow rust across all environments (site x year) somewhat points towards the former. There were only two instances of a change in resistance levels among MAGIC lines, from <10 %IF in 2015 to 40-67 %IF in 2016.

3.5.2 Characterisation of QTL conferring YR resistance

Four YR QTLs, *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B.1* and *QYr.niab-2D.1,* accounted for 8, 7, 16 and 18 % of variation on average, respectively. These were the most consistent of the QTL identified, detected in at least two scores in all pathology trials and with all genetic mapping methods (*QYr.niab-1A.1* detected with SMA approach only in all scores at ROTH15). QTLs *QYr.niab-2B.1* and *QYr.niab-2D.1* were the most consistent loci identified, as the same SNP was identified as the peak marker in at least two scores in all environments This is in contrast to *QYr.niab-1A.1* and *QYr.niab-2A.1,* for which the respective peak marker varied in location between two closely located loci, despite detection of this QTL in all environments tested. In addition to these four major QTL, several minor QTL

were detected on 3A, 6A and 6B, each with a percentage variation explained ranging between 3-8 %. Five QTL each locating on 5A, 3D, 4D, 5B and 4B, were found to fall short of the significance threshold. This section focuses on the four most consistently identified QTLs across scores and environments, in the context of previously reported *Yr* genes and YR resistance QTLs.

QYr.niab-2D.1 located at the distal end of the long arm of chromosome 2D and mapped to one of four YR resistance loci described by Chen & Kang (2017). This region comprises two YR resistance QTLs, *QYr.niab-2D.2* (Powell et al., 2013) and *QPst.jic-2D* (Jagger et al., 2011), and *QYr.tam-2D*, now catalogued as *Yr54* (Basnet et al., 2014). *QYr.niab-2D.2* and *QPst.jic-2D* both mapped to the same genomic location on 2D (Powell et al., 2013), while *Yr54* and *QPst.jic-2D* share *Xgwm301* as an interval-defining marker. Markers closely linked to *QPst.jic-2D* and *QYr.niab-2D.2* were located in close proximity (<4 Mbp) to *QYr.niab-2D.1*. Claire is the resistant allele contributor for *QYr.niab-2D.2* (Powell et al., 2013). This QTL was inconsistently expressed across environments, in contrast with *QYr.niab-2D.1* in the MAGIC population. However, it explained 13.7-31.8 % of the phenotypic variation explained, bearing slightly closer resemblance to those levels observed for *QYr.niab-2D.1*. Furthermore, the first score of NIAB16 (data not shown) indicates that *QYr.niab-2D.1* is expressed at an early stage in adult plant development. This aligns with observations made by Powell et al. (2013), who reported the detection of QTLs on 2DL as early as the tillering stage. Unlike the latter Claire QTL, *QYr.jic-2D* explained 30-36 % phenotypic variation and was consistently detected (Jagger et al., 2011). The right flanking marker of *Yr54*, *wPt-667054*, located within the *QYr.niab-2D.1* physical interval. Despite this similarity in genomic location, there are differences in the extent of explained phenotypic variation between *Yr54* and *QYr.niab-2D.1*. *Yr54* explained a much higher percentage of the phenotypic variation in the Avocet x Quaiu mapping population (49-54 %, Phenotypic Variation Explained, PVE) compared to *QYr.niab.2D.1* in the MAGIC population (up to 18.3 % PVE). Taken together, evidence suggests that *QYr.niab-2D.1* maps to a previously characterised interval on 2DL. Nevertheless, the relationship between Yr54 and *QYr.niab-2D.1* requires further exploration, notably by inferring any ancestral relationship between Alcedo, Claire and

Alchemy. The other QTL reported by Powel et al. (2013), *QYr.niab-2D.1*, which mapped to the same region as *Yr16*, was not detected in the MAGIC population. The Claire x Lemhi population was assessed for YR resistance prior to 2011 Warrior race incursion that led to a drastic genetic shift in the *Pst* population. It is therefore likely that *QYr.niab-2D.1/Yr16* is no longer effective against the post-2011 *Pst* population virulence profile, which would explain why it was not detected in the MAGIC population.

QYr.niab-2B.1 closely followed *QYr.niab-2D.1* in explaining phenotypic variation in YR (up to 17 % PVE) and was detected consistently across all environments (year x site). *QYr.niab-2B.1* located to chromosome 2B, a chromosome particularly rich in YR resistance QTLs, distributed over four to five main genomic regions (Chen & Kang, 2017; Rosewarne et al., 2013). *QYr.niab-2B.1* specifically mapped to the long arm of 2B, in the upper region closest to the centromere. According to Chen & Kang (2017), this region comprises 16 QTLs and *Yr* genes altogether conferring HTAP, adult plant and all-stage resistance. Marker *wmc501* was linked to *QYR1*, a QTL from Camp Remy conferring adult plant resistance (Boukhatem et al., 2002) and located 11 Mbp away from *QYr.niab-2B.1*. It was the third closest QTL to *QYr.niab-2B.1*. The relationship between the other two nearby QTLs (*QYrdr.wgp-2BL* and *QYrns.orz-2BL*) and *QYr.niab-2B.1* could not be established. Three all-stage and closely linked resistance genes, *Yr5*, *Yr7* and *YrSP* also cluster in this YR resistance-rich region and share *wmc501* as a linked marker (Feng et al., 2015; Marchal et al., 2018; P. Zhang et al., 2009). The close linkage of these <u>*Yr*</u> genes to *QYR1* through marker *wmc501* and its distance from *QYr.niab-2B.1* suggests that *QYr.niab-2B.1* likely does not coincide with these previously reported YR resistance loci.

Compared to the chromosomes 2B and 2D, less YR QTLs and *Yr* genes have been detected on 1A. Chen & Kang (2017) reported a total of 13 QTL and two temporarily catalogued *Yr* resistance genes. One QTL in particular, *QYr.wsu-1A.1*, physically located within the *QYr.niab-1A.1* interval. A GWAS study of the National Small Grains Collection winter wheat germplasm collection identified *QYr.wsu-1A.1*, a QTL that conferred resistance against YR in all tested environments and explained 4.04 % of the phenotypic variation in disease severity across all trials (Bulli et al.,

2016). IWA5505, the SNP associated with this QTL located within the physical interval of *QYr.niab-1A.1*. The germplasm collection used in Bulli et al. (2016) consisted of 588 Europeans accessions, which could potentially represent the same source of resistance as that observed in the MAGIC population. However, this is unlikely, based on the population structure of the GWAS germplasm collection. The favourable allele IWA5505 was present in the subpopulation group 2, a group that represents 21 accessions only within Europe. The majority of these accessions originate from Eastern Europe, with only six from Western Europe, of which one is from the UK. Accessions are coded and cultivar names could not be retrieved. While it cannot be excluded that some of the European accessions in the Bulli et al. (2016) GWAS study may be related to *QYr.niab-1A.1* YR resistance contributor, it is unlikely to be the case.

Similar to the 1A chromosome, 2A harbours less YR resistance loci compared to 2B and 2D, with the exception of the telomeric region on the short arm, known to carry several *Yr* genes, including the widely deployed *Yr17* in North Western Europe (X. Chen & Kang, 2017). *QYr.niab-2A.1* appears to coincide with *Yrxy2,* a gene conferring HTAP resistance and characterised in the Chinese winter wheat cultivar Xiaoyan 54 (Zhou et al., 2011). One of its flanking markers, Barc5, locates within the *QYr.niab-2A.1* physical interval. Cultivar Xiaoyan 54 is likely to be unrelated to any of the MAGIC parents and it is therefore unlikely *Yrxy2* and *QYr.niab-2A.1* represent the same source of YR resistance. This evidence suggests that *QYr.niab-2A.1* potentially represents a novel source of YR resistance.

In addition to these four QTLs detected consistently across environments and collectively explaining nearly 50 % of the phenotypic variation, a series of small effect QTLs were also identified (*QYr.niab-2A.2, QYr.niab-3A.1, QYr.niab-3D.1, QYr.niab-4B.1, QYr.niab-4D.1, QYr.niab-5A.1, QYr.niab-6A.1, QYr.niab-6A.2, QYr.niab-6A.3, QYr.niab.6B.1*). These cumulatively explained 20-25 % of the phenotypic variation in yellow rust resistance. Except for *QYr.niab-2A.2* (absent in OSG16) and *QYr.niab-3A.1,* which were both consistently detected, the majority were detected in a subset of environments. While a year and/or site effect are unclear, there is the possibility that some may be race-specific, given the emergence of novel virulence combinations in *Pst* races detected in 2016. This

could particularly be the case for *QYr.niab-3D.1* and *QYr.niab-6B.1*, both detected in 2015 only, but also *QYr.niab-4B.1*, *QYr.niab-4D.1* and *QYr.niab-6A.3*, appearing in 2016 only. QTL mapping studies on yellow rust resistance often derive phenotypic data from naturally infected trials (Bulli et al., 2016; Maccaferri et al., 2015). Nevertheless, the potential impact of *Pst* population shifts between years of testing is not something that is usually addressed. It would thus deserve more attention, in order to better understand QTL detection patterns between environments. Repeating the QTL mapping study with the four major QTLs as co-factors, or with a subtracted dataset of MAGIC lines omitting highly resistant lines would be a first step towards better detecting these minor-effect QTLs. Considering that approximately 30 % of the phenotypic variation remains unexplained, the method could potentially shed some light on undetected QTLs. Basnet et al. (2014) reported higher PVE values and the detection of two additional minor-effect QTLs by omitting the subset of RILs harbouring the major effect *QYr.tam-2D*.

The variety of yellow rust resistance QTL observed in the MAGIC population practically spanned all wheat chromosomes. Disease severity distribution under different QTL combinations and variation in disease severity response suggest that quantitative and qualitative resistance underpin yellow rust resistance in the MAGIC population.

3.5.3 Effect of QTL interactions on yellow rust resistance

The classification of MAGIC lines into different genotypes of the four most significant QTL (*QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B, QYr.niab-2D.1*) enabled the examination of the effect of different QTL combinations on disease severity. Combinations were more effective in conferring resistance to YR than QTL in isolation, a strong indication of interaction effects between QTLs in the MAGIC population. Additivity was observed in ten out of eleven combinations. Interactions between QTL and/or *Yr* genes are a key element of an effective breeding programmes. Additivity of minor effect QTL has successfully been exploited in the CIMMYT wheat breeding programme, demonstrating that near immunity can be achieved with this approach (Singh, Huerta-Espino, & Rajaram,

2000). The *Yr18* complex, made up of the APR resistance gene *Yr18* and up to four additional resistance genes, has been particularly successful in combatting yellow rust (Singh et al., 2005). In this study, dual combinations were sufficient in maximising resistance to 15 % disease severity or less, a good example of which is the 1A-2D combination. Exploiting dual combinations of yellow rust resistance loci would be more desirable to breeders, rather than attempting to combine three or more QTL, an option likely to be more expensive. The 1A-2B-2D combination proved particularly effective as it conferred almost complete immunity against yellow rust (0.05-1.1 % DS across environments). In contrast, the only instance where additivity was not observed was with the 1A-2B combination. Similarly, it did not confer any significant reduction in disease severity based on the pairwise interaction statistical test. Only two MAGIC lines carried both the 1A and 2B resistant allele. One of them was highly resistant, as expected, but the other one was highly susceptible. These extremes are not discernible in the evenly distributed box plot, so the nature of the 1A-2B interaction cannot be sustained. Interestingly, while both MAGIC lines in question are indistinguishable based on the resistant allele at the four major QTLs, the susceptible line carries the *QYr.niab-3A.1* resistant allele. This example highlights the importance of carefully interpreting such data when the allocated resistance allele may not have accurately captured the phenotypic resistance observed in each instance. Since the small effect QTL were not considered in examining QTL interactions, the presence of resistant MAGIC lines in the 'no QTL' category was somewhat expected. Considering the 30 % phenotypic variation remaining unexplained, it is therefore possible that some small effect QTL have been unaccounted for in this study, which would explain the average disease severity score being at approximately 70 %. A number of susceptible outliers (60-100 % DS) were observed in some QTL combinations. The misallocation of the resistance allele, a potential error in phenotypic assessment or the lack of inclusion of additional QTL could be three potential causes of the presence of these outliers.

With an increasing focus on utilising yellow rust resistance genes and QTLs in combination via allele stacking, QTL mapping studies should not only focus on identifying QTLs but also the effect that they have on disease resistance when

combined. This study has identified a number of beneficial QTL combinations, for which the underlying mechanisms of interactions merit further exploration.

3.5.4 Future work

The genetic analysis conducted here was based on a multi-founder population estimated to capture around 80% of the SNP diversity in UK wheat. Accordingly, efficient marker assisted selection of the adult plant resistance QTLs identified will likely be of immediate interest for wheat breeders and researchers. A number of next steps can be envisaged:

(1) *Validation of QTL in independent experimental populations and in different genetic backgrounds*. Bi-parental populations from breeding programmes, known to carry one or more of the QTL characterised, would be a first step. This can be complemented by cross-referencing with existing YR resistance datasets from the WAGTAIL association mapping panel of around 480 lines, and the development of genetic markers identified as closely linked to the QTLs identified, for marker assisted selection within existing breeding pipelines.

(2) *Exploiting Heterogeneous Inbred Families*. The MAGIC RILs were genotyped at the F_5 stage, at which around 2% heterozygosity is expected. Therefore, of the 643 RILs genotyped, around 12 are expected to be heterozygous across any given QTL. Analysis of the genotype calls of the parental lines versus the selected RILs possessing regions of heterozygosity across the target QTL allows individual RILs carrying strongly contrasting alleles to be identified. By sampling sib F_5 seed and genotyping with co-dominant markers within the QTL interval, individual homozygous lines for contrasting alleles at the target QTL can be identified. Therefore, HIFs contrasting essentially just for allelic state across the target QTL can be generated in just one generation (Tuinstra, Ejeta, & Goldsbrough, 1997). These germplasm resources can form the basis for downstream studies, including precise evaluation of Mendelised QTL effect, and can also be crossed to generate recombinations across the target interval for further refinement of QTL location, and ultimately, map-based cloning of causal genes.

(3) *Exploration of additional QTL mapping strategies*. The recent surge in statistical methods for mapping QTLs in multiparent populations would benefit this study

and increase its robustness. Specifically, the Whole Genome Average Interval Mapping (WGAIM) approach utilises a one-stage approach and enables the characterisation of QTL by environment interactions (Verbyla et al., 2014). While the implementation of this statistical method was problematic at the time of analysis in this study (due to glitches in the code), it has now been troubleshot to a sufficient extent for it to be used to analyse phenotypic datasets generated in the NIAB Elite MAGIC population. The use of WGAIM is expected to result in more precise QTL mapping intervals, and so help downstream analyses and exploitation. Additional approaches and/or statistical packages for mapping and analysing QTL in multiparental populations have been developed, including a random-model approach (J. Wei & Xu, 2016), R/mppR (Garin et al., 2018) and R/qtl2 (Broman et al., 2019), which was recently updated to include QTL analysis of multiparental populations.

(4) *Phenotyping Infection Type.* It is worth noting that this study did not include the evaluation of infection type (IT), an assessment that categorises how a host responds to pathogen colonisation host response to infection and has frequently been used in YR QTL mapping studies (Roelfs et al., 1992). QTLs detected only with IT phenotypes have previously been reported (Boukhatem et al., 2002). It is therefore possible that, in this study, some QTLs may have been missed or that the minor effect QTLs could have been more efficiently detected. Evaluating experimental trials consisting of 1,000-2,000 lines at different locations required a significant investment in time, thus preventing the accurate assessment of IT in addition to percentage infection. Nevertheless, information on IT would be particularly useful in future fine-mapping or additive effects experiments by investigating HIFs or MAGIC lines with interesting QTL combinations.

(5) *Exploring slow rusting characteristics.* Although not explored further in this study, slow rusting in the MAGIC population merits further exploration in controlled environment experiments. In a first instance, AUDPC values could be calculated and used to investigate disease progression and further experiments under controlled conditions would enable the characterisation of latent infection, infection frequency and stripe length (Singh et al., 2005).

3.6 Conclusion

The diversification of the *Pst* population in the last decade is posing a challenge to wheat resistance breeding in the UK. I conducted replicated and partially replicated pathology trials to investigate the genetic basis of YR resistance in the breeding relevant NIAB Elite MAGIC population in five trials across three sites and two years. Nine YR adult plant resistance QTLs were identified, with four consistently detected across environments and explaining nearly 50 % of the phenotypic variation, and the other five explaining 15-20 % with inconsistent detection across environments. The most significant QTL, *QYr.niab-2D.1*, was localised to a 10.8 cM interval and explained 18.9 % of the phenotypic variation. There was a strong indication of additivity effects between the four strong-effect QTL. Overall, this study demonstrates the benefits of using a multi-founder population to dissect quantitative disease resistance traits effectively, providing a rapid basis for translating research outcomes into YR resistance breeding programmes.

4 GENOME ANNOTATIONS TO ACCELERATE CANDIDATE GENE CHARACTERISATION

To further understand the molecular processes underpinning the different types of yellow rust resistance mechanisms, it is important to characterise the functionality of *Yr* genes. At present, only a small fraction of Yellow rust resistance (*Yr*) genes have been cloned, from the hundreds of yellow rust resistance loci that have been published. The recently published IWGSC RefSeq v1.0 genome assembly, along with its annotations, provide the opportunity to mine those loci further and identify candidate genes. With this purpose, I further examine the eight principal yellow rust resistance QTL identified in Chapter 3 (*QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2A.2, QYr.niab-2B.1, QYr.niab-2D.1, QYr.niab-3A.1, QYr.niab-6A.1, QYr.niab-6A.2*).

Five QTL were characterised by NLR (Nucleotide-Binding Site Leucine-Rich Repeats) clusters within their intervals (*QYr.niab-2B.1, QYr.niab-2D.1, QYr.niab-3A.1* and *QYr.niab-6A.2*), or were in close proximity (*QYr.niab-2A.1*), indicating that these loci are likely to confer major, race-specific resistance against yellow rust. A subset of NBS-LRR candidate genes were characterised by the presence of integrated domains. In contrast, three QTL (*QYr.niab-1A*.1, *QYr.niab-2A.1* and *QYr.niab-6A.1*) lacked NBS-LRR annotations completely, demonstrating that non-race specific Adult Plant Resistance (APR) may also be at play. Taken together, this study demonstrates how the latest genomic tools, like the IWGSC RefSeq v1.0 genome assembly, can be rapidly implemented to further inspect genetic loci. Functional annotations revealed a variety of putative genes as candidates with diverse functionalities, suggesting the role of both all-stage and adult plant resistance in the NIAB Elite MAGIC population.

4.1   Introduction

To date, 80 *Yr* genes have been characterised but only eight have been cloned. Cloned genes range from classic R genes encoding Nucleotide-Binding Site Leucine-rich Repeats (NLRs): *Yr5, YrSP, Yr7* (Marchal et al., 2018), *Yr10* (Liu et al., 2014) with a currently debated map position (Yuan et al., 2018), to genes encoding a variety of proteins, including an ABC transporter (*Yr18/Lr34*) (Krattinger et al., 2009), a hexose transporter (*Yr46/Lr67*) (Moore et al., 2015), a gene encoding a putative kinase-pseudokinase (*Yr15*) (Klymiuk et al., 2018), and a gene encoding a protein kinase and a START domain (*Yr36*) (Fu et al., 2009). These genes have predominantly been isolated, cloned and functionally characterised using map-based cloning approaches. More recently, novel cloning approaches combining gene targeted enrichment sequencing and mutational genomics have started to emerge, enabling the rapid cloning of disease resistance genes (Sánchez-Martín et al., 2016; Steuernagel et al., 2016). These approaches rely on sequence capture of specific targets and comparative genomics between loss-of-function mutants. The first of these approaches, termed MutRenSeq, was implemented to clone the stem rust resistance genes *Sr22* and *Sr45* (Steuernagel et al., 2016), and more recently for the cloning of *Yr5, Yr7* and *YrSP* (Marchal et al., 2018). What both the classical map-based and more recent cloning-by-sequencing approaches have in common is the need for high quality genome sequence information. Map-based cloning requires a physical map onto which segregating markers can be anchored, while the more recent cloning approaches rely on sequencing of reduced complexity genomic templates, captured using DNA baits designed from the reference genome sequence for example.

4.1.1   Wheat genome assemblies

Wheat is a crop whose genome has been notoriously difficult to sequence, primarily due to its large size (17 Gb), its polyploid structure and high levels of repetitive DNA elements. Historically, the lack of a wheat reference genome sequence was a major limiting factor in narrowing down QTLs to their causal genes and genetic variants. However, the increasing affordability of next-generation sequencing technologies and availability of bioinformatic tools and algorithms for assembling the resulting data have revolutionised wheat genomics and opened the

door to relatively rapid and inexpensive delimitation of physical regions and gene content within QTL intervals. Initial attempts to generate large scale wheat genome sequence resources using bacterial artificial chromosome (BAC) libraries, shotgun, chromosome and whole genome-based sequencing approaches, have provided valuable insight into the composition of the wheat genome (Brenchley et al., 2012; Chapman et al., 2015; Marcussen et al., 2014; Paux et al., 2008). However, the large number of genomic scaffolds generated meant that while the information was extremely useful, these assemblies were highly fragmented. The genome assemblies released in the past couple of years build on more efficient assembling algorithms, longer sequencing reads, paired end reads and improved library preparation methodologies, and are therefore of better quality (Clavijo et al., 2017; Zimin et al., 2017). The most recent wheat genome assembly was released in 2018: IWGSC RefSeq v1.0, from cultivar Chinese Spring 42 (Appels et al., 2018).

The above assemblies primarily focus on the reference hexaploid wheat variety Chinese Spring, thought to have been a selection from a Chinese Landrace (Liu et al., 2018). Chinese Spring has become the global reference for wheat genome sequence, largely due to its early use in the development of genetic stocks, such as the aneuploidy germplasm developed by Earnest Sears in the 1930s (Sears, 1954). More recently, with costs of genome sequencing reducing drastically and advances in genome assembly software and approaches, researchers are moving beyond reliance on a single reference towards a pan-genome era. For example, the 10+ Wheat Genomes Project (http://www.10wheatgenomes.com/) is an international consortium involved in re-sequencing 15 wheat varieties. Of these, genome sequences for five varieties, including the MAGIC founders Claire and Robigus, have already been made available through the Grassroots Genomics initiative (Bian, Tyrell, & Davey, 2017). Eight are available for BLASTn interrogation.

4.1.2 Gene annotations

Genome assemblies provide the backbone for functional annotations, for which gene models are the crucial starting point. Gene models are available for the CSS, TGAC and RefSeq v1.0 genome assemblies for Chinese Spring 42 and are based on *de novo* gene prediction, RNA-seq data, and sequence similarity with related

species gene models and/or proteins (Appels et al., 2018; Clavijo et al., 2017; Marcussen et al., 2014). A major limiting factor in gene model accuracy has been contig and scaffold length, since these are the physical contigs the gene models are based on. The fragmented nature of the earlier wheat genome assemblies elevated the risk of inaccuracy in the gene model predictions. With an increase in contig and scaffold length in the latest genome assemblies, gene model predictions have become more accurate. The most recent RefSeq v1.1 gene annotation consists of the largest number of gene models to date (269,328 genes), categorised into high (107,891) and low (161,537) confidence genes. This classification was based on levels of sequence homology between the putative protein encoding the gene model and proteins searchable on public databases. These annotations have been used to anchor molecular markers and genetic variants to the assembly to characterise candidate genes for important agronomic and disease resistance traits (Appels et al., 2018; Downie et al., 2018; Marchal et al., 2018). The potential of wheat functional gene annotations has been enhanced with additional genetic resources, such as transcription atlases for gene expression (Borrill, Ramirez-Gonzalez, & Uauy, 2016; Pearce et al., 2015; Ramírez-González et al., 2018) and Targeting Induced Local Lesions IN Genomes (TILLING) mutants for reverse genetics research (Krasileva et al., 2017).

4.2 Aims and overview

The IWGSC RefSeq v1.1 gene models provide a valuable resource for the rapid identification of candidate genes, aiding subsequent cloning of target genes. This chapter takes a deeper look at the physical intervals of the eight major YR resistance QTLs identified in Chapter 3, based on the IWGSC ordered physical sequence and annotated gene models. For each QTL, I start by systematically defining the physical regions that form the basis for subsequent investigation. This was followed by a description of the high confidence gene models extracted from the IWGSC RefSeq v1.1 gene annotations, and classification of the gene models according to predicted molecular function. Finally, this information was used to examine which annotated genes are likely to be involved in disease resistance pathways, and the results compared the literature.

4.3 Materials and methods

The selection of QTLs for further exploration was based on (i) consistency of detection across environment, (ii) significance levels, and (iii) consistency across mapping methods. Based on these criteria, this chapter thus describes the physical space for a subset of QTLs conferring YR resistance in the MAGIC population, as characterised in Chapter 3: *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2A.2, QYr.niab-2B.1, QYr.niab-2D.1, QYr.niab-3A.1, QYr.niab-6A.1* and *QYr.niab-6A.2*.

4.3.1 MAGIC map anchoring to the IWGSC physical map

SNP anchoring was carried out by Keith Gardner and Bruno Santos (unpublished). Briefly, SNPs from the MAGIC genetic map (Gardner et al., 2016) were aligned to the IWGSC RefSeq v1.0 annotated Chinese Spring cultivar (cv) genome assembly (Appels et al., 2018) using BLASTn (Altschul et al., 1990). Following the marker order from the genetic map, the average deviance of a marker from the physical map was used as an indicator for filtering markers, in a rolling window of 20 markers. A cut-off of 50 Mbp was used as a deviance threshold, above which markers were filtered out. Manual curation was necessary in areas with lower recombination rates such as the centromere and chromosome ends. Known introgressions were also considered for this purpose. In addition, the following criteria were also used:

- Markers resulting in hits on chromosomes different to the chromosome to which they were genetically mapped in the MAGIC map were filtered out.

- In case of matches within the same chromosome (<1,000 bases apart), the first hit was selected.

For each of the eight YR QTL investigated in this Chapter, I plotted the genetic map positions (cM) for markers on the relevant chromosome against their corresponding physical map position (bp), based on the anchoring described above. Data were plotted using R/ggplot2 (Wickham, 2009), and the peak marker positions for each QTL highlighted.

4.3.2 Delineation of the peak marker interval for the physical space description

Within the QTL interval based on the CIM results using 10 covariates, the peak marker was used as the starting point for each characterised QTL. Flanking markers were selected based on the physical map positions, following re-ordering of markers from each recombination bin onto the physical map. The process of flanking marker selection is outlined in Figure 4.1.
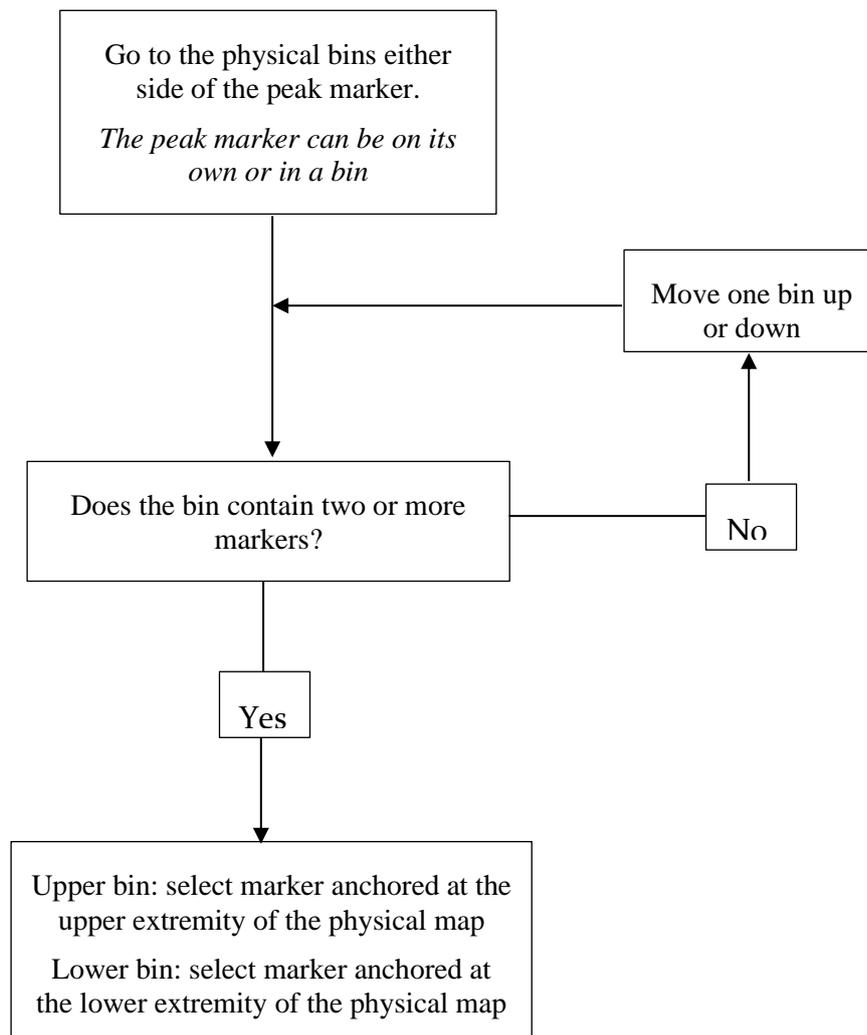
Figure 4.1 Process of selection for the flanking markers of each physical interval

4.3.3 Physical space description

Gene models within the physical intervals determined were based on the IWGSC RefSeq v1.1 gene annotations. These are further classified into high and low confidence genes. In this study, only the high confidence category was considered. IWGSC functional annotations were based on two independent gene prediction pipelines further supported with transcript evidence (PacBio transcripts and RNA-seq assemblies), as outlined in Appels et al. (2018). Protein domains and putative functions were identified using Pfam 31.0 (Finn et al., 2016), Interpro 69.0 (Finn et al., 2017) and GO terms (Ashburner et al., 2000). In addition to this publicly available information, I performed manual tBLASTn searches of the wheat genome, using protein sequences of cloned *Yr, Lr, Sr* and *Pm* resistance genes (Table 4.1) as queries (e-value cut-off used = E-10).

Table 4.1 Summary of *R* gene-encoding protein sequences used as queries for tBLASTn search against the *Triticum aestivum* genome.

Cloned resistance gene	NCBI protein accession number	Gene Functional annotation	Reference
Lr1	ABS29034	CC-NBS-LRR	Cloutier et al., 2007
Lr10	AAQ01784	CC-NBS-LRR	Feuillet et al., 2003
Lr21	ACO53397	NBS-LRR	Huang et al., 2003
Pm3b	AAQ96158	CC-NBS-LRR	Yahiaoui, Srichumpa, Dudler, & Keller, 2004
Pm8	AGY30894	CC-NBS-LRR	Hurni et al., 2013
Sr22	CUM44200	CC-NBS-LRR	Steuernagel et al., 2016
Sr33	AGQ17384	CC-NBS-LRR	Periyannan et al., 2013
Sr35	AGP75918	CC-NBS-LRR	Saintenac et al., 2013
Sr45	CUM44213	CC-NBS-LRR	Steuernagel et al., 2016
Sr50	ALO61074	CC-NBS-LRR	Mago et al., 2015
Yr46/Lr67	ALL26331	Hexose transporter	Moore et al., 2015
Yr18/Lr34	ACN41354	ABC transporter	Krattinger et al., 2009
Yr36	ACF33187	Kinase-START	Fu et al., 2009

A gene was considered as a candidate when its associated functional annotations and protein domain descriptions based on Pfam and Interpro terms exhibited one of the following:

- Explicit mention of disease resistance or that of known cloned *Yr, Lr, Sr* and *Pm* resistance genes in the functional description

- NB-ARC and leucine-rich repeat protein domains

- Sugar transporter

- ABC transporter

- Kinase START domain

4.3.4 Visualisation of QTL location in relation to total and NBS-LRR gene content of the reference wheat genome assembly

A type of circular diagram termed 'Circos plot' (Krzywinski et al., 2009) was used to visualise the physical position of the eight QTL in relation to the total and NBS-LRR gene content of the 21 wheat chromosomes using the IWGSC ReSeq v1.1 wheat gene annotation dataset (Appels et al., 2018). The gene annotations used consisted of those categorised as 'High Confidence' only. QTL flanking markers described in Table 5.2 were used to delineate each QTL on the diagram. The NBS-LRR gene content consisted of all gene models annotated as 'NBS-LRR, 'NB-LRR' and 'NB-ARC' and was generated using a bespoke python script developed by Larry Percival-Alwyn (NIAB). Total and NBS-LRR gene counts were generated every 10 Mb for each of the 21 wheat chromosomes. Finally, NBS-LRR gene density in relation to total gene content was calculated as NBS-LRR gene count/total gene count. This ratio, together with the two types of gene counts were plotted using Circos v0.69 (Krzywinski et al., 2009).

4.4 Results

4.4.1 Physical versus genetic intervals

Plotting physical versus genetic map positions for all SNPs genetically mapped to the target chromosomes revealed that all QTL peaks were located outside of the highly non-recombining chromosomal regions spanning the centromeres (Figure 4.2). *QYr.niab-2A.2, QYr.niab-2D.1* were located to the telomeric regions on the long arms of their respective chromosomes. In contrast, *QYr.niab-3A.1* (located via peak SNP Kukri_c28650_111) was situated towards the end of the short arm of chromosome 3A. Similarly, *QYr.niab-6A.1* and *QYr.niab-6A.2* were located on the short arm of chromosome 6A. Finally, *QYr.niab-2B.1* lies close to the start of the boundary between high and low recombining regions on the short arm of chromosome 2B.

Visual inspection of the physical map in relation to the genetic map also provides insight into SNP marker alignments and order. Clusters and other non-linear patterns indicate poor marker ordering on the genetic map, or structural rearrangements such as inversions or translocations. This is notably the case for *QYr.niab-1A.1,* which falls into a small cluster of markers whereas *QYr.niab-2A.2* falls into a local inversion.

The physical intervals (as defined here in section 4.3.2) for all QTLs ranged between 0.59-5.61 Mbp and averaged 3.04 Mbp (Table 4.2). Notably, the physical interval versus genetic interval ratios often differed. For example, *QYr.niab-6A.2* had the second largest genetic interval (16.66 cM) but the smallest physical interval (0.59 Mbp), equating to 0.04 Mbp/cM (Table 4.2). Similarly, the smallest genetic interval (0.50 cM; *QYr.niab-2A.1*) had the third largest physical interval (0.50 cM versus 4.28 Mbp = 8.51 Mbp/cM).
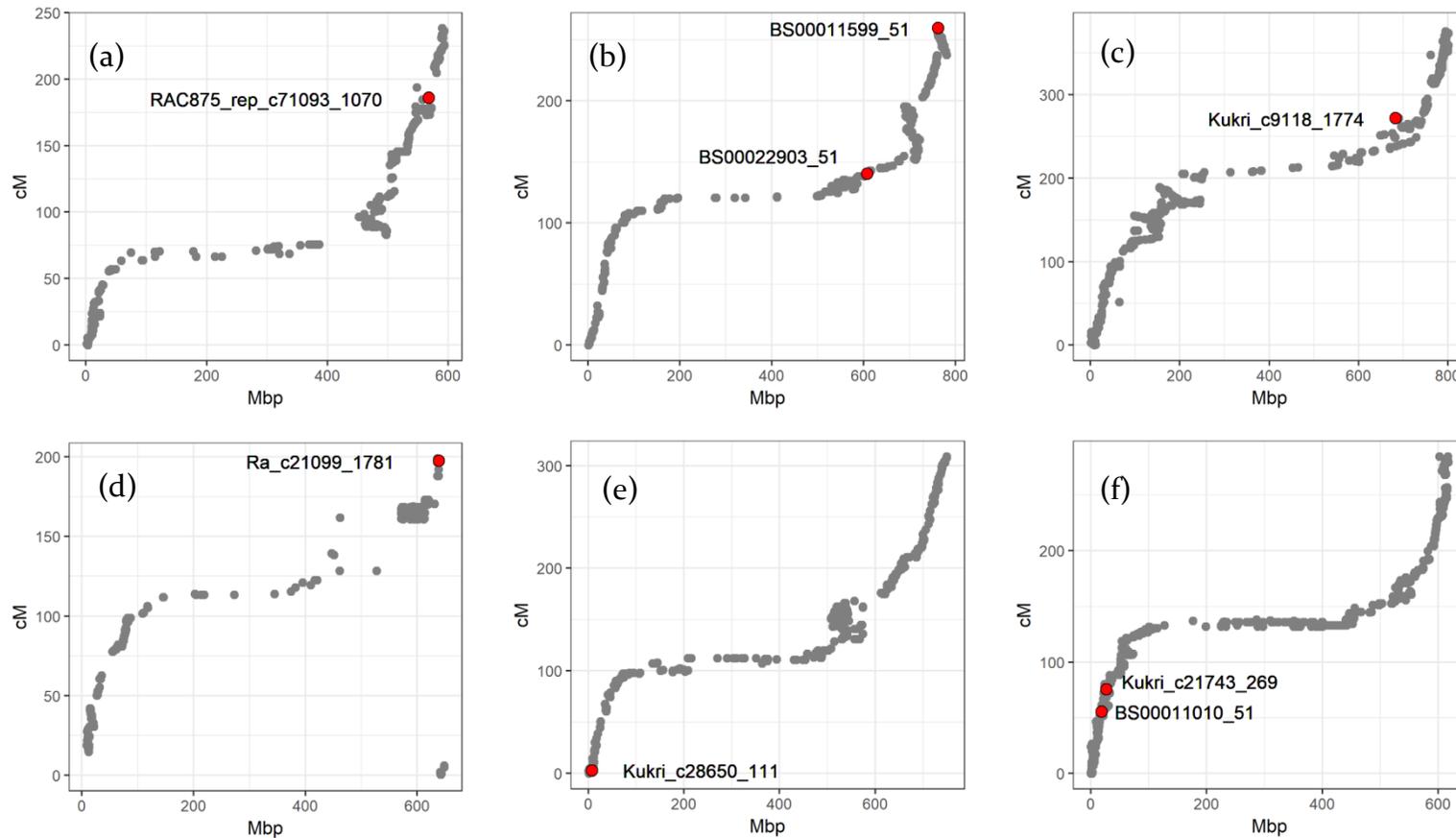
Figure 4.2 Location of the eight most significant YR QTLs on plots of physical versus genetic map SNP locations for chromosomes 1A (a), 2A (b), 2B (c), 2D (d), 3A (e) and 6A (f).

Peak markers are indicated in red and labelled on the figure. The physical map is plotted in Mbp while the genetic map is plotted in cM. Markers BS00022903_51 and BS00011599_51 in (b) correspond to *QYr.niab-2A.1* and *QYr.niab-2A.2* respectively. BS00011010_51 and Kukri_c21743_269 in (f) correspond to *QYr.niab-6A.1* and *QYr.niab-6A.2*, respectively.

Table 4.2 Summary of genetic and physical map intervals for the eight most significant QTL conferring YR resistance in the MAGIC population

QTL name	Chr	Peak marker	P (min)	Genetic map (cM)	Physical map start, end (Mbp)	Flanking markers left, right	Flanking marker genetic start, end (cM)	Total genetic start, end (cM)*	Total physical start, end (Mbp)	Total genetic interval (cM)	Total physical interval (Mbp)
QYr.niab-1A.1	1A	RAC875_rep_c71093_1070	7.7E-10	185.804	568.0128	CAP12_c6629_301	172.96	172.96	564.28	14.89	5.19
					568.0257	CAP12_c8163_118	173.46	187.84	569.47		
QYr.niab-2A.1	2A	BS00022903_51	2.4E-13	140.255	607.8272	Kukri_c24064_2095	140.76	140.76	606.83	0.50	4.28
					607.8273	BS00067276_51a	141.26	141.26	611.10		
QYr.niab-2A.2	2A	BS00011599_51	9.5E-10	259.386	762.2900	RAC875_c26809_453	255.82	252.80	762.29	6.58	3.43
					762.2901	Excalibur_c28885_363	256.32	259.39	765.71		
QYr.niab-2B.1	2B	Kukri_c9118_1774	0	271.917	683.0475	BobWhite_c18540_351	238.54	238.54	682.85	33.38	5.61
					683.0476	CAP8_rep_c8162_101	271.92	271.92	688.46		
QYr.niab-2D.1	2D	Ra_c21099_1781	0	197.357	638.3764	IACX9095	196.35	0.50	638.37	†	3.24
					638.3765	Tdurum_contig32203_281	0.50	197.36	641.61		
QYr.niab-3A.1	3A	Kukri_c28650_111	9.4E-10	3.0151	7.9207	Jagger_c6722_104	3.02	1.01	7.48	6.53	1.21
					7.9208	CAP12_c1860_280	7.54	7.54	8.69		
QYr.niab-6A.1	6A	BS00011010_51	3.9E-08	55.5076	18.7131	Excalibur_rep_c102994_1689	51.94	51.94	17.64	9.64	1.07
					18.7132	Tdurum_contig13287_203	56.51	61.57	18.71		
QYr.niab-6A.2	6A	Kukri_c21743_269	7.7E-07	75.6858	27.0718	Excalibur_c14222_179	68.61	59.03	26.93	16.66	0.59
					27.0719	CAP7_rep_c8019_110	75.69	75.69	27.52		

Intervals were determined for each QTL based on the anchoring of genetically mapped SNP markers on the physical map. *After the ordering of genetically mapped markers within QTL intervals based on the physical map, for some QTLs, the resulting genetic interval does not correspond to the genetic interval obtained using flanking markers alone. This is distinguished by the terms 'flanking' and 'total'. Chr: chromosome. *P* min represents the most significant *P* value ($P < 0.0001$) for each QTL, with 0 corresponding to $P<2.2E\text{-}16$. † Due to pronounced re-ordering between the genetic and physical map, an accurate genetic interval could not be established for *QYr.niab-2D.1*.

4.4.2 QTL location in relation to NLR gene models across the wheat genome

Plotting the genomic locations of the eight yellow rust resistance QTL in relation to NBS-LRR gene models from the IWGSC RefSeq v1.0 genome assembly revealed that all QTL located to regions of relatively high NBS-LRR gene content, which are also located within regions of high total gene content (Figure 4.3, tracks 4 and 5). These regions however were not the ones exhibiting the highest overall NBS-LRR gene density (quantified as NBS-LRR gene count over total gene count)– those were the telomeric regions of chromosomes 1B, 2B, 4A, 5B, 7B and 7D. Of the eight QTL examined here, *QYr.niab-2B.1* exhibited the highest NBS-LRR gene density (11.3 % of total genes per Mb are NBS-LRR), closely followed by *QYr.niab-2D.1* (10.8 %). The QTL exhibiting the lowest density was *QYr.niab-6A.1* (4.5 %).

4.4.3 Gene annotations and candidate genes within QTL intervals

The eight QTL were then examined in more detail for their gene content based on (i) IWGSC RefSeq v1.1 functional annotations, protein domain terms (Pfam, Interpro), and (ii) BLAST analysis (based on using predicted protein sequences of 13 known *Yr, Lr, Sr* and *Pm* resistance genes as queries for tBLASTn searches against the RefSeq v1.0 wheat reference genome. The number of candidate genes is outlined in Table 4.3. Further details for each candidate can be found in Table 4.5 (gene physical locations and functional annotations), and Appendix D (tBLASTn results).

The number of high confidence genes within QTL intervals varied between 15 and 75, with most QTL exhibiting a number ranging between 50 and 75 (Table 4.3). Three of the QTL intervals (*QYr.niab-2A.2, QYr.niab-6A.1, QYr.niab-6A.2*) exhibited a high number of genes relative to their physical size, as illustrated by gene densities ≥20 genes/Mbp (Table 4.3). However, a large physical size did not always translate into a high number of gene models in the region (Table 4.3), as was the case for *QYr.niab-2B.1* (lowest gene density of 9.6).
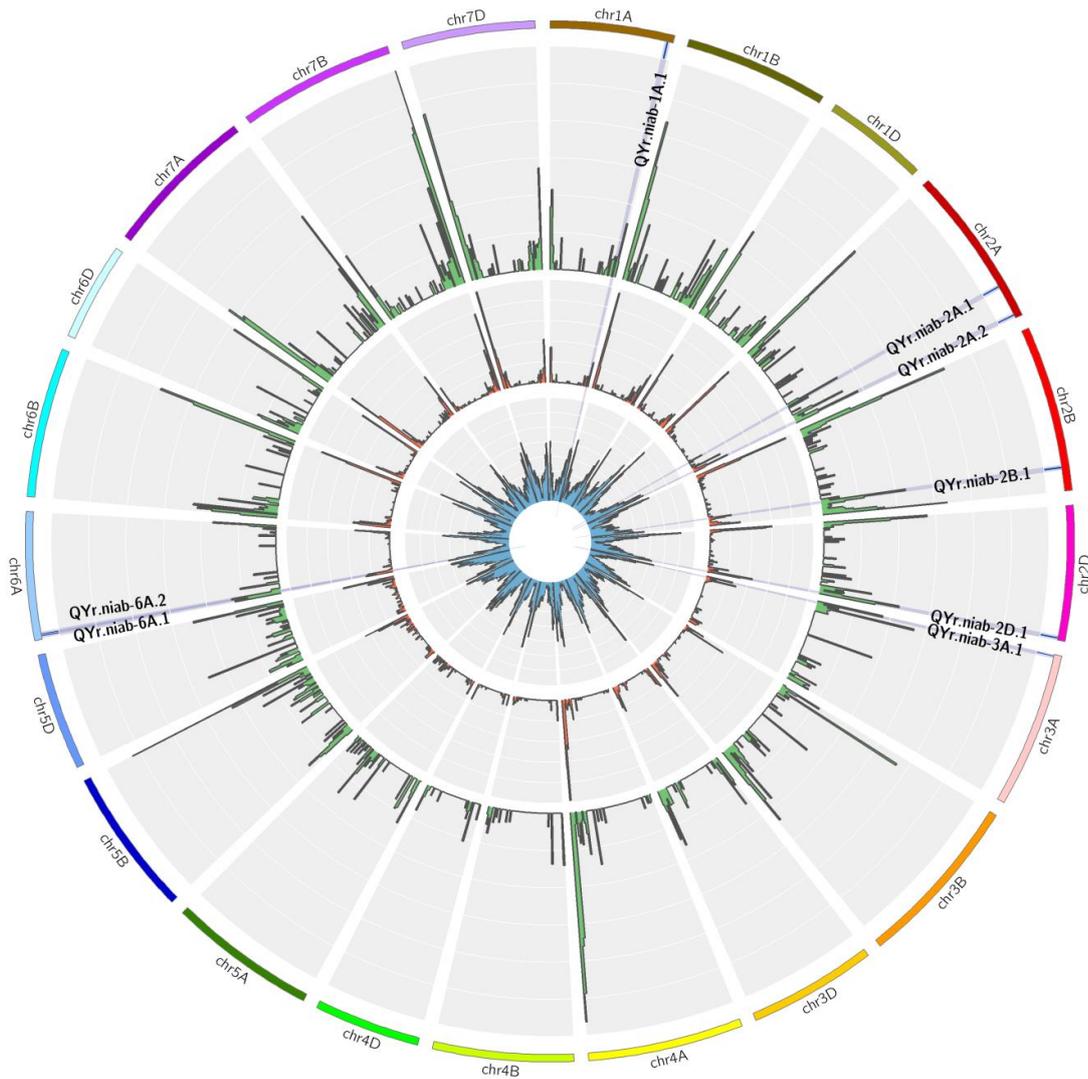
Figure 4.3 Circular diagram representing the position of yellow rust resistance QTL identified in the MAGIC population, in relation to the wheat cv. Chinese Spring reference genome sequence (IWGSC RefSeq v1.0) and gene annotation (RefSeq Annotation v1.1).

Starting from the outermost track; Track 1: chromosome name with different colours for each of the 21 chromosomes. Track 2: Position of QTL (in Mbp), as indicated by a blue tick, with grey bars going down the track, each labelled according to the QTL in question. Track 3 (green): NBS-LRR gene density as calculated by NBS-LRR gene model count over all gene models count (scale 0-0.5). Track 4 (red): NBS-LRR gene models count (scale 0-49). Track 5 (blue): All gene models count (scale 0-304). All counts per 10 Mb, with only High Confidence genes considered.

Table 4.3 Summary of IWGSC RefSeq v1.1 gene annotation metrics for the eight most significant QTL conferring YR resistance in the MAGIC population

QTL name	Chr	Total physical interval (Mbp)	Gene models	Gene density (genes/Mbp)	Average spacing between genes (Kb)	Genes with *NLR* gene motif[a]	Genes with non-*NLR* gene motif[b]
QYr.niab-1A.1	1A	5.19	68	13.1	73.3	0	3
QYr.niab-2A.1	2A	4.28	53	12.4	93.9	0	0
QYr.niab-2A.2	2A	3.43	75	21.9	43.8	12	0
QYr.niab-2B.1	2B	5.61	54	9.6	101.8	13	1
QYr.niab-2D.1	2D	3.24	64	19.8	46.9	15	0
QYr.niab-3A.1	3A	1.21	24	19.8	49.6	8	0
QYr.niab-6A.1	6A	1.07	22	20.5	47.5	1	0
QYr.niab-6A.2	6A	0.59	15	25.5	37.6	6	0

Gene models quantified correspond to IWGSC RefSeq v1.1 high confidence gene models only and exclude splice variants. a. Number of putative disease resistance genes with any functional annotation that explicitly specifies disease resistance in its description and is evidenced by at least one of the following protein domains, when available: NB-ARC, LRR, RPM, RGA. b. Number of putative disease resistance genes with any functional annotation that explicitly specifies disease resistance in its description and is evidenced by at least one of the following protein domains, when available: ABC transporter, START kinase, sugar transporter, BED zinc finger.

QYr.niab-1A.1

Based on the reference genome sequence of wheat cv. Chinese Spring, the physical interval did not contain any NBS-LRR gene models. However, three gene models were identified as candidate genes for *QYr.niab-1A.1*: an ABC transporter (TraesCS1A01G404300) and two sugar transporter proteins (TraesCS1A01G405600, TraesCS1A01G405800). The putative ABC transporter is the closest candidate gene to the peak marker, located just 33.5 Kb away. *Yr46* and *Yr18* represent two cloned YR resistance genes known to encode a hexose transporter and an ABC transporter, respectively. tBLASTn analysis revealed no homology between these cloned genes and the candidate genes identified in the *QYr.niab-1A.1* interval.

QYr.niab-2A.1

The physical region for *QYr.niab-2A.1* was one of the largest among the QTLs investigated in this study (4.28 Mbp). Within this region, none of the gene functional annotations and protein domain terms (Pfam, Interpro) were associated with disease resistance. Looking beyond the physical region as defined here, TraesCS2A01G362300 was the closest predicted NBS-LRR gene, located approximately 1.1 Mbp away from the physical location of the peak marker, and showed some protein sequence similarity to Lr1 (41 % identity, e-value=2.50E-54). This gene was followed by two clusters of two and five disease resistance genes respectively. TraesCS2A01G361200, one of the NBS-LRR genes in the cluster of five genes, was located close (13.6 kb) to SNP BS00049644_51, identified as the peak marker at this QTL location in six YR assessments (Table 4.4). This NBS-LRR gene showed most protein sequence similarity to Sr22 (48.5 % identity, e-value=5.5E-105).

Table 4.4 Instances of BS000049644_51 as peak marker for *QYr.niab-2A.1*

Trial	Mapping method	Score	*P* value	% phenotypic variation explained
NIAB15	IM	S2	1.16E-06	5.69
NIAB15	HA	S2	2.34E-07	na
OSG15	IM	S2	2.78E-09	8.11
OSG15	CIM	S2	3.28E-11	8.11
OSG15	IM	S3	3.70E-09	8.65
OSG15	HA	S3	5.46E-12	na
ROTH15	IM	S3	6.26E-10	7.25
NIAB16	IM	S2	1.85E-14	8.37
NIAB16	IM	S4_log	5.74E-07	6.66
OSG16	IM	S2_log	6.77E-10	6.23
OSG16	HA	S2_log	7.61E-08	na

For each trial, mapping method and disease score were considered independently. Interval mapping was carried out with 0 (Interval Mapping, IM) and 10 covariates (Composite Interval Mapping, CIM) using founder probabilities. HA: Haplotype Analysis. Na: not applicable. Trials name abbreviations are described in Chapter 2, section 2.1.

QYr.niab-2A.2

The physical interval for *QYr.niab-2A.2* contained 12 NBS-LRR encoding gene models. The closest cluster of NBS-LRR candidates to the peak marker BS00011599_51 (0.69-0.47 Mb away) were TraesCS2A01G560600, TraesCS2A01G560700, TraesCS2A01G560900 and TraesCS2A01G561100. All four candidates are most similar at the protein level to Lr1 (40-47 % identity, 5.2E-69<e value<5.8E-52). Similarity to one rather than several of the cloned genes suggests that these four candidates share some ancestry. Indeed, TraesCS2A01G560600 is paralogous to the closely located genes TraesCS2A01G560700 (53 % protein identity), TraesCS2A01G560900 (54 % protein identity) and TraesCS2A01G561100 (64 % protein identity). In contrast, the other eight NBS-LRR genes, located much further away from the peak marker shared protein sequence similarity with a broader range of map-based cloned disease resistance genes (Appendix D). TraesCS2A01G563200 was the only gene that shared no similarity with any of the 13 cloned disease resistance genes (Table 4.1).

In much closer proximity to the peak marker was gene model TraesCS2A01G559300 (3.4 Kb away), predicted to encode a sugar phosphate transporter domain (IPR004853). Immediately adjacent to TraesCS2A01G559300 was gene model TraesCS2A01G559400, encoding a receptor kinase. These gene models were considered of note since two of the map-based cloned *Yr* genes encode a sugar transporter (*Yr46*) and a kinase-START (*Yr36*).

Two markers, BobWhite_c5756_532 and BS00064055_51, co-segregated with the consensus marker (i.e. most frequent peak marker identified from analysis of the multiple YR trials described in Chapter 3) BS00011599_51 (physical map position 762.3 Mbp) on the MAGIC genetic map at 259.39 cM. Their physical positions as predicted by BLASTn however, differed considerably. BobWhite_c5756_532 and BS00064055_51 were located at 763.1 and 765.3 Mbp, 0.76 and 2.98 Mb away from the peak marker, respectively. SNP BobWhite_c5756_532 was located within gene model TraesCS2A01G561300, a putative GATA transcription factor Two genes upstream is TraesCS2A01G561100, a predicted NBS-LRR gene. The more physically distant co-segregating SNP marker BS00064055_51 was located 93.7 Kb away from TraesCS2A01G565800, the closest gene model predicted to encode an NBS-LRR

(and described as disease resistance protein RPM1 in the IWGSC RefSeq v1.1 functional annotations).

QYr.niab-2B.1

The 5.61 Mbp physical interval for QYr.niab-2B.1 contained 13 candidate genes, representing one of the highest numbers of putative disease resistance gene models among the QTLs explored in this study (Table 4.5). This included ten NBS-LRRs, all of which shared significant protein sequence similarity with Lr1 (37-45 % identity, 2.8E-118<e-value<52.5E-43). Additionally, one ABC-like candidate gene was identified (TraesCS2B01G485600), which did not show significant protein sequence similarity to the protein of the cloned ABC transporter *Yr18/Lr34*. Finally, two Resistance Gene Analogues (RGA; genes considered to be candidate *R* genes due to their sequence similarity to conserved disease resistance domains and motifs. TraesCS2B01G487600, TraesCS2B01G488600) were identified but showed no tBLASTn similarity to the previously cloned *Yr*, *Sr*, *Lr* and Pm disease resistance genes.

Interestingly, the physical interval for *QYr.niab-2B.1* contained four gene models that each comprised a zinc finger BED domain (TraesCS2B01G488000, TraesCS2B01G488400, TraesCS2B01G488600, TraesCS2B01G489400). The recent cloning of *Yr7* and *Yr5/YrSP* demonstrated the involvement of such domains in conferring yellow rust resistance (Marchal et al., 2018). Indeed, the equivalent Chinese Spring gene model for *Yr7* is TraesCS2B01G488000 and the equivalent Chinese Spring gene model for *Yr5/YrSP* is TraesCS2B01G488600 (Marchal et al., 2018).

The peak marker co-segregated with three additional markers (BobWhite_c22728_78, BS00010687_51, BS00035276_51) that were distributed across the physical interval under study (683-687.25 Mbp, a 4.2 Mbp interval). All candidates but the ABC transporter fall within this physical interval, further highlighting them as good candidates for *QYr.niab-2B.1*.

QYr.niab-2D.1

The physical interval for *QYr.niab-2D.1* had the most candidate genes out of all of the QTL analysed in this study, totalling 15 NBS-LRR encoding genes (Table 4.5). Four were annotated as encoding RPM1-like proteins (RPM-1 is an NBS-LRR protein providing resistance against the bacterial pathogen *Pseudomonas syringae*). The NBS-LRR genes were present in clusters, with one large cluster comprising eight genes, while upstream of it, a smaller cluster of two genes was located. The closest NBS-LRR gene model to the peak marker within the region was located 0.26 Mbp away. The RPM1 genes were located much further away (1.31-2.27 Mbp distance), with three present in a cluster and one in isolation. tBLASTn analysis indicated that the candidate gene products in this interval were similar in sequence to a wide range of classes of the 13 disease resistance proteins (Table 4.5). This is unlike other QTL intervals investigated here, within which protein sequence similarity between gene products of candidate and cloned genes was restricted to a single cloned gene (e.g. the best tBLASTn hits for *QYr.niab-2B.1* candidate all show homology with Lr1). The level of similarity in protein sequence was comparable to other instances in this study (3.70E-163<e value<5.30E-08, 30.8-66.7 % ID).

QYr.niab-3A.1

Eight candidate genes were identified for *QYr.niab-3A.1*. TraesCS3A01G008100 was annotated as RPM1-like, with the remaining seven also belonging to the NBS-LRR family. As was the case for other QTLs investigated here, candidate genes occurred in clusters, with one larger core of four candidates (three NBS-LRR and one RPM-like) located at 7.9-8.1 Mb, and a second consisting of two NBS-LRR-encoding genes at 8.3-8.5 Mb. The peak marker Kukri_c28650_111 lies within TraesCS3A01G008000, an NBS-LRR gene from the former cluster. Within the same candidate gene cluster, TraesCS3A01G008100, annotated as encoding an RPM1-like gene, consists of an NB-ARC domain (PF00931), as well as a thioredoxin domain (PF00085), a common integrative domain (Sarris et al., 2016). The remaining candidates were scattered across the physical interval.

Furthermore, the peak marker co-segregated with Jagger_c6722_104 on the MAGIC genetic map. Its physical position overlaps with that of TraesCS3A01G007300, a gene model predicted to encode a 50S ribosomal protein. Of note is TraesCS3A01G007400, an NBS-LRR located immediately adjacent to TraesCS3A01G007300 (positioned 20.6 Kb away from the co-segregating marker). tBLASTn analysis indicated that all candidate gene products in this interval most resemble Sr22 (1.30E-64<e value<3.80E-24, 34-55.6 % ID). The only exceptions were TraesCS3A02G009300 showing protein sequence similarity with Sr35 (e value = 1.30E-64, % ID = 55.6), and TraesCS3A02G008200 and TraesCS3A02G008300, which did not resemble any of the cloned gene predicted proteins.

QYr.niab-6A.1

The consensus peak marker BS00011010_51 overlaps TraesCS6A01G037800 and TraesCS6A01G037900, two genes encoding an RNA-binding NOB1 and a Kelch repeat-containing protein. Its position overlaps with that of gene model TraesCS6A01G041500, predicted to encode a transmembrane protein consisting of an EXPERA domain (IPR033118) and a Sigma intracellular receptor 2 (IPR016964).

Within the physical interval of *QYr.niab-6A.1*, there was one gene model annotated as an NBS-LRR: TraesCS6A01G035900. However, the corresponding set of terms describing the protein it encodes consisted only of the Leucine Rich Repeat feature (IPR032675) and lacked the NB-ARC domain. This gene model is located 1.06 Mbp away from the peak marker and showed no significant tBLASTn hits with the 13 cloned disease resistance genes.

QYr.niab-6A.2

Two distinct clusters of NBS-LRR genes were located within the *QYr.niab-6A.2* physical region. Each cluster consists of either two or four genes annotated as NBS-LRRs. The gene models (TraesCS6A01G051900, TraesCS6A01G052200, TraesCS6A01G052300 and TraesCS6A01G052500) were described by the IWGSC annotations as potentially being transposable elements, most likely due to the presence of the 'DDE superfamily endonuclease' (PF13359) domains. The predicted proteins of these four gene models shared protein sequence similarity with Sr22 (e

value = 1.30E-17, % ID = 43.9), Pm3b (e value = 2.30E-15, % ID = 29.5), Lr1 (e value = 2.50E-54, % ID = 41) and Sr45 (e value = 3.70E-18, % ID = 28.6), respectively. Gene models TraesCS6A01G051800 and TraesCS6A01G052400 were not annotated as containing a DDE superfamily endonuclease domain. Based on tBLASTn analysis, the predicted proteins of both genes showed protein sequence similarity to Pm3b (TraesCS6A01G051800: e value = 1.40E-13, 35.4 % ID; TraesCS6A01G052400: e value = 2.20E-18, 34.6 % ID

Table 4.5 Candidate genes for all eight investigated YR resistance QTL and their corresponding IWGSC RefSeq v1.1 functional annotation

QTL	Chr	Gene model	Physical start (Mb)	IWGSC functional annotation
QYr.niab-1A.1	1A	TraesCS1A01G404300	567.97	Multidrug resistance protein ABC transporter family protein
		TraesCS1A01G405600	568.41	Sugar transporter protein
		TraesCS1A01G405800	568.54	Sugar transporter protein
QYr.niab-2A.2	2A	TraesCS2A01G560600	762.76	CC-NBS-LRR disease resistance protein
		TraesCS2A01G560700	762.77	Disease resistance protein (NBS-LRR class) family
		TraesCS2A01G560900	762.80	Disease resistance protein (NBS-LRR class) family
		TraesCS2A01G561100	762.98	NBS-LRR disease resistance protein
		TraesCS2A01G563200	764.05	NBS-LRR resistance-like protein
		TraesCS2A01G564200	764.66	Disease resistance protein (NBS-LRR class) family
		TraesCS2A01G564300	764.68	NBS-LRR resistance-like protein
		TraesCS2A01G564500	764.72	disease resistance family protein / LRR family protein
		TraesCS2A01G564600	764.73	disease resistance protein (TIR-NBS-LRR class)
		TraesCS2A01G564800	764.78	Disease resistance protein (NBS-LRR class) family
		TraesCS2A01G564900	764.82	NBS-LRR resistance-like protein
		TraesCS2A01G565800	765.17	Disease resistance protein RPM1
QYr.niab-2B.1	2B	TraesCS2B01G485600	682.85	Multidrug resistance protein ABC transporter family protein
		TraesCS2B01G486100	683.04	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G486200	683.05	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G486300	683.07	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G486400	683.13	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G486500*	683.13	Disease resistance protein (TIR-NBS-LRR class) family
		TraesCS2B01G486700	683.16	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G487600*	683.75	Disease resistance protein RGA2
		TraesCS2B01G487700	683.75	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G488000	685.27	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G488400	685.74	NBS-LRR disease resistance protein-like protein
		TraesCS2B01G488600*	686.05	Disease resistance protein RGA2
		TraesCS2B01G488700	686.05	Disease resistance protein (TIR-NBS-LRR class) family
		TraesCS2B01G489400	686.81	NBS-LRR disease resistance protein-like protein
QYr.niab-2D.1	2D	TraesCS2D01G573600	638.64	Disease resistance protein (NBS-LRR class) family
		TraesCS2D01G573800	638.77	Disease resistance protein (NBS-LRR class) family

QTL	Chr	Gene model	Physical start (Mb)	IWGSC functional annotation
		TraesCS2D01G573900	638.78	disease resistance protein (TIR-NBS-LRR class)
		TraesCS2D01G574300	638.86	disease resistance protein (TIR-NBS-LRR class)
		TraesCS2D01G574400	639.09	Disease resistance protein (NBS-LRR class) family
		TraesCS2D01G574500	639.14	Disease resistance protein (TIR-NBS-LRR class) family
		TraesCS2D01G574600	639.15	Disease resistance protein (NBS-LRR class) family
		TraesCS2D01G574700	639.28	NBS-LRR-like resistance protein
		TraesCS2D01G574800	639.29	Disease resistance protein (TIR-NBS-LRR class) family
		TraesCS2D01G574900*	639.31	NBS-LRR resistance-like protein
		TraesCS2D01G575000	639.36	Disease resistance protein (NBS-LRR class) family
		TraesCS2D01G576100	639.69	Disease resistance protein RPM1
		TraesCS2D01G577300	640.58	Disease resistance protein RPM1
		TraesCS2D01G577400	640.63	Disease resistance protein RPM1
		TraesCS2D01G577500	640.65	Disease resistance protein RPM1
QYr.niab-3A.1	3A	TraesCS3A01G007400	7.50	Disease resistance protein (TIR-NBS-LRR class)
		TraesCS3A01G008000	7.92	Disease resistance protein (NBS-LRR class) family
		TraesCS3A01G008100	7.95	Disease resistance protein RPM1
		TraesCS3A01G008200	8.03	Disease resistance protein (NBS-LRR class) family
		TraesCS3A01G008300	8.06	Disease resistance protein (TIR-NBS-LRR class) family
		TraesCS3A01G008600	8.16	disease resistance protein (TIR-NBS-LRR class)
		TraesCS3A01G009200	8.32	NBS-LRR-like resistance protein
		TraesCS3A01G009300	8.55	Disease resistance protein (TIR-NBS-LRR class) family
QYr.niab-6A.1	6A	TraesCS6A01G035900	17.65	NBS-LRR resistance-like protein
QYr.niab-6A.2	6A	TraesCS6A01G051800	26.95	NBS-LRR-like resistance protein
		TraesCS6A01G051900	26.96	NBS-LRR-like resistance protein
		TraesCS6A01G052200	27.00	NBS-LRR-like resistance protein
		TraesCS6A01G052300	27.06	NBS-LRR-like resistance protein
		TraesCS6A01G052400	27.08	NBS-LRR-like resistance protein
		TraesCS6A01G052500	27.10	NBS-LRR-like resistance protein

* No tBLASTn results in Appendix 5A

4.5 Discussion

Adult plant yellow rust resistance in the MAGIC population was found to be controlled by over ten QTL. To date, the current literature indicates that one of these MAGIC QTL overlaps with the cloned yellow rust resistance genes *Yr7* and *Yr5/YrSP* (Marchal et al., 2018), indicating that the gene underlying *QYr.niab-2B.1* may represent either *Yr7* or *Yr5/YrSP*. All the remaining QTL represent uncloned *Yr* genes and are of interest for further characterisation. This chapter therefore set out to (i) examine the physical genomic space for the eight most significant QTL conferring yellow rust resistance in the UK wheat MAGIC population and (ii) identify candidate genes underlying yellow rust resistance for further investigation via comparison to functional annotation terms and protein sequences for known rust and powdery mildew *R* genes.

In this discussion, I start by putting the physical intervals into the wider context of the latest wheat genome assembly (Appels et al., 2018) and recent findings on the physical distribution of NBS-LRR genes in the wheat genome. This is then followed by a more detailed discussion on candidate genes of note, with reference to known cloned biotrophic disease resistance genes. Then, the candidate gene selection approach taken herein is critically assessed and compared to other available approaches, leading to a final section on suggested future work.

4.5.1 Physical characteristics of the YR QTL investigated

Each physical interval was based on the peak markers for the QTLs identified in Chapter 3 and were delineated based on the re-ordering of the most significant markers according to the physical map. Intervals varied significantly in size between the eight QTLs, from less than 1 Mbp to over 5 Mbp. This variation reflects the number, size of, and spacing between genes within each interval. *QYr.niab-1A.1* and *QYr.niab-2B.1* had the two largest intervals and exhibited some of the largest average spacing between genes. Interval size did not always correlate with number of genes. *QYr.niab-2B.1* for instance had the lowest gene density of all intervals investigated. The B sub-genome, however, is known to have a higher gene content compared to the A and D sub-genomes (Appels et al., 2018). However, it is worth noting that the gene content number reported in the IWGSC RefSeq v1.1

annotation included both HC and LC genes, whereas here I consider the HC gene models only. Omission of LC genes from this study could therefore explain those differences.

Over half of the QTL investigated in this study located to physical regions of the wheat genome that exhibit relatively high numbers of NBS-LRR genes. These findings are in line with published information about the physical location of R genes in the wheat genome. For example, the IWGCS consortium looked at the gene distribution of NBS-LRR genes, which were found to cluster in distal, mostly telomeric regions of chromosomes. These distal locations and clustering patterns of distribution are further confirmed by Steuernagel et al. (2018), who specifically focused on improved annotation of NBS-LRR loci in the wheat IWGSC Refseq v1.0 genome assembly. The physical intervals for *QYr.niab-1A.1* and *QYr.niab-6A.1* were located within regions with a lower density of NBS-LRR-encoding genes compared to the other QTLs investigated here. Again, I only considered genes annotated as high confidence genes. Considering that low confidence genes make up approximately 60 % of the total annotated gene content in the IWGSC RefSeq v1.1 annotation (Appels et al., 2018), it is thus likely that low confidence NBS-LRR genes may be located in the QTL intervals of *QYr.niab-1A.1* and *QYr.niab-6A.1*.

4.5.2 NBS-LRR as candidate genes

Out of the eight QTL selected for further investigation in this chapter, five were characterised by NBS-LRR clusters within their physical interval (*QYr.niab-2B.1, QYr.niab-2D.1, QYr.niab-3A.1* and *QYr.niab-6A.2*) or in very close proximity (*QYr.niab-2A.1*). These findings may reflect the predominance of NBS-LRR-encoding *R* genes in mediating plant immunity against pathogens (Kourelis & van der Hoorn, 2018). In this section, I discuss NBS-LRR genes, with additional attention paid to those proteins predicted to contain integrated domains, recently shown to be involved in plant immunity as pathogenicity sensors (Cesari et al., 2014; Sarris et al., 2016; Wu et al., 2015).

Effector-triggered plant immunity is mostly governed by intracellular receptors featuring NBS-LRR domains, which recognize pathogenicity effectors either directly or indirectly (Jones & Dangl, 2006; Kourelis & van der Hoorn, 2018). This

type of resistance leads to the well-characterised hypersensitive response, underpinned by programmed cell death (Jones & Dangl, 2006). NBS-LRR proteins are encoded by genes belonging to some of the largest and most diverse gene families in plants (Monteiro & Nishimura, 2018; Sarris et al., 2016). They are often organised as clusters in the genome, a characteristic that was also observed in the five yellow rust QTL identified here as being characterised as containing NBS-LRR clusters. This phenomenon has been extensively investigated and reviewed in plant genomes (Hulbert et al., 2001; Michelmore & Meyers, 1998; Noël et al., 1999; Seo et al., 2016). In cereals, two important examples of NBS-LRR clusters include the barley *Mla* locus (Seeholzer et al., 2010; F. Wei et al., 1999) and its homologous counterpart in wheat, which includes *Sr33* and *Sr50* (Mago et al., 2015; Periyannan et al., 2013).

NBS-LRR function can also be aided by 'integrated' domains for effector recognition (Baggs, Dagdas, & Krasileva, 2017; Cesari et al., 2014; Kroj et al., 2016; Sarris et al., 2016). This additional domain is suggested to serve as a bait or decoy to pathogenicity effectors, and some of the candidate genes identified notably for *QYr.niab-2B.1, QYr.niab-3A.1* and *QYr.niab-6A.2* in this study appear to encode such domains.

QYr.niab-2B.1 is located on the long arm of chromosome 2B, a region known to be particularly dense in yellow rust resistance genes (Feng et al., 2015; Luo et al., 2009). The peak marker that was most frequently identified across all trials and scoring dates for this QTL (Kukri_c9118_1774, 70 % of trials and scoring dates) co-segregated with 24 other markers – this is as might be expected from the physical location of the QTL peak on the edge of the highly non-recombining region on the short arm of chromosome 2B. Within the *QYr.niab-2B.1* physical interval, 14 genes were identified as candidates and ten of these were annotated as NBS-LRRs. Four were predicted to encode a zinc finger BED domain (TraesCS2B01G488000, TraesCS2B01G488400, TraesCS2B01G488600, TraesCS2B01G489400). The recently cloned *Yr7, Yr5* and *YrSP* genes also encode NBS-LRRs with this integrated BED domain (Marchal et al., 2018). While both *Yr7* and *YrSP* no longer provide adequate resistance in the field, *Yr5* continues to remain effective. A syntenic analysis of this *Yr7/Yr5/YrSP* region, performed across the wheat genomes and

related grasses, revealed that three of the characterised BED-NBS-LRRs (Ta_2B9, Ta_2B10, Ta_2B11) overlapped with three IWGSC RefSeq v1.1 gene models identified as candidates in this study (TraesCS2B01G488000, TraesCS2B01G488400 and TraesCS2B01G488600). Notably, Marchal's study (2018) seems to suggest that candidate genes TraesCS2B01G488000 and TraesCS2B01G488600/TraesCS2B01G488700 are the Chinese Spring homologs of *Yr7* and *Yr5/YrSP* respectively, indicating that *QYr.niab-2B.1* may be encoded by one of these previously cloned *Yr* genes. However, the peak marker for *QYr.niab-2B.1* (Kukri_c9118_1774) locates within gene model TraesCS2B01G486100, upstream of the BED-NBS-LRR cluster that includes *Yr7* and *Yr5/YrSP*. While the IWGSC RefSeq v1.1 functional annotations do not describe this gene model as encoding a BED domain, the NLR-Annotator programme used in Marchal et al. (2018) and Steuernagel et al. (2018) characterised it as a BED-containing NBS-LRR. Overall, these findings indicate that all canonical and non-canonical NBS-LRR genes within this region are worthy of further investigation.

Marchal's study also tested *Yr7* and *YrSP* diagnostic markers in global wheat collections, which included Soissons, the main contributing founder of yellow rust resistance for *QYr.niab-2B.1* in the MAGIC population used here. In Marchal's presence/absence study, both *Yr7* and *YrSP* were absent from Soissons, suggesting that the resistance conferred by *QYr.niab-2B.1* is likely not determined by either of these *Yr* genes. Soissons is postulated to carry the seedling resistance genes *Yr3a* and *Yr4a* (de Vallavieille-Pope, Picard-Formery, Radulovic, & Johnson, 1990), but its durable adult plant resistance remains uncharacterised (de Vallavieille-Pope et al., 2011).

QYr.niab-3A.1 localised to the short arm of chromosome 3A. Eight RefSeq v1.1 gene models were identified as candidates, based on their functional annotations. Of note was TraesCS3A01G008100, a gene model recently annotated to contain a thioredoxin domain (Appels et al., 2018; Steuernagel et al., 2018), a class of integrative domain. In a recently reported study, of the 1,540 NLR investigated for their integrated domains, only two contained a thioredoxin domain (Steuernagel et al., 2018), including TraesCS3A01G008100 identified here. The other wheat gene found to contain a thioredoxin domain, TraesCS3D02G002700, was located

on chromosome 3D and is homeologous to TraesCS3A01G008100. Thioredoxins are key disulphide proteins involved in metabolic regulation and oxidative stress mitigation (Schürmann & Jacquot, 2000; Vieira Dos Santos & Rey, 2006). Sarris et al. (2016) reported this family of proteins as one of the most prevalent integrated domains in flowering plants and as having been involved in plant-pathogen interaction activity.

QYr.niab-6A.2 provides another example of a physical interval consisting of NBS-LRRs containing integrated domains. However, these domains contain the more prevalent DDE_Tnp_4 domain, which belongs to the DDE superfamily endonucleases. These domains commonly form part of transposons, where they act as a catalyst for transposition. DDE domains have been shown to integrate with NBS-LRRs in wheat (Sarris et al., 2016; Steuernagel et al., 2018) and related grasses (Sarris et al., 2016). In the integrated decoy model (Cesari et al., 2014), unusual fused domains are hypothesised to act as sensors or targets of pathogenicity effectors. However, there is no evidence to date that would suggest transposons playing such a role during elicitation of an immune response in the host.

In conclusion, the presence of canonical and non-canonical NBS-LRR with a broad range of integrated domains either within or in close proximity to the peak marker for each QTL investigated may provide some clue about the functional mechanisms underlying yellow rust resistance in the MAGIC population.

4.5.3 Beyond NBS-LRR as candidate genes

There were three notable instances when NBS-LRR were not identified as the obvious candidates, either because they were too distantly located from the QTL peak marker within a physical interval or because they were absent from the interval completely. Lack of NBS-LRR gene models within physical intervals was a feature of *QYr.niab-1A.1, QYr.niab-6A.1* and *QYr.niab-2A.1.* For these, a number of candidate genes belonging to other types of gene family were identified, and their potential role in YR resistance is discussed here.

The *QYr.niab-1A.1* physical interval contains an ABC transporter and two sugar transporters, proteins that have been shown to confer durable yellow rust adult

plant resistance (Krattinger et al., 2009; Moore et al., 2015). A fourth likely and perhaps more plausible candidate is TraesCS1A01G404500, an aspartic proteinase, in which the two most significant markers were located. In plants, this family of proteolytic enzymes is characterised by two aspartic residues required for catalysis and a plant-specific insert that resembles saposin-like proteins, when the enzyme is in its precursor form (Simões & Faro, 2004). The biological function of aspartic proteases remains for the most part speculative but generally, they have been implicated in protein processing in a range of processes, including senescence, programmed cell death, reproduction and stress responses (reviewed in Simões & Faro, 2004). More specifically, aspartic proteases have also been involved in plant innate immunity signalling, a role which has recently been the subject of several reviews (Balakireva & Zamyatnin, 2018; Hou, Jamieson, & He, 2018; Thomas & van der Hoorn, 2018). In *Arabidopsis* and rice (*Oryza sativa*), overexpression of aspartic protease-encoding genes enhance resistance against bacterial and fungal pathogens (Prasad et al., 2009; Xia et al., 2004). Interestingly, *Arabidopsis* CDR1 does not contain the plant-specific insert characteristic of aspartic proteases. In potato (*Solanum tuberosum*), aspartic proteases have been shown to exhibit antimicrobial activity against several fungal pathogens (Mendieta et al., 2006; Muñoz et al., 2010). To date, while aspartic proteases have been characterised in wheat seeds and are speculated to play a role in gluten degradation (Belozersky, Sarbakanova, & Dunaevsky, 1989; Tamura et al., 2007), their implication in disease resistance signalling has not been demonstrated. In addition to the aspartic proteinase, tplb0021i12_383, the third co-segregating marker lies within a gene predicted to encode a subtilisin-like protease, a serine protease with pathogen recognition properties (Figueiredo, Sousa Silva, & Figueiredo, 2018).

The second example of the lack of NBS-LRR genes within the physical region came from *QYr.niab-2A.1*. Here, the consensus peak marker lies close (6.8 kb) to TraesCS2A01G363500, an isocitrate dehydrogenase. These enzymes catalyse the production of NADPH and have been suggested to play a role in redox signalling, nitrogen assimilation and oxidative stress prevention (Hodges et al., 2003). More recently, isocitrate dehydrogenases have been shown to provide antioxidative properties under abiotic stress conditions in pea (Leterrier et al., 2012, 2007). In

wheat, an isocitrate dehydrogenase was recently genetically characterised and expression studies under nutrient stress conditions suggest that the enzyme could be linked to nitrogen assimilation (Rani et al., 2019). While the vast majority of studies point towards a biological function in metabolic fluxes, there is some evidence of isocitrate dehydrogenase activity being linked to pathogen colonisation (Mhamdi et al., 2010). In addition to TraesCS2A01G363500, surrounding proteins of note also included a glutaredoxin (TraesCS2A01G363600), and two hydrolases (TraesCS2A01G363400 and TraesCS2A01G363200). The proximity of several NBS-LRR at the upmost edge of the interval and their omission from the physical interval is discussed further in section 4.6.

Finally, *QYr.niab-6A.1* represents an additional instance whereby NBS-LRRs may not be the most likely candidate genes. While the physical interval did include an NBS-LRR, its distance from the peak (1.2 Mbp) marker suggests it may not be a likely candidate. The physical interval of *QYr.niab-6A.1* is characterised by a cluster of kinase-encoding genes at its proximal boundary. While a particular type of kinase (Kinase-START) has been shown to confer durable yellow rust resistance (Fu et al., 2009), the location of the putative kinases in relation to the peak marker would suggest that they may be less likely to be implicated in the resistance conferred by *QYr.niab-6A.1*. The consensus marker for *QYr.niab-6A.1* overlaps with TraesCS6A01G037900, a gene model for which the transcript is annotated as a Kelch repeat-containing protein. The Kelch domain is prevalent in flowering plants and has been shown to be a likely decoy domain for pathogens (Sarris et al., 2016).

4.5.4 Physical intervals to reconsider

In some instances, taking a systematic and thus more general approach to selecting the physical region for all eight QTL may have omitted some important genomic areas and therefore misguided the selection of candidate genes involved in yellow rust resistance. Based on the specificities of each of the QTL, different explanations are possible and are discussed in this section.

For *QYr.niab-2A.1,* the initial and perhaps stringent delineation may have omitted NBS-LRR clusters worth exploring. Indeed, unlike other yellow rust QTL characterised in the MAGIC population, *QYr.niab-2A.1* exhibited considerable

variation in peak marker across the different field trials and YR scores. While the agreed consensus based on occurrence was BS00022903_51, there were four additional markers identified as most significant in at least ten instances (trials and score confounded): BS00049644_51, BS00022641_51, TA006396_1406, and CAP12_rep_c7918_56. When expanding the physical interval to include those trial and score-specific peak markers, additional candidates come into play, including 14 NBS-LRR-encoding genes. This expansion would see the physical interval increase to 27 Mbp, from the much smaller 5.6 Mbp examined in this study. Further work is thus likely required to determine a more robust interval for this QTL.

In contrast, *QYr.niab-2D.1* is a good example of how the lack of high marker density in a genetic map and the chromosomal region in question can impact candidate gene selection. Firstly, the considerable segregation distortion in the 2D chromosome, as evidenced by the PCA analysis in Gardner, Wittern & Mackay (2016) and the scattered distribution of marker blocks, may have biased the QTL mapping. This hypothesis is supported by the physical position of the five markers below the consensus peak. All of them appear to have been anchored at a physical position that does not match the order of the genetic map. Secondly and more importantly, the last marker of the genetic map for chromosome 2D is still highly significant for YR resistance and stops at 638-642 Mb on the physical map. However, the genome assembly for chromosome 2D extends beyond this, to 651.8 Mb, meaning that approximately 10 Mb are unaccounted for by the MAGIC genetic map. It is therefore highly likely that the region of significance extends beyond that the current genetic map accounts for. A number of QTL mapping studies focusing on 2DL have been undertaken in the last ten years or so (Basnet et al., 2014; Jagger et al., 2011; Mallard et al., 2005; Melichar et al., 2008; Powell et al., 2013), and point towards a similar degree of difficulty in mapping to a precise location. This potential bias aside, the physical interval for *QYr.niab-2D.1* was characterised by a number of NBS-LRR clusters. The peak marker is located approximately 260 kb away from the first NBS-LRR gene model within this region. The distal telomeric region of the 2D chromosome is evidently rich in NBS-LRR (Appels et al., 2018;

Steuernagel et al., 2018) (Figure 5.3), indicating that they are a likely candidate gene family for *QYr.niab-2D.1*.

Finally, *QYr.niab-2A.2* represented a situation whereby markers co-segregating with the peak markers on the genetic map were distant from each other on the physical map. NBS-LRR genes were present within the physical interval, but their considerable distance from the peak marker suggested that they might be considered lower priorities as candidate genes. Instead, candidate genes more closely located to the QTL peak included TraesCS2A01G565800, a disease resistance protein RPM1 closely linked to the co-segregating marker BS00064055_51. With over 70 gene functional annotations within a 3.4 Mb interval, this physical interval requires narrowing down to then characterise a smaller and more robust set of candidate genes for further investigation.

In addition to the physical specificities of each QTL interval, the IWGSC RefSeq v1.1 NLR annotations can also be questioned. In regions where no NLR were reported in the RefSeq v1.1 gene annotations, there is the possibility that the annotation pipeline used in the latest assembly of Chinese Spring may have not have identified all NLR genes present. NLR-Annotator identified over 3,000 NLRs in Chinese Spring, whereas the IWGSC annotations only identified around 1,500. Although it is speculated that some of these are pseudogenes rather than functional ones, there is a possibility that the additional NLRs could have functional homologues and orthologues not present in Chinese Spring. Furthermore, significant sequence and NLR copy number variation between wheat varieties exists, as is now becoming evident from recent studies resequencing additional wheat accessions.

This study has provided an initial insight into the genes potentially underlying yellow rust resistance for eight of the QTL characterised in the MAGIC population. For some QTL, the process of candidate gene selection proved more straight forward, either because the region was a lot smaller to start with, or because it comprised the 'obvious' NBS-LRR candidate genes. For others, the physical location of the QTL on a chromosome meant that the physical region was difficult to delineate. In either case, further work is required at different levels to confirm and/or narrow down candidate genes.

In addition, low confidence genes were not considered in this study. For QTL with little (*QYr.niab-6A.1*) or no NBS-LRR (*QYr.niab-1A.1, QYr.niab-2A.1*), this would be particularly important to either include or rule out as potential candidate genes.

The physical interval of a number of QTL requires further refinement e.g. *QYr.niab-2A.2* and *QYr.niab-2D.1*. This could be done by fine-mapping the region further for example, with Heterogeneous Inbred Families (HIF), taking advantages of the residual heterozygosity present in MAGIC F_5 lines to rapidly develop heterogeneous inbred lines (Tuinstra et al., 1997). Alternatively, RNAseq data could provide the transcript evidence required to investigate differential expression, similar to that done with the investigation of the *SSt1* region (1.6 Mb), narrowed down six-fold, from 160 genes altogether present in the region to 26 differentially expressed genes (Appels et al., 2018). The Wheat Expression Browser 'expVIP' is an example of an additional source of publicly available RNAseq data that could be exploited here, which now comprises yellow rust infection data at the seedling stage (Borrill et al., 2016; Ramírez-González et al., 2018).

Finally, this study relied on functional annotations based on the assembly of the Chinese Spring genome. It is therefore likely that some of the candidates identified are either not functional in either of the eight MAGIC founders or that some candidates have been missed altogether because they are absent, have not been annotated as a gene model, or the underlying genomic sequence has been mis-assembled in the Chinese Spring reference genome. All MAGIC founder genomes are being sequenced as part of an ongoing project (https://gtr.ukri.org/projects?ref=BB%2FP010741%2F1) and this sequencing data will serve as a source of potential validation for the candidate genes characterised in this study. In addition, high throughput annotation platforms for NBS-LRR genes such as NLR-Parser (Steuernagel et al., 2015) and the more recent NLR-Annotator (Steuernagel et al., 2018) ought to complement the fine-mapping. The latter allows for *in silico* annotations and is independent of transcript data.

4.6 Conclusion

I set out to describe the physical space for each of the eight most significant QTL conferring yellow rust resistance in the MAGIC population in the 2015-16 field seasons, and to identify candidate genes underlying this resistance. Functional annotations and tBLASTn analyses using previously cloned resistance genes revealed a range of putative *R* genes within QTL intervals as potential candidate genes. This study provides initial insights into the possible mechanisms of resistance at play in the MAGIC population at the gene level.

The application of the latest genomic tools and that of the reference genome have largely focused on deciphering brown and stem rust resistance genes (with the exception of the study by Marchal et al., 2018). Overall, 14 QTLs have been identified to confer yellow rust resistance in the MAGIC population. In the light of recurrent instances of PST emergence, this data provides a striking opportunity to implement the latest genomic and transcriptomic approaches for the rapid advancement of wheat yellow rust resistance gene cloning and a deeper understanding of the molecular mechanisms of resistance.

5 CHARACTERISATION OF QTLS CONFERRING YELLOW RUST RESISTANCE IN WHEAT EARS

While much research has focused on yellow rust infection of wheat leaves, comparatively little is known on the genetic control of yellow rust (YR) resistance in other structures, such as the ears. Here, genetic analysis of glume infection in two YR infected MAGIC field trials identified five QTL, each explaining between 3.4% and 6.8% of the phenotypic variation. Of these, three QTL (termed *QYrg.niab-2D.1, QYrg.niab-4D.1 and QYrg.niab-5A.1*) were found to co-locate with QTL for leaf YR resistance identified in Chapter 3 (*QYr.niab-2D.1, QYr.niab-4D.1 and QYr.niab-5A.1,* respectively). The remaining two glume YR resistance QTL were linked to flowering time loci. The first, *QYrg.niab-2D.Ppd_D1*, mapped to the major flowering time locus *PPD-D1*, with the early flowering allele originating from the founder Soissons conferring reduced glume YR resistance. The second, *QYrg.niab-4A.1*, was identified in one trial only, and was located close to the previously reported flowering time QTL *QFt.niab-4A.1*. This indicates that earlier heading date potentially results in increased glume YR susceptibility, possibly due to exposure of tissues during environmental conditions that are more favourable for *Pst* infection. Collectively, the results provide insights into the genetic control of YR resistance in glumes, providing breeders with targets to ensure resistance in both the leaves and the glumes, and may be especially relevant in agricultural environments where earlier flowering is favoured, or where high *Pst* infection is common early on in the season.

5.1 Introduction

Symptoms of yellow rust infection such as sporulation are typically observed on leaves. However, given the right conditions, *Pst* urediniospores will also colonise leaf sheaths, glumes and awns. Infection in those tissues tends to be less frequent and to correlate with periods of moderate to severe epidemics of yellow rust. While the main concern for wheat growers is foliar infection, glume infection can become a more significant problem during severe epidemics. Nevertheless, other than being reported when observed in unusually high levels, it is a trait that has received relatively little attention compared to its foliar counterpart.

5.1.1 Yellow rust in wheat ears

Pst uredinospores infect glumes from heading (Zadoks Growth Stage (GS) 55; Zadoks, Chang, & Konzak, 1974) until flowering (GS61), with symptoms appearing 10 to 20 days after (Wellings, 2003). Inside each floret, *Pst* germination and sporulation will occur on glumes, lemma and palea. At a first glance, infected wheat heads appear discoloured or bleached, and symptoms can be mistaken for that of other diseases such as Fusarium head blight. Upon closer inspection, pustules inside the spikes are more evident and opening the floret exposes urediniospores, thereby confirming *Pst* infection. These symptoms are short-lived compared to foliar infection, mainly because cycles or re-infection are more difficult when spores are encased within the florets. An additional symptom includes shrivelled grains, although this is largely dependent on the timing on infection. Generally, the earlier the infection at the onset of flowering, the higher the extent of shrivelled grain. In severe cases, glume infection can lead to grain quality downgrading, as exemplified by reports of a 77% reduction in grain weight relative to uninfected plants (Cromey, 1989b) and a yield reduction of 20 % in a resistant variety (Purdy & Allan, 1965).

Pst glume infection is mostly observed in moderately to highly susceptible varieties and this usually correlates with how varieties respond to leaf infection (Wellings, 2003). A variety with foliar resistance will generally also show resistance in the ear, while a susceptible variety is much more likely to exhibit symptoms of glume infection. However, there have been reports of moderately resistant varieties

exhibiting unusually high levels of glume infection (Cromey, 1989a; Purdy & Allan, 1965; Wellings, 2003, 2009). Little is known about the effects of environmental conditions, ear morphology and timing of flowering on the yellow rust infection in glumes, although it has been suggested that moisture and cool temperatures at flowering are key factors in the onset of significant outbreaks, and that florets are least susceptible to yellow rust after anthesis (Cromey, 1989b).

5.1.2 Relevance to wheat breeding

Pst infection of the leaves prevents leaf tissue from photosynthesizing. Extensive infection of the ears may have a similar detrimental effect on photosynthesis, particularly when leaf photosynthetic area is already significantly reduced due to extensive sporulation, necrosis and/or chlorosis. Ear photosynthesis significantly contributes towards grain filling by providing the developing grain with an important source of photoassimilates (10-76 %) and has shown to exhibit similar levels of CO_2 fixation compared to the flag leaf (Maydup et al., 2010; Tambussi et al., 2007). These characteristics are of particular relevance under drought conditions, when ears exhibit higher water use efficiency than flag leaves (Tambussi et al., 2007). The photosynthetic capacity of ears has therefore been recognised as an area of relevant to wheat breeding and is the target of further optimisation as part of international research efforts looking to increase wheat yield potential, such as the Wheat Yield Consortium (Parry et al., 2011).

In addition, if we consider the ways in which *Pst* uredinospores can colonise plant tissue, glume infection could potentially pose a threat to hybrid wheat breeding. As wheat is a predominantly inbreeding crop, developing hybrids requires male sterility in female parents, which receive pollen from male parents. This critical cross-pollination step relies on open florets during anthesis, providing an easy entry point for *Pst* urediniospores. In the context of a moderate to severe yellow rust outbreak, glume infection could result in non-negligible levels of shrivelled grain, posing a serious problem for the large-scale seed multiplication efforts required by commercial hybrid breeding.

5.2 Aims and overview

Knowledge of yellow rust glume infection in wheat is mostly limited to its phenotypic symptoms. The genetic components underlying glume resistance have currently not been explored. This chapter aims to use the NIAB Elite MAGIC population to (i) characterise the genetic basis of glume yellow rust resistance, and (ii) compare glume resistance QTL to the genetic basis of yellow rust resistance in leaves.

5.3 Material and Methods

Glume infection was quantitively assessed in the 2016 MAGIC pathology trials described in Chapters 2 and 3. This chapter briefly reiterates which plant material was used, how the phenotypic trait was assessed and which statistical methods were employed for analysis of phenotypic data and QTL mapping.

5.3.1 Plant material and experimental design

The MAGIC population was assessed in two untreated, partially replicated trials sown at two locations in autumn 2015, and phenotyped in the 2016 growth season: Osgodby (OSG16) and NIAB, Cambridge (NIAB16). Number of lines used and experiment design are summarised in Table 5.1. For further information, refer to Chapter 2, Section 2.1.

NIAB16 was artificially inoculated with a mixture of YR races: Solstice (isolate 08/21 virulent on *Yr 1, 2, 3, 4, 6, 9, 17, 25, 32*) and Warrior (isolate 11/08 virulent on *Yr 1, 2, 3, 4, 6, 7, 9, 17, 25, 32, Sp*). For OSG16, no inoculation was undertaken due to high levels of natural *Pst* inoculum early in the season. Similar, although slightly lower, levels of natural infection occurred for NIAB16. It is therefore highly likely that both sites were exposed to natural *Pst* infection.

Table 5.1 2016 pathology trial summary

	NIAB16	OSG16
Location	Cambridge, Cambridgeshire	Osgodby, Lincolnshire
Genotyped MAGIC lines[a]	678	678
Ungenotyped MAGIC lines	48	48
MAGIC founders[b]	8	8
Positive control (Oakley)[b]	1	1
Negative control (Cougar)[b]	1	1
Spreader positions	Every traverse	Every 3 traverses
Trial design	Unbalanced incomplete random block design	Unbalanced incomplete random block design

a. 444 MAGIC lines were replicated twice while the remaining 234 were present once in each trial; b: MAGIC founders and varietal controls were replicated three times. A traverse consists of *n*th plots, with each plot made up of six one metre long rows.

5.3.2 Phenotypic assessment of YR glume infection

A glume infection assessment scale was devised as part of the study. A quantitative scale was implemented (

Table 5.2), where 0 represents no infection and 4 represents all glumes per head infected. 10 to 12 ears selected at random were considered, ensuring that they were not emerging from the same plant. The number of spikelets per ear exhibiting glume infection was counted. This was converted into a percentage to obtain an estimate of the proportion of ear infection, which was assigned a number on the 0-4 scale. A score was given when ≥ 50 % of the plot corresponded to that score. Glume infection was assessed at the late milk development stage (Zadoks GS77).

Table 5.2 Glume infection scale

Score	% infection per wheat ear
0	0
1	25
2	50
3	75
4	100

Percentage infection in wheat heads is estimated by counting the number of spikelets per ear that exhibit YR pustules in the glumes. For example, 9 infected spikelets out of 12 means that 75 % of the ear is infected. If this is the case for ≥ 50 % of the plot, a score of 3 will be given.

5.3.3 Statistical analysis of glume infection

Inspection of the glume infection distribution and residual plots revealed a dataset skewed towards resistance. The natural log was applied to improve the normality of the data:

$$y = \ln (x + 1) ,$$ (7)

where x is the original score and y is the transformed score. The glume infection severity scale starts at zero, a value not taken by the natural log equation. A value of one was thus added to all original scores.

Linear mixed models based on Restricted Maximum Likelihood (REML) were used to adjust the infection data and take in account any spatial variation. Best Linear Unbiased Estimators (BLUEs) were computed following a step-wise model selection approach. Baseline models considered included:

- Model 1 - Blocking. Genetic effects are estimated based only on the inter- and intra-block variation recovered from the model.

- Model 2 - Spatial. Only considers global and/or local field trends.

- Model 3 - Spatial + blocking. Combination of the above models.

Models were fitted using Genstat 18th edition (VSN International 2015). For further details on each of the baseline models, the optimisation process and best model selection process, refer to Chapter 3, Section 3.3.2.

5.3.4 Correlation analysis

Pearson correlation coefficients for glume infection across environments were computed to determine the consistency of the phenotype across environments. The coefficient was also calculated to assess the phenotypic relationship between yellow rust infection in the glumes and leaves, across both environments. The 'cor.test()' function in the statistical software R 3.6.1 was used to compute the coefficient tests (R Core Team, 2019).

5.3.5 QTL mapping approach

Similar to the leaf infection phenotypic data, four different statistical methods based on marker allelic state or founder haplotypes were used to characterise QTL underlining glume infection. These are outlined below and further details on each of the methods used and peak marker selection can be found in Chapter 3, Section 3.3.3.

- Single Marker Analysis (SMA, Method 1): Regression analysis on allelic state of 7,369 mapped SNP markers from the MAGIC genetic map (Gardner et al., 2016).

- Haplotype Analysis (HA, Method 2): Regression analysis on founder haplotype probabilities.

- Interval Mapping (IM, Method 3): IM with founder probabilities and no covariates.

- Composite Interval Mapping (CIM, Method 4): CIM with founder probabilities and 10 marker covariates.

For IM and CIM, the empirical $p=0.05$ significance threshold was computed in R/mpMap using the sim.sig.thr function, producing a cut-off at $p=1.6E-05$ (-log10(p) = 4.8) for NIAB16 and $p=3.7E-05$ (-log10(p) = 4.2) for OSG16. For SMA and HA, the package R/qvalue was used to correct for multiple testing, with a q-value threshold of 0.05 (R Core Team, 2017). QTL intervals were defined as a genomic region which, at a given genetic position had a q value of $q \leq 1.6E-05$ (NIAB16) and $q \leq 3.7E-05$ (OSG16).

The genetic versus physical map location of each QTL was determined as described in Chapter 4. Briefly, for each chromosome, the location of SNPs on the genetic map (Gardner et al., 2016) were plotted against their position on the physical map from the IWGSC RefSeq v1.0 genome assembly (Appels et al. 2018) and the location of the most significant SNP for a given QTL highlighted. Data were plotted using R/ggplot2 (Wickham, 2009).

5.4 Results

5.4.1 MAGIC population response to *Pst* glume infection

Genotyped MAGIC lines were assessed for disease severity in the ears in two environments in 2016 (NIAB16 and OSG16), following the observation of significant glume infection in the resistant MAGIC founder Soissons (Figure 5.1). All other founders were found to be relatively resistant, with mean glume infection (GIF) scores ≤1 (Figure 5.2). Among the MAGIC lines however, phenotypic variation was observed over the full range of the assessment scale (Figure 5.3). Nearly 70 % of the lines showed little or no infection in the glumes (GIF = 0) while almost 10 % exhibited *Pst* infection in most of the ear (GIF = 3-4). Upon inspection of the phenotypic relationship between glume infection and the leaf infection data reported in Chapter 3, a positive correlation was found in both environments, with OSG16 showing a stronger correlation than NIAB16 (Figure 5.4, NIAB16: $R = 0.49$, $p < 2.2E-16$; OSG16: $R = 0.62$, $p < 2.2E-16$; based on Pearson's correlation coefficients). MAGIC lines susceptible in the glumes generally exhibited high susceptibility (leaf IF > 60 %). There were a few notable exceptions, particularly at NIAB16, where six lines (5954, 6216, 5770, 5883, 6110, 6374) were found to be moderately resistant on leaves (IF ≤ 25 %) but susceptible in the glumes (GIF > 3.5). One of those lines, 5954, displayed the same phenotypic combination at OSG16.

Statistical analysis of the glume infection phenotypic data using REML showed that the Wald statistic for genotype effect was highly significant (Chi pr < 0.001) at both environments (Table 8.13, Appendix E). High genotypic performance is further supported by high broad sense heritability values (Table 8.13, NIAB16: 0.80, OSG16: 0.82), mirroring values observed for leaf infection ($H^2 = 0.79$-0.94). This evidence points towards stable genotypic performance for glume infection in the MAGIC population at each site. However, the fairly weak correlation of glume infection scored between OSG16 and NIAB16 (Figure 5.5.; $R = 0.47$, $p < 2.2E-16$) highlights possible genetic x environmental interaction.

Figure 5.1 Glume infection in the MAGIC population

A: The MAGIC founder Soissons is highly resistant to yellow rust in the leaves but exhibits glume infection during the mid-to hard-dough growth stage in wheat. B: Close-up of a MAGIC line susceptible to yellow rust in the glumes. The arrows point to the formation of pustules on the glumes, on the inner side of the lemma.
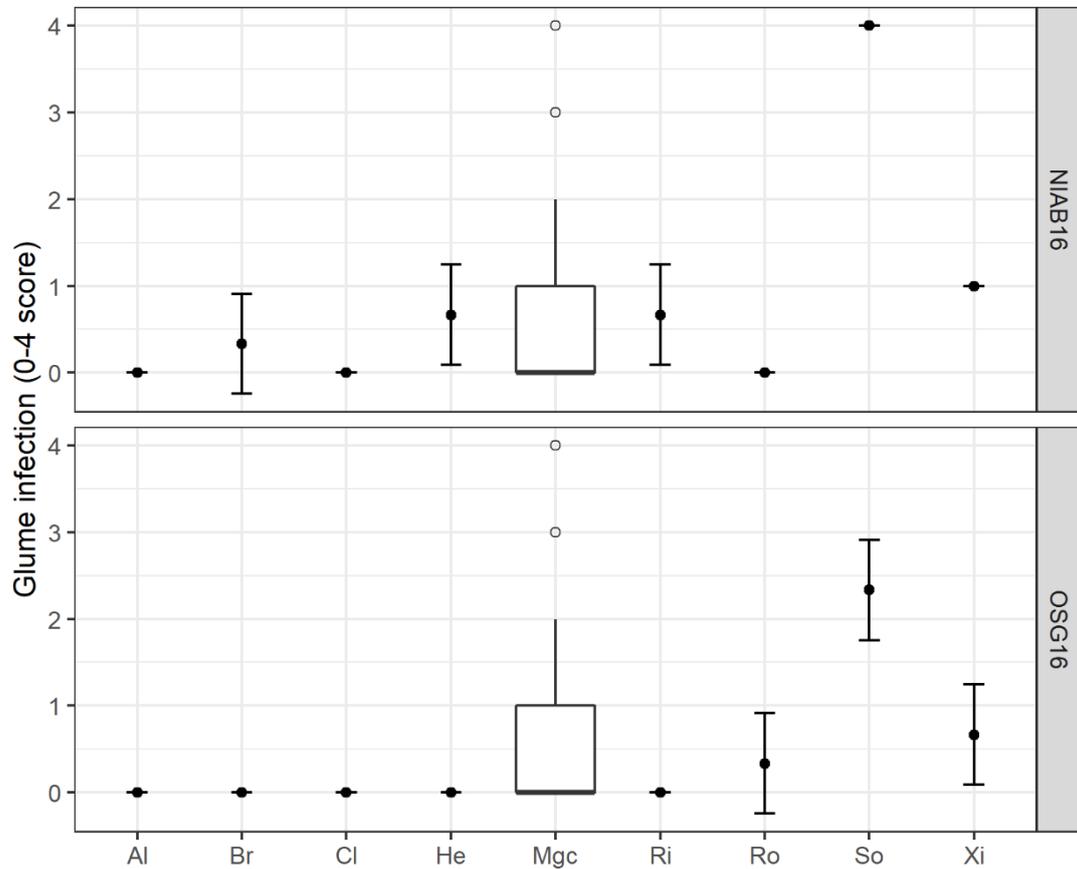
Figure 5.2 Distribution of glume infection for the MAGIC founders relative to the MAGIC population.

Mean glume infection is represented by filled black dots for all parents, with standard deviation plotted as error bars. The central boxplot represents the distribution of the MAGIC population (single counts for unreplicated lines and total counts for replicated lines). Black unfilled circles correspond to outliers. Al: Alchemy, Br: Brompton, Cl: Claire, He: Hereward, Mgc: MAGIC lines, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19. NIAB16 and OSG16 represent the two trial sites from 2016.
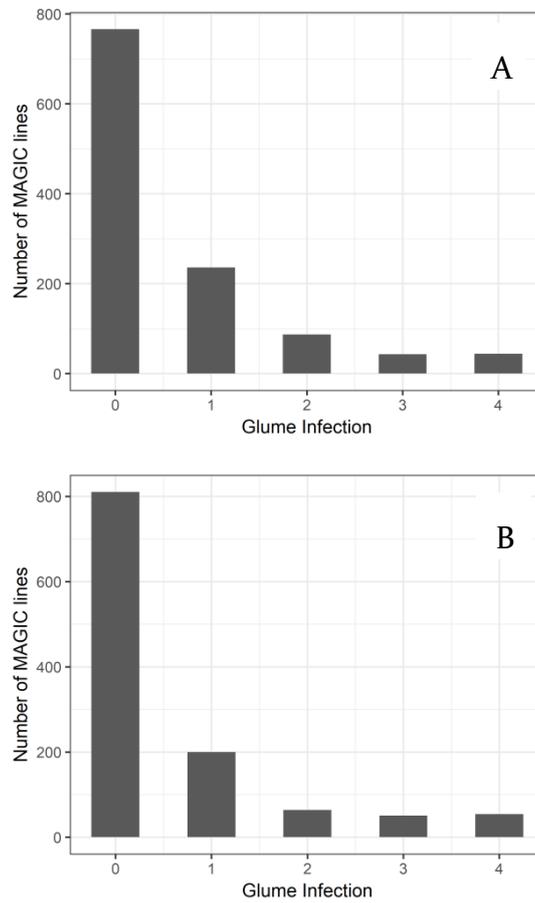
Figure 5.3 Histogram of glume infection data identified a skew towards resistance.

Glume score distribution (glume score = 0 to 4) at NIAB16 (A) and OSG16 (B). The number of MAGIC lines includes single counts for unreplicated lines and total counts for replicated lines.
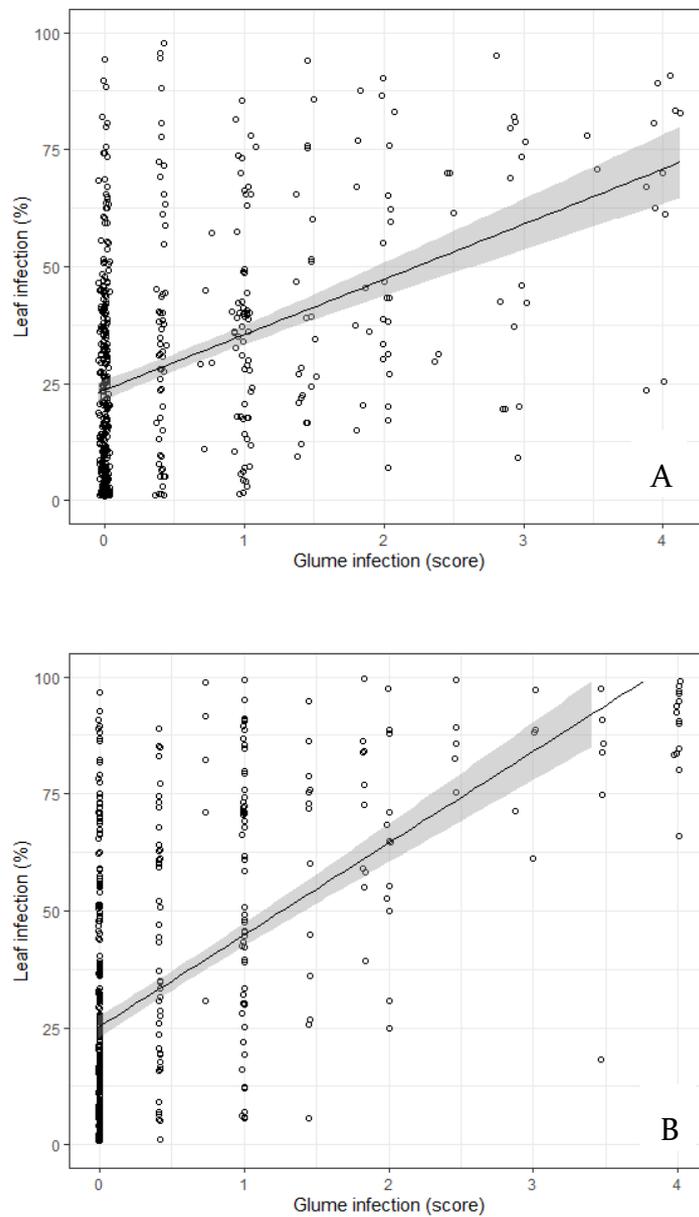
Figure 5.4 Glume infection (score) positively correlates with leaf infection (%).

Correlation between leaf and glume yellow rust infection at NIAB16 (A) and OSG16 (B). Back-transformed adjusted means were used for both scatter plots. The black line is the linear regression line and the grey area the 95 % confidence level interval. Correlation coefficients: NIAB16: $R = 0.49$, $p < 2.2E\text{-}16$; OSG16: R = 0.62, p < 2.2E-16.
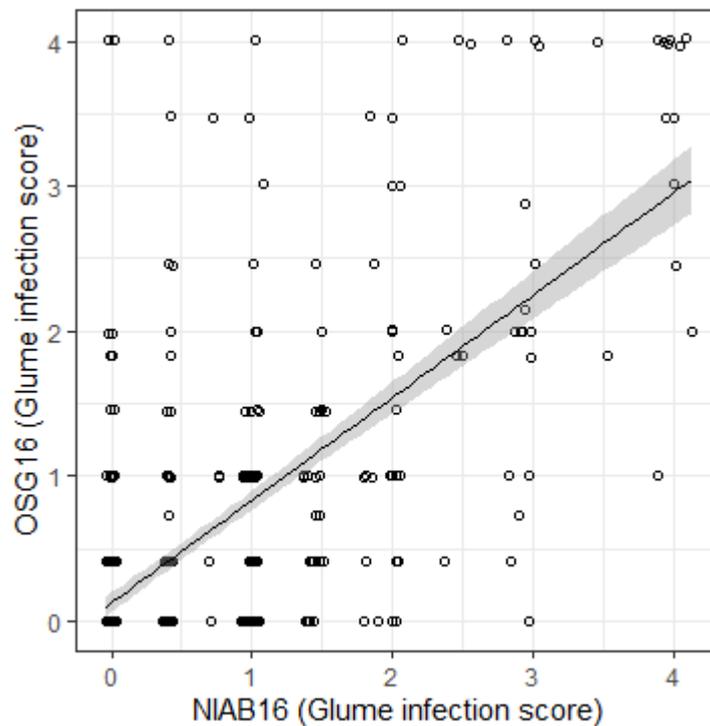
Figure 5.5 Glume infection between trial sites NIAB16 and OSG16 positively correlates.

Scatter plot shows back-transformed adjusted means. The black line is the linear regression line and the grey area the 95 % confidence level interval. Correlation coefficients: $R = 0.47$, $p < 2.2E\text{-}16$.

5.4.2 QTL mapping of YR resistance in the glumes

Five glume YR resistance QTLs were identified, located on chromosomes 2D, 4A, 4D and 5A, as summarised in Table 5.3 and illustrated in Figure 5.6. Marker-by-marker results of all genetic mapping methods are listed in Appendix E. MAGIC founder contributions are highlighted in Table 5.4. The proportion of phenotypic variation explained by each QTL ranged from 3.4 to 6.8 %. Jointly, they explained 23.2 % of the phenotypic variation (average of maximum values). The glume resistance QTLs *QYrg.niab-4D.1* and *QYrg.niab-5A.1* were found to account for the highest percent of phenotypic variation explained (PVE), at 4.8 % and 6 %,

respectivel. Both of these QTLs were also found to confer yellow rust resistance in the leaves (Chapter 3, Section 3.4.2) and were found to locate to the telomeric regions of their respective long arm chromosomes (Figure 5.6). The glume QTL *QYrg.niab-4D.1,* accounting for 4.2% PVE, was found to co-locate with the leaf resistance QTL *QYr.niab-4D.1* and harbour the same peak marker (Figure 4.2 and Figure 5.6). For *QYrg.niab-5A.1* however, the leaf and glume marker are 8 cM apart. Similar values for the percent of the phenotypic variation explained, and the proximity of markers, are nonetheless a strong indication that the leaf and glume QTL both map to the same location on chromosome 5A. Similar to *QYrg.niab-4D.1,* glume QTL *QYrg.niab-2D.1* (accounting for 3.9 PVE %) was found to co-locate with its leaf QTL counterpart *QYr.niab-2D.1* (Figure 5.6). Of the remaining two glume resistance QTL, the peak marker for *QYrg.niab-2D.Ppd_D1* (4.19 % PVE) was Ppd_D1, causative diagnostic marker for the photoperiod insensitive *Ppd-D1a* allele that originates from the MAGIC founder Soissons. This was further confirmed upon examination of MAGIC founder contributions (Table 5.4). Finally, *QYrg.niab-4A.1,* identified in NIAB16 only and explaining 3.4 % PVE, was not found to co-locate with any QTLs for yellow rust resistance in leaves. *QYrg.niab-4A.1* located to the edge of a cluster of SNP markers on the long arm of chromosome 4A, potentially indicating poor marker ordering or structural rearrangements.

Table 5.3 Five glume YR resistance QTLs (5 % significance threshold $p < 1.6E-05$) identified in the MAGIC population in 2016

QTL	Chr	Mapping method		Peak marker	Genetic map position, cM	P value, origin (trial, mapping method)		% PVE		QTL interval, cM
		NIAB16	OSG16			min	max	min	max	
QYrg.niab-2D.Ppd1_D1	2D	3, 4	4	Ppd_D1	60.67	4.61E-07† NIAB16, 3	6.44E-05† NIAB16, 4	.	4.2†	11.88
QYrg.niab-2D.1	2D	4	3, 4	Ra_c21099_1781	197.36	6.70E-06† OSG16, 4	3.40E-04† NIAB16, 4	.	3.9†	10.85
QYrg.niab-4A.1	4A	3, 4	.	BS00067903_51	180.38	.	3.80E-04† NIAB16, 3	.	3.5†	5.53
QYrg.niab-4D.1	4D	1 – 4	2 – 4	D_GDRF1KQ02H66WD_341	125.78	6.32E-08† NIAB16, 4	1.32E-05 OSG16, 4	.	4.8†	.
QYrg.niab-5A.1	5A	1 – 4	2 – 4	BS00065714_51	309.03	9.12E-09 NIAB16, 4	3.98E-09† OSG16, 4	6.2	6.8†	16.97

Underlined QTLs indicate those with co-locating leaf YR QTLs identified in Chapter 3. Mapping methods are numbered 1-4: Single Marker Analysis (1), Haplotype Analysis (2), Interval Mapping (3), Composite Interval Mapping with 10 covariates (4); min and max in superscript indicate the corresponding minimum and maximum p values for each QTL. Genetic map position and QTL intervals are from the MAGIC genetic map (Gardner et al. 2016). Minimum (min) and maximum (max) p values and percentage phenotypic variation explained (% PVE) were selected amongst environments and QTL mapping methods. † PVE % corresponds to indicated p value for each QTL. QTL detected by method 4 only across both environments were not included.

Table 5.4 Predicted founder effects at the five QTL conferring glume YR resistance in the MAGIC population in 2016

QTL	Chr	Founder effects								Origin
		Al	Br	Cl	He	Ri	Ro	So	Xi	Trial, mapping method
QYrg.niab-2D.Ppd_D1	2D	0.09	0.01	-0.05	-0.04	0	0	0.3	NA	NIAB16, 3
QYrg.niab-2D.1	2D	-0.24	0.09	-0.05	0.43	-0.32	-0.26	0.32	NA	NIAB16, 4
QYrg.niab-4A.1	4A	-0.27	-0.18	-0.2	-0.34	-0.33	-0.32	-0.41	NA	NIAB16, 3
QYrg.niab-4D.1	4D	0.23	-0.08	NA	0.32	0.55	0.52	0.4	NA	OSG16, 3
QYrg.niab-5A.1	5A	-0.18	-0.12	-0.15	-0.31	0.07	-0.1	0.13	NA	OSG16, 4

Founder effects presented here are for the consensus marker (outlined in Table 5.5) with the most significant p value identified with QTL mapping methods 3 (IM) and 4 (CIM, cov=10). Al: Alchemy, Br: Brompton; Cl: Claire, He: Hereward, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19. NA: Not Applicable
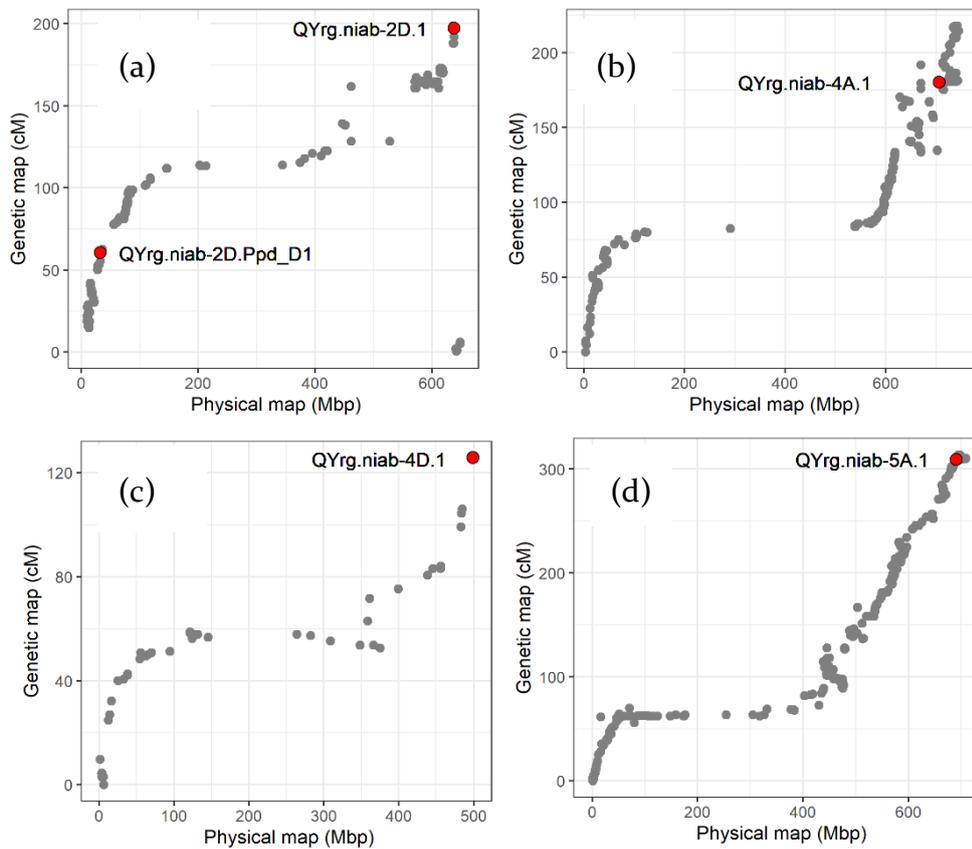
Figure 5.6 Location of the five QTLs conferring resistance against glume yellow rust superimposed onto plots of physical versus genetic map SNP locations for chromosomes 2D (a), 4A (b), 4D (c) and 5A (d).

QTL are indicated by their respective peak markers as listed in Table 5.3.

5.5 Discussion

Yellow rust infection in the ears was explored in the MAGIC population in 2016, following the observation of high infection of the glumes in Soissons, the founder that exhibited the highest foliar resistance to the disease in Chapter 3.

5.5.1 Phenotypic response of the MAGIC population to yellow rust infection in ears

Variation for ear infection was skewed towards resistance, with most MAGIC lines showing little to no yellow rust infection. This pattern of distribution is different to that observed for foliar infection (Chapter 3, Section 3.1.5), which exhibited a wider range of phenotypic variation. This difference in distribution is reflected in the significant, but relatively weak, correlation observed between the leaf and ear infection phenotypes. Assessing the extent of yellow rust infection in ears is more difficult compared to assessing foliar infection as disease symptoms are more easily and rapidly identified on leaves. In addition, the sometimes bleached appearance of florets as a result of ear infection can be confused with other disease symptoms like that of Fusarium head blight in the initial stages. The degree of difficulty in phenotypic assessment had to be balanced with the practicalities of scoring large numbers of plots in the field. Therefore, a simple 0-4 scale was developed, and the results obtained compared to the percentage scale used for scoring leaf yellow rust. The simpler scale used for glume infection will inevitably have grouped different levels of response together. This may have had an effect on the correlation between leaf and glume infection, and on the power and precision to identify QTL. The correlation between the leaf and the ear phenotype was more pronounced in OSG16 (R = 0.62) than in NIAB16 (R = 0.49). The >10 % difference in correlation between these environments was perhaps due to a climatic difference, at the OSG16 trial was in a cooler and wetter region than was NIAB16. This will have likely led to a longer window of infection for urediniospores, making disease symptoms more pronounced and easier to score later in the season, thus resulting in a stronger correlation between leaf and ear infection.

Leaf and ear infection not only differed in phenotypic distribution, but also in broad sense heritability. Yellow rust infection in ears exhibited over 10 % lower

heritability values in both environments tested, indicating that additional non-genetic factors have a larger impact on the phenotypic response in ears than in leaves, and/or that the limitations of the 0-4 glume infection scale combined with the inherent difficulty in accurately scoring glume infection impacted phenotyping accuracy. This difference in heritability is in agreement with a recent yellow rust resistance study in Triticale that also reported a lower heritability value in ears compared to leaves (Losert et al., 2017). Furthermore, the Losert et al. (2017) study argues that the severe ear infection observed in Triticale may have been a feature of the prevalent *Pst* 'Warrior' race and its potential ability to develop at warmer temperatures. Although temperature adaptation of the 'Warrior' race variants have not been experimentally confirmed, as they have for the aggressive race *Pst*S2 for example (Milus et al., 2009), it is plausible that warmer temperatures not previously conducive of germination and subsequent colonisation enable the pathogen to prolong its growth into remaining photosynthetic tissue i.e. the glumes (Mboup et al., 2012).

2016 was a year that saw a large number of yellow rust infected leaf samples received by the UK Cereal Pathogen Virulence Survey, with some samples received as early as January (A. Hubbard et al., 2017). Yellow rust was observed early on in the trial but it would seem that the disease pressure reported by the UKCPVS was not as pronounced in the 2016 MAGIC pathology trials (Chapter 3, Section 4.1.1). The fact that variation in ear infection was still observed would indicate that, in addition to inoculum levels, other factors come into play. Weather conditions conducive of *Pst* spore development (high humidity and cool temperatures) coupled with the onset of ear emergence up to anthesis would have also had to have been in tune. Overall, these observations are in line with reports describing that yellow rust infection in the ears is prominent when significant levels of inoculum are present and are coupled with suitable growing conditions and timely flowering (Cromey, 1989a; Wellings, 2003).

Ear infection by *Pst* commonly affects susceptible wheat varieties. However, yellow rust ear infection of relatively resistant wheat varieties has been reported in New Zealand (Cromey, 1989a). This phenomenon was observed in the MAGIC founder Soissons, as well as a subset of MAGIC lines across both environments tested (leaf

infection ≤ 25% coupled with ear % infection≥ 3.5). This difference in phenotypic response to Pst could be due to several factors. Cromey (1989a) reported environmental variation for this trait, which was also the case here. Taking into consideration that this phenotype was distinctively observed in the MAGIC founder Soissons, it has the potential to be genetically inherited in the MAGIC population, rather than being a result of environmental factors. Further work would be required to determine which of these hypotheses holds true. It should be noted that Soissons is very early flowering, due largely to the photoperiod insensitive allele it carries at the major flowering time locus. The importance of flowering time is discussed further in Section 5.5.2 below.

5.5.2 QTL mapping of yellow rust resistance in wheat ears

Following the interrogation of the glume infection data in relation to foliar resistance, the genetic basis of yellow rust resistance in ears was characterised by genetic mapping in the MAGIC population. Each QTL identified was then compared to the yellow rust leaf resistance QTLs identified in Chapter 3. I also take into consideration morphological and phenology traits and discuss their potential role in the context of other wheat ear fungal diseases.

Five QTL conferring yellow rust resistance in the ears were identified in the MAGIC population, explaining 23.2 % of the phenotypic variance in total. *QYrg.niab-5A.1* (6.2-6.8 % PVE) and *QYrg.niab-4D.1* (4.8 % PVE) were the main contributors. Both these QTLs either had the same peak marker or mapped close to their leaf resistance QTL counterparts, which also were small effect QTLs (Chapter 3). *QYrg.niab-4D.1* shared the same peak marker as *QYr.niab-4D.1*. The peak marker for *QYrg.niab-5A.1* was located < 10 cM away from that of *QYr.niab-5A.1* and thus likely represents the same source of yellow rust resistance. A common peak marker was also observed for *QYrg.niab-2D.1* and its leaf counterpart *QYr.niab-2D.1*. Taken together, this evidence is a strong indication that, in the MAGIC population in 2016, three QTL conferred yellow rust resistance in both leaves and ears. Another interesting finding was the absence of glume YR resistance QTL corresponding to the leaf YR resistance QTLs *QYr.niab-1A.1, QYr.niab-2A.1* and *QYr.niab-2B.1*, despite their high significance and PVE values for leave resistance, and their

consistent expression across environments. The leaf resistance QTL *QYr.niab-2B.1* was possibly not expected to be identified as being involved in glume resistance, since resistant alleles at this locus originate from Soissons, which was found to be particularly susceptible to yellow rust in ears. However, the susceptibility of Soissons to YR glume infection is controlled by multiple loci and therefore additional susceptibility QTLs will be at play. The observation that three robust QTLs for YR resistance in the leaves are not implicated in glume YR resistance is perhaps an indication that the resistance mechanisms they control are not relevant to ears.

Finally, of the seven QTL detected in this study, one QTL was found to be linked to flowering time, indicating that phenology may play a role in glume YR resistance. The peak marker for *QYrg.niab-2D.Ppd1_D1* is located in the photoperiod response gene *PPD-D1*, with the early flowering *Ppd-D1a* allele carried by the founder Soissons conferring susceptibility to glume YR infection. Soissons is known to be the earliest flowering of the MAGIC founders. Based on this evidence, it could be hypothesised that the ears of earlier flowering MAGIC lines are more prone to yellow rust because they are exposed to *Pst* inoculum for a longer period of time than later flowering lines. It was therefore expected that the subset of six MAGIC lines that combined glume susceptibility and foliar resistance would be early flowering like Soissons. However, none of them carried the early flowering Soissons allele at the *PPD-D1* locus. Nevertheless, further evidence of a link between yellow rust ear infection and flowering time was provided by the co-location of *QYrg.niab-4A.1* with the previously identified multi-year and multi-location flowering time QTL *QFt.niab-4A.1,* identified in the same MAGIC population (Mackay et al., unpublished). The involvement of flowering time on glume YR infection was somewhat expected, as this and other morphological traits are often associated with other fungal diseases of the wheat ears such as Fusarium (Gervais et al., 2003) and Septoria nodorum Glume Blotch (SGB, (Aguilar et al., 2005; Scott, Benedikz, & Cox, 1982). To account for such traits and improve the detection of true resistance QTL, a previous study on SGB used spore suspensions at a similar developmental stage to inoculate flag leaves and glumes (Shankar et al., 2008). This effectively reduced the pleiotropic effect of flowering time and

height that usually confound QTL analyses of SGB resistance (Aguilar et al., 2005). This phenotyping strategy could easily be envisaged and tested for wheat yellow rust, with adult plants grown in controlled environment conditions. QTLs conferring SGB resistance in leaves and glumes have been suggested to be independently inherited (Aguilar et al., 2005; Bostwick, Ohm, & Shaner, 1993; Nelson & Gates, 1982; Shankar et al., 2008). This mode of inheritance could potentially also apply to wheat yellow rust resistance despite the opposing nature of host-pathogen interactions between both diseases (Tan et al., 2010). In light of the importance of morphological and phenological traits in other wheat ear diseases and the early flowering phenotype of the MAGIC founder Soissons, it would have been beneficial to have phenotyped flowering time in the MAGIC glume infection study presented here. However, practical limitations did not make this possible. Height was assessed but not included as a covariate in the QTL analysis. Although height has been shown to be correlated with Septoria Nodorum Glume Blotch (Aguilar et al., 2005), this has not been the case for yellow rust (Maccaferri et al., 2015).

5.5.3 Future work

Evidence presented in this study provides some initial clues as to whether leaf and glume yellow rust resistance should be considered as independent traits in breeding programmes. Further work is however required to dissect the importance of morphological traits and explore the use of genomics tools to gain a better understanding of the differences between leaf and glume yellow rust resistance. A number of next steps can be considered:

(1) Assessment of morphological traits in controlled environment experiments. A small subset of MAGIC lines (approximately 300 lines randomly selected) could be selected to determine the correlation between yellow rust resistance in ears and morphological traits including height and heading date. Stomatal density in ear tissues could also be explored as this is an important photosynthetic trait and is also the means of entry of *Pst* urediniospores into plant tissue.

(2) Validation of QTL in independent experimental populations. The Soissons x Beaver bi-parental population (Kumar et al., 2011) and/or the WAGTAIL association mapping panel (Gardner, Cockram, unpublished) represent examples of initial steps towards independent validation of QTL.

(3) Near isogenic lines. These are currently under construction for the MAGIC leaf resistance QTLs. YR QTL NILs could be phenotyped to confirm whether leaf and glume YR resistance are two independently inherited traits, on a QTL by QTL basis.

(4) Comparative RNASeq and other transcriptomics studies. Exploring differential gene expression patterns of yellow rust resistance between leaves and glumes will shed light on any differing molecular mechanisms underlying resistance between both plant tissues. This could most usefully be done using the near isogenic lines listed above.

(5) Spread of glume infection. While it is assumed that the *Pst* urediniospores that infect wheat ears originate from inoculum produced in surrounding canopy or the plant itself, it is lesser known whether infection may arise from hyphae developing in the leaves and travelling up to the ear. A microscopy experiment with fluorescent tagged *Pst* isolates infection an adult wheat plant would enable the visualisation of such a phenomenon.

5.6 Conclusion

Ear photosynthetic capacity plays a significant role in grain filling, particularly under stress conditions. With temperatures predicted to rise significantly during the crucial developmental stages that are head emergence and flowering, pests and diseases pose an even greater threat. While ear yellow rust infection may not be as much of a recurrent problem as leaf infection, it can become detrimental to yields provided suitable growing and environmental conditions. Despite this, the genetic basis of ear yellow rust resistance ears has not been sufficiently researched to date.

The QTL mapping study undertaken here has provided a starting point towards the elucidation of the genetic components underlying yellow rust resistance in glumes, how it relates to leaf resistance QTL, and demonstrated that phenology traits also have pleiotropic effect. The knowledge of which QTLs are relevant to both leaf and glume resistance, along with the genetic markers tagging these traits identified here, will help inform marker-assisted breeding approaches best suited for wheat varieties for growth in agricultural environments in which glume *Pst* infection is prevalent.

6 GENERAL DISCUSSION

Yellow rust is one of the most detrimental diseases of wheat, causing recurrent yield losses in all wheat-producing regions globally (Wellings, 2011). At the same time, yellow rust is also one of the most challenging wheat diseases to control, with current methods relying on the growth of resistant varieties and fungicide application. Of these, the primary control method is growing resistant varieties, a method that is sustainable and saves costs on fungicide use. There is therefore a need to identify genes conferring different types of resistance and to develop suitable tools to track beneficial alleles at these loci in breeding programmes. This is especially true due to the changing nature of two key factors:

(1) New regulations restricting the use of fungicides: fungicide application is effective, but it is also expensive and limited by weather conditions. However more recently, the use of many effective fungicides is becoming increasingly restricted via legislation, due to environmental concerns.

 (2) A diversifying and dynamic *Pst* population. *Pst*'s adaptability given the right selection pressure, and ability to disperse over long distances, have resulted in outbreaks and epidemics of regional, national and continental significance. Recent studies have shown that exotic incursions of genetically diverse *Pst* races into Europe have largely replaced the more clonal *Pst* populations that preceded them (Hovmøller et al., 2016; Hubbard et al., 2015). This, combined with the recurrent yellow rust epidemics of the past 60 years, have shown that deploying varieties largely underpinned by major gene resistance is not an effective long-term strategy.

A more durable and resilient approach is one that is based on the combination of different yellow rust resistance (*Yr*) genes governed by major and minor resistance, with adaptation to local and regional *Pst* populations. It should be noted that combining several *Yr* genes is not new and is in fact a well-established approach in CIMMYT's global wheat breeding programme for brown and yellow rust (Singh et

al., 2005). Nevertheless, with *Pst* populations constantly evolving and the threat of changing regulations in fungicide use, host resistance must keep up and continue to provide breeders with new and different options for the development of resilient wheat varieties. To this end, additional sources of resistance must be identified, along with adequate molecular and genomic tools to aid effective deployment in breeding programmes and to increase our understanding of the underlying molecular mechanisms of yellow rust resistance. To date, while over a hundred yellow rust resistance loci have been identified, including more than 80 different *Yr* genes, not all of them are effective against the *Pst* races that have emerged over the past two decades and only eight have been cloned (*Yr36* (Fu et al., 2009), *Yr18/Lr34* (S. G. Krattinger et al., 2009), *Yr10* (Liu et al., 2014), *Yr46/Lr67* (Moore et al., 2015), *Yr7, Yr5, YrSP* (Marchal et al., 2018), and *Yr15* (Klymiuk et al., 2018)).

In this thesis, I used the NIAB Elite MAGIC population in conjunction with different genetic and genomic approaches to identify and characterise genetic loci controlling resistance to yellow rust in adult wheat plants. Findings and conclusions are summarised in the below 6.2 section. I reflect on those findings in subsequent sections, discussing them in the context of current research on wheat disease resistance (Section 6.3), as well as current wheat genetic and genomics resources and the relevance of the results to wheat breeding (Section 6.4). Finally, I conclude with some suggestions on the direction of future research efforts based on the work presented in this thesis (Section 6.5).

6.1 Summary of key findings

6.1.1 Characterising QTLs conferring YR resistance in wheat leaves

- The MAGIC population exhibited a range of responses to *Pst* infection that were highly heritable (broad sense heritability = 0.77-0.94 %) in all trials and locations.

- Approximately 10-13% of the MAGIC population remained highly resistant (<10 % infection) throughout the infection season, in both 2015 and 2016. Of those, 20 lines showing no signs of YR infection across environments (site x year).

- Four strong-effect QTLs, *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B.1* and *QYr.niab-2D.1,* were detected consistently across environments and altogether explained nearly 50 % of the phenotypic variation.

- Five small-effect QTLs, *QYr.niab-2A.2, QYr.niab-3A.1, QYr.niab-6A.1, QYr.niab-6A.2* and *QYr.niab.6B.1*, were also detected, cumulatively explaining 15-20 % of the phenotypic variation, with slightly less consistency in detection with the different mapping methods used, across YR scores and environments.

- *QYr.niab-3D.1, QYr.niab-4B.1, QYr.niab-4D.1, QYr.niab-5A.1, QYr.niab-6A.3* were borderline QTLs. Their detection likely fell short due to the masking from strong effect QTLs.

- The phenotypic effects of different allelic combinations between *QYr.niab-1A.1, QYr.niab-2A.1, QYr.niab-2B, QYr.niab-2D.1* were investigated. Combinations were more effective in conferring resistance to YR than single QTL in isolation, a strong indication of interaction effects between QTLs in the MAGIC population. Additivity was observed in the majority of combinations.

- All MAGIC founders except Robigus, the most susceptible, contributed towards yellow rust resistance either singly or in different combinations, revealing the untapped resistance potential of some founders.

6.1.2 Accelerating candidate gene characterisation with genome annotations

- Five QTL were characterised by NBS-LRR clusters within their intervals (*QYr.niab-2B.1, QYr.niab-2D.1, QYr.niab-3A.1* and *QYr.niab-6A.2*), or were in close proximity (*QYr.niab-2A.1*), indicating that these loci are likely to confer major, race-specific resistance against yellow rust.

- There is potential evidence for effector triggered immunity based on indirect recognition within the NIAB Elite MAGIC population, as indicated by the presence of NBS-LRR-encoding genes with integrated domains within the physical intervals of *QYr.niab-2B.1* (BED domain) and *QYr.niab-3A.1* (thioredoxin domain).

- *QYr.niab-2B.1* located to a genomic region dense in NBS-LRR, which includes the recently cloned *Yr7* and *Yr5/YrSP*.

- Yellow rust resistance in the MAGIC population may also be underpinned by non-NBS-LRR-encoding genes, as exemplified by the absence of NBS-LRR candidate genes in the physical intervals of *QYr.niab-1A*.1, *QYr.niab-2A.1* and *QYr.niab-6A.1* of the wheat reference Chinese Spring genome.

6.1.3 Characterising QTLs conferring YR resistance in wheat ears

- This study represents the first genetic characterisation of yellow rust resistance in wheat ears. The trait exhibited a high heritability of 72 %.

- Five QTL explained between 3.4% and 6.8% of the phenotypic variation observed in the NIAB Elite MAGIC population.

- *QYrg.niab-2D.1, QYrg.niab-4D.1* and *QYrg.niab-5A.1* co-located with QTLs for leaf yellow rust resistance identified in Chapter 3 (*QYr.niab-2D.1, QYr.niab-4D.1* and *QYr.niab-5A.1* respectively).

- *QYrg.niab-2D.Ppd_D1* and *QYrg.niab-4A.1* were linked to flowering time QTL, indicating that earlier ear emergence potentially results in increased susceptibility to yellow rust in wheat ears.

6.2 How far have we come in dissecting the genetics of YR resistance?

The genetic inheritance of yellow rust resistance was demonstrated at the turn of the last century (Biffen, 1905). Since then, cytogenetic studies enabled the first *Yr* genes to be characterised (Lupton & Macer, 1962) and advances in QTL mapping and experimental population development have led to the cataloguing of nearly 100 *Yr* genes and identification of over 200 QTL (Maccaferri et al., 2015; Rosewarne et al., 2013). These genetic resources have formed the basis for the discovery of different types of host resistance. In this section, I discuss my findings, more specifically candidate gene annotations and glume infection, in relation to some recent publications.

Today, we know that yellow rust resistance is governed by different types of genes. These broadly fall into two different categories. The first category is based on NBS-LRR-encoding *R* genes, which largely confer major, mostly race-specific resistance, expressed at both the seedling and adult-plant stages and elicit effector-triggered immunity in the host (Jones & Dangl, 2006; Liu et al., 2014; Marchal et al., 2018). In wheat, these genes are found in high abundance and tend to cluster in the distal parts of chromosomes (Appels et al., 2018). The recent meta-analysis of over 300 cloned *R* genes by Kourelis & van der Hoorn (2018) identified nine different molecular mechanisms underlying plant immunity, highlighting how diverse those mechanisms are. This diversity is reflected in the candidate genes identified in Chapter 4. Of the eight QTL selected for further investigation, five were defined by NBS-LRR clusters within their physical interval. Three of those QTL were annotated with candidate genes that appeared to encode NBS-LRR genes with integrated domains (*QYr.niab-2B.1, QYr.niab-3A.1* and *QYr.niab-6A.2*). Research is only starting to emerge as to what the role of these integrated domains exactly is. One notable example comes from the recent cloning of *Yr7* and *Yr5/Sp*, which both encode NBS-LRR proteins with BED domains (Marchal et al., 2018). Interestingly, the genomic region for *QYr.niab-2B.1* contained *Yr7* and *Yr5/YrSP*, indicating that these loci may be in current use in UK wheat varieties.

The second category is based on all other genes that do not encode NBS-LRR proteins. Most of the recently cloned *Yr* genes fall into this category and have shed

some light on the underlying molecular mechanisms driving resistance. There are two instances in particular that I would like to discuss here. Firstly, the case of *Yr18/Lr34* and *Yr46/Lr67,* which encode ABC and hexose transporters respectively (Krattinger et al., 2009; Moore et al., 2015). Krattinger & Keller (2016) hypothesise the potential 'artificial creation' of durable resistance via targeted mutagenesis, following the conclusion that both genes likely evolved following the domestication of wheat. This presents a novel and exciting avenue of research for durable resistance. Secondly, the case of *Yr36* and *Yr15,* which both encode kinases (Fu et al., 2009; Klymiuk et al., 2018).

In Chapter 5, candidate gene annotations for three of the eight QTL suggested that these may not governed by NBS-LRR genes (*QYr.niab-1A.1, QYr.niab-6A.1, QYr.niab-2A.1*), based on the reference genome assembly of Chinese Spring. Even though further work would need to be done to identify and clone each of the causal gene, this is an indication that non-race specific, adult plant resistance is potentially segregating in the NIAB Elite MAGIC population and thus merits further investigation. Overall, this group of genes is particularly interesting to pursue from a breeding perspective because they tend to provide durable (*Sr2* and *Lr34* have been resistant against most races of stem rust and leaf rust respectively) and/or non-race specific yellow rust resistance at the adult plant stage, with resistance sometimes extending to multiple pathogens like is the case for *Yr46/Lr67* and *Yr18/Lr34,* which both confer resistance to stem rust and powdery mildew (Herrera-Foessel et al., 2014; Singh, 1992; Spielmeyer et al., 2005).

The vast majority of genetic studies on yellow rust resistance have focused on foliar resistance specifically. However, leaves are not the only tissue that *Pst* urediniospores have the ability to colonise. Ears can also become infected and if left uncontrolled, this can have a significantly detrimental impact on yield (Purdy & Allan, 1965). Despite this, glume infection is a trait that has received relatively little attention, other than reports of infection during severe epidemics (Cromey, 1989b; Wellings, 2003, 2009). The work reported in Chapter 5 therefore represents an important first step in filling in this gap. The glume infection scenario provides the opportunity to explore some interesting research questions. Two are of particular interest. (1) Is glume resistance governed by the same disease resistance

mechanisms as in leaves? In chapter 5, I showed that wheat ears do exhibit yellow rust resistance and that this is governed by QTLs also expressed in leaves, suggesting that this is indeed probably the case. (2) Are foliar and glume yellow rust resistance two independently inherited traits, like it has been demonstrated in other wheat ear diseases? In Chapter 5 I found that while glume-specific YR resistance QTL were identified, these co-located with known flowering time QTL indicating a pleiotropic effect of the timing of ear emergence and development on YR infection in the glumes. The identification of the photoperiod insensitive allele at the *Ppd-D1* locus as a glume YR resistance locus is further discussed in Section 6.5.

6.3   The NIAB Elite MAGIC population as a resource for the wheat research and breeding communities

Based on analysis of SNP markers, the NIAB Elite MAGIC population used in this thesis captures approximately 80 % of the genetic variation present in the WAGTAIL panel, which is representative of UK elite commercial varieties (Gardner et al., 2016). As the genetic analysis undertaken was conducted in UK-relevant germplasm, the QTLs and alleles identified can easily be selected for within germplasm currently in use in ongoing breeding programmes, given the development of appropriate molecular markers from the SNPs identified.

Another advantage of using the NIAB Elite MAGIC population lies in its strong baseline of genetic, genomic and statistical tools and resources that have been developed. The NIAB Elite MAGIC high-density genetic map generated by Gardner, Wittern, & Mackay (2016) provided the foundation for the QTL analysis undertaken in this thesis. The map was based on genotyping the eight founders and a subset of 643 MAGIC lines with the iSelect 90K SNP array. Since then, the remainder of the MAGIC lines have been genotyped using this array (Cockram & Gardner, unpublished), and their incorporation into future studies will provide additional power and precision when mapping QTLs. Statistical methods for genetic analysis based on founder probabilities, such as the R/mpMap package (Huang et al 2011), have been developed specifically for analysis multi-parental populations, and these were implemented in Chapters 3 and 5 for conducting

haplotype-based genetic analyses. This particular method, in combination with a single-marker approach, provides a robust starting point for genetic analysis and has been implemented in several wheat MAGIC QTL mapping studies (Downie et al., 2018; Huang et al., 2012; Sannemann et al., 2018). Additional resources such as founder genome assemblies, genotyping the population via skim sequencing, and the generation of NILs via exploitation of residual heterozygosity in the MAGIC RILs, are being developed for the NIAB Elite MAGIC population. Their use is discussed in Section 6.5.

The community spirit is at the heart of the widespread use of MAGIC populations for dissecting complex traits, and this is perhaps best captured by recently launched initiatives such as the MAGIC-WHEAT consortium comprising of scientists and breeders (Sannemann et al., 2018), or the MAGIC workshop held in 2019, which presented the latest resources and advancements in MAGIC populations for different crops (http://mtweb.cs.ucl.ac.uk/mus/www/MAGICdiverse/MAGIC_workshop.htm). Indeed, the use of MAGIC populations is becoming increasingly commonplace. For example, at the onset of this thesis, only two wheat MAGIC populations were available: a 4-founder MAGIC derived from Australian varieties (Huang et al., 2012), and the 8-founder NIAB Elite MAGIC population constructed from varieties grown in the UK (Mackay et al., 2014). NIAB Elite MAGIC was therefore the only available MAGIC population of direct relevance to UK wheat improvement. More recently, three additional wheat MAGIC populations have been developed. Two originate from Europe and were developed predominantly using German wheat cultivars selected for their advantageous agronomic attributes and relevance to breeding programmes (Sannemann et al., 2018; Stadlmeier et al., 2018). The third one is predominantly based on spring wheat cultivars originating from six different countries (China, Mexico, USA, Australia, Canada, Israel), making it the most international wheat MAGIC population generated thus far (Shah et al., 2019). These five populations encapsulate the genetic diversity of over 30 founders. This represents an unprecedented opportunity to compare the genetic basis of complex quantitative traits of interest in a wide array of genetic backgrounds. For example, any commonalities in the genetic basis of resistance to different biotrophic fungi

could be investigated by comparing the yellow rust resistance data collected in this study to the powdery mildew resistance data generated in one of the German MAGIC populations (Stadlmeier et al., 2019).

Taken together, the strengths of the MAGIC population used in this study ultimately lie in its direct relevance to UK breeding germplasm, the comprehensive set of community-focused analytical tools and genomic resources that have been developed or initiated, and the possibility to explore multi-QTL and multi-trait interactions; all in a single population.

6.4 Future work: how best to continue the delivery of molecular tools for the development of wheat varieties with more effective yellow rust resistance?

A number of exciting avenues of research are now possible to better understand the genetic basis of yellow rust resistance and further characterise genetic resources of use for deployment in wheat breeding programmes. In a first instance, the NIAB Elite MAGIC population should be further explored thanks to the additional genomic and population resources that are undergoing development (Section 6.5.1). Secondly, it is necessary to move beyond the MAGIC population and take a multi-population and multi-disciplinary approach towards developing durable yellow rust resistance (Section 6.5.2).

6.4.1 Further genetic analysis of yellow rust resistance using the NIAB Elite MAGIC population

(1) *Exploitation of historical disease data*. Yellow rust resistance was assessed in the MAGIC population in the 2012 and 2014 seasons, via phenotyping of the MAGIC nursery plots grown for seed multiplication at the adult plant stage. Both of these periods shortly follow the 'Warrior' race incursion in the UK and should therefore provide valuable insight into host response to key *Pst* population changes. It should however be noted that these nurseries were unreplicated and fungicide treated, and so QTL characterisation is likely to prove less powerful than in the pathology trials used in this study.

(2) *Fine-mapping and candidate gene analysis with tailored resources*. There are two key MAGIC resources that could be exploited to identify the genes and allelic

variants underlying the QTLs identified in this study: Near Isogenic Lines (NILs) and founder genome assemblies. NILs can be developed for each of each QTLs. One rapid way to do this in self-pollinating crops is to generate Heterogeneous Inbred Families (HIFs), exploiting the residual heterozygosity typically present in a population segregating for a specific QTL (Tuinstra et al., 1997). Briefly, HIFs are developed via the following steps: (1) Identification of residual heterozygosity across a target QTL using the 90k SNP genotypic data available for the F5 MAGIC RILs, (2) identifying a subset of these RILs for which the region of heterozygosity present is predicted to originate from founders carrying alleles of contrasting effect, (3) development of SNP markers to screen selfed progeny of the F5 line for individuals that are homozygous for each of the two alleles present as a het in the F5 individual, (4) selfing these individuals to produce a bulk of seed. These two lines represent a HIF pair. This process has now been initiated for the QTLs identified in this study. These HIF resources open the door for detailed genetic studies: (a) Effects of a single QTL can be explored in an isogenic background, so that interesting resistant phenotype can be studied at the microscopic and molecular levels, with the aim to determine their underlying immunity mechanisms. (b) A HIF pair can be crossed to generate recombination across the locus of interest, thus allowing further fine-mapping to be undertaken, ultimately to a resolution that enables the identification of the causal gene and genetic variant. This process will be greatly aided by genome assemblies of MAGIC founders, the second key resource. These will provide the basis for gene annotations tailored to the MAGIC population, and the identification of allelic variants for the gene models identified. Genome assemblies for the MAGIC founders Claire and Robigus are now available (www.earlham.ac.uk/grassroots-genomics), and those for the remaining six founders are under construction as part of the 'Wheat MAGIC Pangenome' project (BBSRC projects BB/P010741/1, BB/P010733/1, BB/P010768/1). Taken together, these resources will provide robust resources to underpin future fine-mapping and map-based cloning of the yellow rust resistance QTLs identified in this study.

(3) *Determining the influence of phenology and morphology on host response to Pst*. Traits like height and flowering time have been shown to be associated with

resistance to wheat ear diseases (Gervais et al., 2003; Schnurbusch et al., 2003; Scott et al., 1982). For foliar yellow rust resistance, the potential confounding effect of flowering time and height have been considered in GWAS studies conducted by Maccaferri et al. (2015) and Bulli et al. (2016). Both studies identified markers significantly associated with those traits ($P < 0.05$), but these were removed from the final GWAS results. Such an approach fails to take into account the pleiotropic effects that these markers may potentially have. In Chapter 5, I reported two ear YR resistance QTL that co-located with a previously reported flowering time QTL identified in the same population, strongly indicating that the timing of ear emergence has an impact on resistance or susceptibility to yellow rust infection in the ear. Analysis of the historical data available in the MAGIC population on flowering time and height, together with the morphology and phenology data collected over the course of this study, would allow the relationship between agronomic traits and yellow rust resistance to be examined in the NIAB Elite MAGIC population.

(4) *Development of a MAGIC disease resistance resource*. Over the course of the growing season in the field, wheat is simultaneously and/or sequentially exposed to several different diseases. Investigating the underlying genetic mechanisms of host resistance to individual pathogens, and how different diseases and host resistance pathways interact with each other will provide a better understanding of which pathways and/or mechanisms to target for the development of wheat varieties resistant to multiple diseases. Such an approach would allow the identification of resistance genes potentially effective against several pathogens, like *Yr18/Lr34*, and merits further exploration. Developing an all-encompassing MAGIC disease resource could use the YR datasets generated here, and be built up using datasets generated for other diseases in the same population notably for the hemi-biotrophic fungal diseases Stagonospora nodorum blotch caused by *Parastogonospora nodorum* (Downie et al., 2018; Lin et al., 2019) and tan spot caused by *Pyrenophora tritic-repentis* (Lin et al., 2020). The work presented here is a first step towards building such a resource, that ideally would also include resistance data for the other two main rust diseases of wheat, brown rust and stem rust.

6.4.2 Exploitation of the latest population and genomic tools and resources

Taking a multi-population and multi-disciplinary approach is key to tackle the dynamic and genetically diverse *Pst* population that has now come to dominate North Western Europe, following the incursion of the 'Warrior' and 'Kranich' races in 2011 (Hovmøller et al., 2016; Hubbard, Lewis, et al., 2015). This requires an equally dynamic response from researchers and breeders, so that current levels of yellow rust resistance in elite varieties can be improved and become more robust. Here, two principle approaches are discussed that both involve the NIAB MAGIC Elite population.

(1) *Combining forces to combat changing Pst populations, for the benefit of European wheat disease resistance breeding*. YR resistance QTL and molecular marker resources from the NIAB Elite MAGIC population identified in this thesis, and similar datasets for the WAGTAIL wheat association mapping panel (Gardner, Cockram et al., unpublished) are now being utilised in the BBSRC and industry-funded research project 'Yellowhammer' (https://gtr.ukri.org/projects?ref=BB%2FR019231%2F1) which started in 2018 and will continue over the next three years. Devised in close collaboration with the North European wheat breeding industry, this project provides a comprehensive strategy to explore resistance gene combinations and identify which would be most effective against the various lineages present in the European *Pst* population. A complementary approach could be envisaged that exploits the recently developed European MAGIC populations (Sannemann et al., 2018; Stadlmeier et al., 2018), allowing further analysis of relatively large numbers of founders within these structured genetic mapping populations.

(2) *Exploiting the latest genomic technologies for the rapid cloning of yellow rust resistance genes*. Genomic technologies and gene cloning approaches have rapidly moved on in recent years for wheat. Traditional map-based cloning approaches now being complemented with direct cloning approaches based on mutational genomics and next generation sequencing (Bettgenhaeuser & Krattinger, 2019; Periyannan, 2018; Thind et al., 2017). These include approaches based on exome capture for genome simplification in combination with mutant populations (MutRenSeq, Steuernagel et al., 2016; AgRenSeq, Arora et al., 2019), approaches

based on mutant populations in combination with chromosome flow sorting (MutChromSeq, Sánchez-Martín et al., 2016) and approaches based on chromosome sorting with long-linkage assemblies (TACCA, Thind et al., 2017). Thind et al. (2017) 's and Bettgenhaeuser & Krattinger's (2019) recent comparisons of all these gene cloning approaches highlights that, inevitably, there is no perfect one-size-fits-all method. Rather, the selected method would depend on what the end goal is, the germplasm resources developed and the budget that is available. What Thind et al. (2017)'s comparison does not consider is the relative ease of deployment into elite germplasm material following gene isolation. Two parallel avenues of research can be envisaged for the rapid cloning of yellow rust resistance. One would focus on major (largely NBS-LRRs) resistance genes, as exemplified by the recent cloning of *Yr7* and *Yr5/YrSp* (Marchal et al., 2018), and the other on minor (largely non-NBS-LRR) yellow rust resistance genes. For MutRenSeq and ChromSeq, it could be possible to use HIFs developed for YR resistance QTL of interest as the basis for a mutant population. Alternatively, AgRenSeq could make use of the genetic diversity present in the WAGTAIL association mapping and NIAB Elite MAGIC populations.

6.5 Concluding remarks

The phenotypic, genotypic and genomic characterisation of yellow rust resistance QTL identified in this thesis provide the resources with which to help develop new molecular tools for the development of resilient wheat varieties with resistance to genetically diverse *Pst* races - both in the leaves and ears. The identification of these QTL in germplasm of direct relevance to UK and north-western European wheat breeding, and the ability to understand how these QTL interact, will be of direct benefit of researchers and breeders aiming to deploy these resistance loci in live wheat breeding programmes. The next steps will be to use the HIF resources developed to fine-map each of the QTL with the ultimate aim of identifying the underlying genes and allelic variants. Additionally, future work should also focus on the characterisation of QTL interactions to identify effective combinations.

7 REFERENCES

Agenbag, G. M., Pretorius, Z. A., Boyd, L. A., Bender, C. M., & Prins, R. (2012). Identification of adult plant resistance to stripe rust in the wheat cultivar Cappelle-Desprez. *Theoretical and Applied Genetics*, *125*(1), 109–120.

Aguilar, V., Stamp, P., Winzeler, M., Winzeler, H., Schachermayr, G., Keller, B., Zanetti, S., & Messmer, M. M. (2005). Inheritance of field resistance to Stagonospora nodorum leaf and glume blotch and correlations with other morphological traits in hexaploid wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, *111*(2), 325–336.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Alexandratos, N., & Bruinsma, J. (2012). *World agriculture towards 2030/2050: the 2012 revision: ESA Working Paper No. 12-03*. Rome.

Ali, S., Gladieux, P., Leconte, M., Gautier, A., Justesen, A. F., Hovmøller, M. S., Enjalbert, J., & de Vallavieille-Pope, C. (2014). Origin, Migration Routes and Worldwide Population Genetic Structure of the Wheat Yellow Rust Pathogen Puccinia striiformis f.sp. tritici. *PLoS Pathogens*, *10*(1), e1003903.

Ali, S., Leconte, M., Rahman, H., Saqib, M. S., Gladieux, P., Enjalbert, J., & Vallavieille-Pope, C. de. (2014). A high virulence and pathotype diversity of Puccinia striiformis f.sp. tritici at its centre of diversity, the Himalayan region of Pakistan. *European Journal of Plant Pathology*, *140*(2), 275–290.

Ali, S., Rodriguez-Algaba, J., Thach, T., Sørensen, C. K., Hansen, J. G., Lassen, P., Nazari, K., Hodson, D. P., Justesen, A. F., & Hovmøller, M. S. (2017). Yellow Rust Epidemics Worldwide Were Caused by Pathogen Races from Divergent Genetic Lineages. *Frontiers in Plant Science*, *8*, 1057.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

Angus, W. J. (2001). United Kingdom Wheat Pool. In P. Bonjean & W. J. Angus (Eds.), *The World Wheat Book - a history of wheat breeding* (pp. 103–126). Paris: Lavoisier Publishing.

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., … Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*(6403).

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., … Wulff, B. B. H. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature Biotechnology*, *37*(2), 139–143.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, *25*(1), 25–29.

Baggs, E., Dagdas, G., & Krasileva, K. V. (2017). NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Current Opinion in Plant Biology*, *38*, 59–67.

Bailey, P. C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., Moscou, M., & Krasileva, K. V. (2018). Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *Genome Biology*.

Balakireva, A. V, & Zamyatnin, A. A. (2018). Indispensable Role of Proteases in Plant Innate Immunity. *International Journal of Molecular Sciences*, *19*(2), 629.

Bansept-Basler, P. G. M. (2013). *A genome-wide association study of resistance to the yellow rust pathogen (Puccinia striiformis f. sp. tritici) in elite UK wheat germplasm*. University of East Anglia, Norwich UK.

Barrero, J. M., Cavanagh, C., Verbyla, K. L., Tibbits, J. F. G., Verbyla, A. P., Huang, B. E., Rosewarne, G. M., Stephen, S., Wang, P., Whan, A., Rigault, P., Hayden, M. J., & Gubler, F. (2015). Transcriptomic analysis of wheat near-isogenic lines identifies PM19-A1 and A2 as candidates for a major dormancy QTL. *Genome Biology*, *16*(1), 1167.

Basnet, B. R., Singh, R. P., Ibrahim, A. M. H., Herrera-Foessel, S. A., Huerta-Espino, J., Lan, C., & Rudd, J. C. (2014). Characterization of Yr54 and other genes associated with adult plant resistance to yellow rust and leaf rust in common wheat Quaiu 3. *Molecular Breeding*, *33*(2), 385–399.

Bayles, R. A., Flath, K., Hovmøller, M. S., & Vallavieille-Pope, C. de. (2000). Breakdown of the Yr17 resistance to yellow rust of wheat in northern Europe. *Agronomie*, *20*(7), 805–811.

Beddow, J. M., Pardey, P. G., Chai, Y., Hurley, T. M., Kriticos, D. J., Braun, H.-J., Park, R. F., Cuddy, W. S., & Yonow, T. (2015). Research investment implications of shifts in the global geography of wheat stripe rust. *Nature Plants*, *1*(10), 15132.

Belozersky, M. A., Sarbakanova, S. T., & Dunaevsky, Y. E. (1989). Aspartic proteinase from wheat seeds: isolation, properties and action on gliadin. *Planta*, *177*(3), 321–326.

Bernoux, M., Ellis, J. G., & Dodds, P. N. (2011). New insights in plant immunity signaling activation. *Current Opinion in Plant Biology*, *14*(5), 512–518.

Bettgenhaeuser, J., & Krattinger, S. G. (2019). Rapid gene cloning in cereals. *Theoretical and Applied Genetics*, *132*(3), 699–711.

Bian, X., Tyrell, S., & Davey, R. P. (2017). The Grassroots life science data infrastructure. Retrieved from https://grassroots.tools

Biffen, R. H. (1905). Mendel's laws of inheritance and wheat breeding. *The Journal of Agricultural Science*, *1*(1), 4–48.

Borrill, P., Ramirez-Gonzalez, R., & Uauy, C. (2016). expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform. *Plant Physiology*, *170*(4), 2172–2186.

Boshoff, W. H. P., Pretorius, Z. A., & van Niekerk, B. D. (2002). Establishment, Distribution, and Pathogenicity of Puccinia striiformis f. sp. tritici in South Africa. *Plant Disease*, *86*(5), 485–492.

Bostwick, D. E., Ohm, H. W., & Shaner, G. (1993). Inheritance of Septoria Glume Blotch Resistance in Wheat. *Crop Science*, *33*, 439–443.

Boukhatem, N., Baret, P. V, Mingeot, D., & Jacquemin, J. M. (2002). Quantitative trait loci for resistance against Yellow rust in two wheat-derived recombinant inbred line populations. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *104*(1), 111–118.

Boyd, L. A. (2005). Can Robigus defeat an old enemy? – Yellow rust of wheat. *The Journal of Agricultural Science*, *143*(04), 233.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L. A., D'Amore, R., Allen, A. M., … Hall, N. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, *491*, 705. Retrieved from https://doi.org/10.1038/nature11650

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X., & Huard, F. (2010). Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Research*, *119*(1), 201–212.

Broman, K. W., Gatti, D. M., Simecek, P., Furlotte, N. A., Prins, P., Sen, Ś., Yandell, B. S., & Churchill, G. A. (2019). R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics*, *211*(2), 495–502.

Brown, J. K. M., & Hovmøller, M. S. (2002). Aerial Dispersal of Pathogens on the Global and Continental Scales and Its Impact on Plant Disease. *Science*, *297*(5581), 537 LP – 541.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., & Kleinhofs, A. (2002). The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America*.

Bryant, R. R. M., McGrann, G. R. D., Mitchell, A. R., Schoonbeek, H., Boyd, L. A., Uauy, C., Dorling, S., & Ridout, C. J. (2014). A change in temperature modulates defence to yellow (stripe) rust in wheat line UC1041 independently of resistance gene Yr36. *BMC Plant Biology*, *14*(1), 10.

Bulli, P., Zhang, J., Chao, S., Chen, X., & Pumphrey, M. (2016). Genetic Architecture of Resistance

to Stripe Rust in a Global Winter Wheat Germplasm Collection. *G3: Genes, Genomes, Genetics*, *6*(8), 2237–2253.

Calderini, D. F., & Slafer, G. A. (1998). Changes in yield and yield stability in wheat during the 20th century. *Field Crops Research*, *57*(3), 335–347.

Camargo, A. V., Mott, R., Gardner, K. A., Mackay, I. J., Corke, F., Doonan, J. H., Kim, J. T., & Bentley, A. R. (2016). Determining phenological patterns associated with the onset of senescence in a wheat magic mapping population. *Frontiers in Plant Science*, *7*(OCTOBER2016).

Cantu, D., Segovia, V., MacLean, D., Bayles, R., Chen, X., Kamoun, S., Dubcovsky, J., Saunders, Diane G. O., & Uauy, C. (2013). Genome analyses of the wheat yellow (stripe) rust pathogen Puccinia striiformis f. sp. tritici reveal polymorphic and haustorial expressed secreted proteins as candidate effectors. *BMC Genomics*, *14*(1), 270.

Cavanagh, C., Morell, M., Mackay, I., & Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology*, *11*(2), 215–221.

Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., & Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. *Frontiers in Plant Science*, *5*, 606.

Chakraborty, S., & Newton, A. C. (2011). Climate change, plant diseases and food security: an overview. *Plant Pathology*, *60*(1), 2–14.

Chapman, J. A., Mascher, M., Buluç, A., Barry, K., Georganas, E., Session, A., Strnadova, V., Jenkins, J., Sehgal, S., Oliker, L., Schmutz, J., Yelick, K. A., Scholz, U., Waugh, R., Poland, J. A., Muehlbauer, G. J., Stein, N., & Rokhsar, D. S. (2015). A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biology*, *16*(1), 26.

Chaves, M. S., Martinelli, J. A., Wesp-Guterres, C., Graichen, Felipe André Sganzerla, Brammer, S. P., Scagliusi, S. M., da Silva, Paulo Roberto, Wiethölter, P., Torres, Gisele Abigail Montan, Lau, E. Y., Consoli, L., & Chaves, Ana Lúcia Soares. (2013). The importance for food security of maintaining rust resistance in wheat. *Food Security*, *5*(2), 157–176.

Chen, W., Wellings, C., Chen, X., Kang, Z., & Liu, T. (2014). Wheat stripe (yellow) rust caused by Puccinia striiformis f. sp. tritici. *Molecular Plant Pathology*, *15*(5), 433–446.

Chen, X., & Kang, Z. (Eds.). (2017). *Stripe Rust*.

Chen, X. M. (2005). Epidemiology and control of stripe rust [Puccinia striiformis f. sp. tritici] on wheat. *Canadian Journal of Plant Pathology*, *27*, 314–337.

Chen, X. M. (2013). Review Article: High-Temperature Adult-Plant Resistance, Key for Sustainable Control of Stripe Rust. *American Journal of Plant Sciences*, *04*(03), 608–627.

Chen, X. M., Moore, M., Milus, E. A., Long, D. L., Line, R. F., Marshall, D., & Jackson, L. (2002).

Wheat Stripe Rust Epidemics and Races of Puccinia striiformis f. sp. tritici in the United States in 2000. *Plant Disease*, *86*(1), 39–46.

Chen, X., Penman, L., Wan, A., & Cheng, P. (2010). Virulence races of Puccinia striiformis f. sp. tritici in 2006 and 2007 and development of wheat stripe rust and distributions, dynamics, and evolutionary relationships of races from 2000 to 2007 in the United States. *Canadian Journal of Plant Pathology*, *32*(3), 315–333.

Chng, S. F., Cromey, M. G., & Shorter, S. C. (2011). Durability of resistance to stripe rust in the wheat cultivar 'Claire' in New Zealand. *New Zealand Plant Protection*, *64*, 17–24.

Churchill, G. A., Airey, D. C., Allayee, H., Angel, J. M., Attie, A. D., Beatty, J., ... Zou, F. (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics*, *36*(11), 1133–1137.

Clavijo, B. J., Venturini, L., Schudoma, C., Accinelli, G. G., Kaithakottil, G., Wright, J., ... Clark, M. D. (2017). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Research*, *27*(5), 885–896.

Cloutier, S., McCallum, B. D., Loutre, C., Banks, T. W., Wicker, T., Feuillet, C., Keller, B., & Jordan, M. C. (2007). Leaf rust resistance gene Lr1, isolated from bread wheat (Triticum aestivum L.) is a member of the large psr567 gene family. *Plant Molecular Biology*, *65*(1), 93–106.

Cockram, J., Jones, H., Leigh, F. J., O'Sullivan, D., Powell, W., Laurie, D. A., & Greenland, A. J. (2007). Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *Journal of Experimental Botany*, *58*(6), 1231–1244.

Cockram, J., & Mackay, I. (2018). Genetic Mapping Populations for Conducting High-Resolution Trait Mapping in Plants. In R. Varshney, M. Pandey, & A. Chitikineni (Eds.), *Plant Genetics and Molecular Biology* (pp. 109–138).

Cromey, M. G. (1989a). Infection and control of stripe rust in wheat spikes. *New Zealand Journal of Crop and Horticultural Science*.

Cromey, M. G. (1989b). Occurrence and effects of stripe rust in wheat spikes in New Zealand. *New Zealand Journal of Crop and Horticultural Science*, *17*(2), 155–158.

Crossa, J., Burgueño, J., Dreisigacker, S., Vargas, M., Herrera-Foessel, S. A., Lillemo, M., Singh, R. P., Trethowan, R., Warburton, M., Franco, J., Reynolds, M., Crouch, J. H., & Ortiz, R. (2007). Association Analysis of Historical Bread Wheat Germplasm Using Additive Genetic Covariance of Relatives and Population Structure. *Genetics*, *177*(3), 1889–1913.

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics*, *11*(4), 381–393.

Dangl, J. L., Horvath, D. M., & Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science*, *341*(6147), 746–751.

Day, L. (2013). Proteins from land plants – Potential resources for human nutrition and food security. *Trends in Food Science & Technology*, *32*(1), 25–42.

De Châtel, F. (2014). The Role of Drought and Climate Change in the Syrian Uprising: Untangling the Triggers of the Revolution. *Middle Eastern Studies*, *50*(4), 521–535.

de Vallavieille-Pope, C., Ali, S., Leconte, M., Enjalbert, J., Delos, M., & Rouzet, J. (2011). Virulence Dynamics and Regional Structuring of Puccinia striiformis f. sp. tritici in France Between 1984 and 2009. *Plant Disease*, *96*(1), 131–140.

de Vallavieille-Pope, C., Picard-Formery, H., Radulovic, S., & Johnson, R. (1990). Specific resistance factors to yellow rust in seedlings of some French wheat varieties and races of Puccinia striiformis Westend in France. *Agronomie*, *10*(2), 103–113.

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., Rudd, J. J., Dickman, M., Kahmann, R., Ellis, J., & Foster, G. D. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology*, *13*(4), 414–430.

Delhaize, E., Rathjen, T. M., & Cavanagh, C. R. (2015). The genetics of rhizosheath size in a multiparent mapping population of wheat. *Journal of Experimental Botany*.

Dolores Vazquez, M., James Peterson, C., Riera-Lizarazu, O., Chen, X., Heesacker, A., Ammar, K., Crossa, J., & Mundt, C. C. (2012). Genetic analysis of adult plant, quantitative resistance to stripe rust in wheat cultivar 'Stephens' in multi-environment trials. *Theoretical and Applied Genetics*, *124*(1), 1–11.

Downie, R. C., Bouvet, L., Furuki, E., Gosman, N., Gardner, K. A., Mackay, I. J., Campos Mantello, C., Mellers, G., Phan, H. T. T., Rose, G. A., Tan, K.-C., Oliver, R. P., & Cockram, J. (2018). Assessing European Wheat Sensitivities to Parastagonospora nodorum Necrotrophic Effectors and Fine-Mapping the Snn3-B1 Locus Conferring Sensitivity to the Effector SnTox3. *Frontiers in Plant Science*, *9*, 881.

Duan, X., Tellier, A., Wan, A., Leconte, M., de Vallavieille-Pope, C., & Enjalbert, J. (2010). Puccinia striiformis f.sp. tritici presents high diversity and recombination in the over-summering zone of Gansu, China. *Mycologia*, *102*(1), 44–53.

Duveiller, E., Singh, R., & Nicol, J. (2007). The challenges of maintaining wheat productivity: pests, diseases, and potential epidemics. *Euphytica*, *157*(3), 417–430.

Eitas, T. K., & Dangl, J. L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Current Opinion in Plant Biology*, *13*(4), 472–477.

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., & Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. *Frontiers in Plant Science*, *5*, 641.

Enjalbert, J., Duan, X., Leconte, M., Hovmøller, M. S., & Vallavieille-Pope, C. de. (2005). Genetic evidence of local adaptation of wheat yellow rust (Puccinia striiformis f. sp. tritici) within France. *Molecular Ecology*, *14*(7), 2065–2073.

Evenson, R. E., & Gollin, D. (2003). Assessing the Impact of the Green Revolution, 1960 to 2000. *Science*, *300*(5620), 758 LP – 762.

FAO. (2017). FAOSTAT statistical database. Retrieved February 23, 2017, from http://www.fao.org/faostat/en/#data/QC

Feng, J. Y., Wang, M. N., Chen, X. M., See, D. R., Zheng, Y. L., Chao, S. M., & Wan, A. M. (2015). Molecular Mapping of YrSP and Its Relationship with Other Genes for Stripe Rust Resistance in Wheat Chromosome 2BL. *Phytopathology*, *105*(9), 1206–1213.

Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A., & Keller, B. (2003). Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (Triticum aestivum L.) genome. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(25), 15253–15258.

Figueiredo, J., Sousa Silva, M., & Figueiredo, A. (2018). Subtilisin-like proteases in plant defence: the past, the present and beyond. *Molecular Plant Pathology*, *19*(4), 1017–1028.

Finckh, M. R., Gacek, E. S., Goyeau, H., Lannou, C., Merz, U., Mundt, C. C., Munk, L., Nadziak, J., Newton, A. C., Vallavieille-Pope, C. de, & Wolfe, M. S. (2000). Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie*, *20*(7), 813–837. Retrieved from https://doi.org/10.1051/agro:2000177

Finckh, M., & Wolfe, M. (2006). *Diversification strategies* (B. Cooke, D. Jones, & B. Kaye, Eds.).

Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., … Mitchell, A. L. (2017). InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Research*, *45*(D1), D190–D199.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., & Bateman, A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, *44*(D1), D279–D285.

Flor, H. H. (1956). *The Complementary Genic Systems in Flax and Flax Rust* (M. Demerec, Ed.). In (pp. 29–54).

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., Sela, H., Fahima, T., & Dubcovsky, J. (2009). A Kinase-START Gene Confers Temperature-Dependent Resistance to Wheat Stripe Rust. *Science*, *323*(5919), 1357–1360.

Gardner, K. A., Wittern, L. M., & Mackay, I. J. (2016). A highly recombined, high-density, eight-founder wheat MAGIC map reveals extensive segregation distortion and genomic locations of

introgression segments. *Plant Biotechnology Journal*, *14*(6), 1406–1417.

Garin, V., Wimmer, V., Borchardt, D., van Eeuwijk, F., & Malosetti, M. (2018). *mppR: Multi-Parent Population QTL Analysis*. Retrieved from https://cran.r-project.org/package=mppR

Garnica, D. P., Upadhyaya, N. M., Dodds, P. N., & Rathjen, J. P. (2013). Strategies for Wheat Stripe Rust Pathogenicity Identified by Transcriptome Sequencing. *PLoS ONE*, *8*(6), e67150.

Gervais, L., Dedryver, F., Morlais, J.-Y., Bodusseau, V., Negre, S., Bilous, M., Groos, C., & Trottet, M. (2003). Mapping of quantitative trait loci for field resistance to Fusarium head blight in an European winter wheat. *Theoretical and Applied Genetics*, *106*(6), 961–970.

Gilmour, A. R., Cullis, B. R., & Verbyla, A. P. (1997). Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. *Journal of Agricultural, Biological, and Environmental Statistics*, *2*(3), 269–293.

Gou, J.-Y., Li, K., Wu, K., Wang, X., Lin, H., Cantu, D., Uauy, C., Dobon-Alonso, A., Midorikawa, T., Inoue, K., Sánchez, J., Fu, D., Blechl, A., Wallington, E., Fahima, T., Meeta, M., Epstein, L., & Dubcovsky, J. (2015). Wheat Stripe Rust Resistance Protein WKS1 Reduces the Ability of the Thylakoid-Associated Ascorbate Peroxidase to Detoxify Reactive Oxygen Species. *The Plant Cell*, *27*(6), 1755–1770.

Guo, Q., Zhang, Z. J., Xu, Y. B., Li, G. H., Feng, J., & Zhou, Y. (2008). Quantitative trait loci for high-temperature adult-plant and slow-rusting resistance to Puccinia striiformis f. sp. tritici in wheat cultivars. *Phytopathology*, *98*(7), 803–809.

Hedrick, P. W. (1985). *Genetics of populations*. Boston, USA: Jones and Bartlett Publishers, Inc.

Herrera-Foessel, S. A., Singh, R. P., Lillemo, M., Huerta-Espino, J., Bhavani, S., Singh, S., Lan, C., Calvo-Salazar, V., & Lagudah, E. S. (2014). Lr67/Yr46 confers adult plant resistance to stem rust and powdery mildew in wheat. *Theoretical and Applied Genetics*, *127*(4), 781–789.

HGCA. (2011). *Protocol CER 12-16: HGCA Recommended List Cereal Trials Protocol*. Kenilworth, Warwickshire.

Hodges, M., Flesch, V., Gálvez, S., & Bismuth, E. (2003). Higher plant NADP+-dependent isocitrate dehydrogenases, ammonium assimilation and NADPH production. *Plant Physiology and Biochemistry*, *41*(6), 577–585.

Hou, S., Jamieson, P., & He, P. (2018). The cloak, dagger, and shield: proteases in plant–pathogen interactions. *Biochemical Journal*, *475*(15), 2491–2509.

Hovmøller, M. S., Walter, S., Bayles, R. A., Hubbard, A., Flath, K., Sommerfeldt, N., Leconte, M., Czembor, P., Rodriguez-Algaba, J., Thach, T., Hansen, J. G., Lassen, P., Justesen, A. F., Ali, S., & de Vallavieille-Pope, C. (2016). Replacement of the European wheat yellow rust population by new races from the centre of diversity in the near-Himalayan region. *Plant Pathology*, *65*(3), 402–411.

Hovmøller, M. S. (2007). Sources of seedling and adult plant resistance to Puccinia striiformis f.sp. tritici in European wheats. *Plant Breeding, 126*(3), 225–233.

Hovmøller, M. S., & Justesen, A. F. (2007a). Appearance of atypical Puccinia striiformis f. sp. tritici phenotypes in north-western Europe. *Australian Journal of Agricultural Research, 58*(6), 518–524. Retrieved from http://dx.doi.org/10.1071/AR06146

Hovmøller, M. S., & Justesen, A. F. (2007b). Rates of evolution of avirulence phenotypes and DNA markers in a northwest European population of Puccinia striiformis f. sp. tritici. *Molecular Ecology, 16*(21), 4637–4647.

Hovmøller, M. S., Justesen, A. F., & Brown, J. K. M. (2002). Clonality and long-distance migration of Puccinia striiformis f.sp. tritici in north-west Europe. *Plant Pathology, 51*(1), 24–32.

Hovmøller, M. S. (2001). Disease severity and pathotype dynamics of Puccinia striiformis f.sp. tritici in Denmark. *Plant Pathology, 50*(2), 181–189.

Hovmøller, M. S, Sørensen, C. K., Walter, S., & Justesen, A. F. (2011). Diversity of Puccinia striiformis on Cereals and Grasses. *Annual Review of Phytopathology, 49*(1), 197–217.

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

Huang, B. E., & George, A. W. (2011). R/mpMap: A computational platform for the genetic analysis of multi-parent recombinant inbred lines. *Bioinformatics*.

Huang, B. E., George, A. W., Forrest, K. L., Kilian, A., Hayden, M. J., Morell, M. K., & Cavanagh, C. R. (2012). A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnology Journal, 10*(7), 826–839.

Huang, B. E., Verbyla, K. L., Verbyla, A. P., Raghavan, C., Singh, V., Gaur, P., Leung, H., Varshney, R. K., & Cavanagh, C. R. (2015). MAGIC populations in crops: current status and future prospects. *Theoretical and Applied Genetics*, 1–19.

Huang, L., Brooks, S. A., Li, W., Fellers, J. P., Trick, H. N., & Gill, B. S. (2003). Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. *Genetics, 164*(2), 655–664. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12807786

Hubbard, A. J., & Bayles, R. A. (2013). *UK Cereal Pathogen Virulence Survey: Yellow rust of wheat – 2012 Annual Report*. NIAB.

Hubbard, A., Lewis, C., Yoshida, K., Ramirez-Gonzalez, R., de Vallavieille-Pope, C., Thomas, J., Kamoun, S., Bayles, R., Uauy, C., & Saunders, D. (2015). Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. *Genome Biology, 16*(1), 23.

Hubbard, A., Pritchard, L., Coventry, E., & Holdgate, S. (2015). *United Kingdom Cereal Pathogen Virulence Survey 2014 Annual Report*. NIAB.

Hubbard, A., Pritchard, L., & Holdgate, S. (2016). *United Kingdom Cereal Pathogen Virulence Survey 2016 Annual Report: Part 1: Wheat Yellow Rust, Wheat Powdery Mildew and Barley Powdery Mildew*. NIAB.

Hubbard, A., Wilderspin, S., & Holdgate, S. (2017). *United Kingdom Cereal Pathogen Virulence survey 2017 Annual Report*. NIAB.

Hulbert, S. H., Webb, C. A., Smith, S. M., & Sun, Q. (2001). Resistance Gene Complexes: Evolution and Utilization. *Annual Review of Phytopathology*, *39*(1), 285–312.

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., Wicker, T., Yahiaoui, N., Mago, R., & Keller, B. (2013). Rye Pm8 and wheat Pm3 are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *The Plant Journal*, *76*(6), 957–969.

International Wheat Genome Sequencing Consortium. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science, 345*(6194), 1251788.

Jagger, L. J., Newell, C., Berry, S. T., MacCormack, R., & Boyd, L. A. (2011). The genetic characterisation of stripe rust resistance in the German wheat cultivar Alcedo. *Theoretical and Applied Genetics, 122*(4), 723–733.

Jighly, A., Oyiga, B., Makdis, F., Nazari, K., Youssef, O., Tadesse, W., Abdalla, O., & Ogbonnaya, F. (2015). Genome-wide DArT and SNP scan for QTL associated with resistance to stripe rust (Puccinia striiformis f. sp. tritici) in elite ICARDA wheat (Triticum aestivum L.) germplasm. *Theoretical and Applied Genetics*, 1–19.

Jin, Y., Szabo, L. J., & Carson, M. (2010). Century-Old Mystery of Puccinia striiformis Life History Solved with the Identification of Berberis as an Alternate Host. *Phytopathology*, *100*(5), 432–435.

Johnson, R. (1992). Past, present and future opportunities in breeding for disease resistance, with examples from wheat. *Euphytica*, *63*, 3–22.

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323–329.

Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G. J., Maclean, D., Cock, P. J., Leggett, R. M., Bryan, G. J., Cardle, L., Hein, I., & Jones, J. D. G. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal*, *76*(3), 530–544.

Kertho, A., Mamidi, S., Bonman, J. M., McClean, P. E., & Acevedo, M. (2015). Genome-Wide Association Mapping for Resistance to Leaf and Stripe Rust in Winter-Habit Hexaploid Wheat Landraces. *PLoS ONE, 10*(6), e0129580.

Klymiuk, V., Yaniv, E., Huang, L., Raats, D., Fatiukha, A., Chen, S., … Fahima, T. (2018). Cloning of the wheat Yr15 resistance gene sheds light on the plant tandem kinase-pseudokinase family. *Nature Communications*, *9*(1), 3735.

Kourelis, J., & van der Hoorn, R. A. L. (2018). Defended to the Nines: 25 Years of Resistance Gene Cloning Identifies Nine Mechanisms for R Protein Function. *The Plant Cell*, *30*(2), 285 LP – 299.

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., Durrant, C., Mott, R., & Mauricio, R. (2009). A Multiparent Advanced Generation Inter-Cross to Fine-Map Quantitative Traits in Arabidopsis thaliana. *PLoS Genetics*, *5*(7), e1000551.

Krasileva, K. V, Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R. H., Wang, X., Borrill, P., Fosker, C., Ayling, S., Phillips, A. L., Uauy, C., & Dubcovsky, J. (2017). Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(6), E913–E921.

Krattinger, S. G., & Keller, B. (2016). Molecular genetics and evolution of disease resistance in cereals. *New Phytologist*, *212*(2), 320–332.

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., Bossolini, E., Selter, L. L., & Keller, B. (2009). A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. *Science*, *323*(5919), 1360–1363.

Krattinger, Simon G., Kang, J., Bräunlich, S., Boni, R., Chauhan, H., Selter, L. L., Robinson, M. D., Schmid, M. W., Wiederhold, E., Hensel, G., Kumlehn, J., Sucher, J., Martinoia, E., & Keller, B. (2019). Abscisic acid is a substrate of the ABC transporter encoded by the durable wheat disease resistance gene Lr34. *New Phytologist*.

Kroj, T., Chanclud, E., Michel-Romiti, C., Grand, X., & Morel, J.-B. (2016). Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *New Phytologist*, *210*(2), 618–626.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., & Marra, M. A. (2009). Circos: an information aesthetic for comparative genomics. *Genome Research*, *19*(9), 1639–1645.

Kumar, B. N. A., Azam-Ali, S. N., Snape, J. W., Weightman, R. M., & Foulkes, M. J. (2011). Relationships between carbon isotope discrimination and grain yield in winter wheat under well-watered and drought conditions. *The Journal of Agricultural Science, 149*(3), 257–272.

Lagudah, E. S. (2011). Molecular genetics of race non-specific rust resistance in wheat. *Euphytica*, *179*(1), 81–91.

Leterrier, M., Barroso, J. B., Valderrama, R., Palma, J. M., & Corpas, F. J. (2012). NADP-dependent isocitrate dehydrogenase from Arabidopsis roots contributes in the mechanism of defence

against the nitro-oxidative stress induced by salinity. *TheScientificWorldJournal*, *2012*, 694740.

Leterrier, M., Leterrier, M., del Río, L. A., & Corpas, F. J. (2007). Cytosolic NADP-isocitrate dehydrogenase of pea plants: Genomic clone characterization and functional analysis under abiotic stress conditions. *Free Radical Research*, *41*(2), 191–199.

Lewis, C. M. (2006). *The genetic basis of Puccinia striiformis resistance in UK wheat germplasm*. University of East Anglia, Norwich UK.

Lillemo, M., Asalf, B., Singh, R. P., Huerta-Espino, J., Chen, X. M., He, Z. H., & Bjørnstad, Å. (2008). The adult plant rust resistance loci Lr34/Yr18 and Lr46/Yr29 are important determinants of partial resistance to powdery mildew in bread wheat line Saar. *Theoretical and Applied Genetics*, *116*(8), 1155–1166.

Lillemo, M., Joshi, A. K., Prasad, R., Chand, R., & Singh, R. P. (2013). QTL for spot blotch resistance in bread wheat line Saar co-locate to the biotrophic disease resistance loci Lr34 and Lr46. *Theoretical and Applied Genetics*, *126*(3), 711–719.

Lin, F., & Chen, X. M. (2007). Genetics and molecular mapping of genes for race-specific all-stage resistance and non-race-specific high-temperature adult-plant resistance to stripe rust in spring wheat cultivar Alpowa. *Theoretical and Applied Genetics*, *114*(7), 1277–1287.

Lin, M, Corsi, B., Ficke, A., Tan, K.-C., Cockram, J., & Lillemo, M. (2019). *Genetic mapping using a wheat multi-founder population reveals a locus on chromosome 2A controlling resistance to both leaf and glume blotch caused by the necrotrophic fungal pathogen Parastagonospora nodorum*. Manuscript submitted for publication.

Lin, Min, Corsi, B., Ficke, A., Tan, K. C., Cockram, J., & Lillemo, M. (2020). Genetic mapping using a wheat multi-founder population reveals a locus on chromosome 2A controlling resistance to both leaf and glume blotch caused by the necrotrophic fungal pathogen Parastagonospora nodorum. *Theoretical and Applied Genetics*.

Line, R. F. (2002). Stripe rust of wheat and barley in North America: A restrospective historical review. *Annual Review of Phytopathology*, *40*(1), 75–118.

Liu, D., Zhang, L., Hao, M., Ning, S., Yuan, Z., Dai, S., Huang, L., Wu, B., Yan, Z., Lan, X., & Zheng, Y. (2018). Wheat breeding in the hometown of Chinese Spring. *The Crop Journal*, *6*(1), 82–90.

Liu, W., Frick, M., Huel, R., Nykiforuk, C. L., Wang, X., Gaudet, D. A., Eudes, F., Conner, R., Kuzyk, A., Chen, Q., Kang, Z., & Laroche, A. (2014). The Stripe Rust Resistance Gene Yr10 Encodes an Evolutionary-Conserved and Unique CC–NBS–LRR Sequence in Wheat. *Molecular Plant*, *7*(12), 1740–1755.

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate Trends and Global Crop Production Since 1980. *Science*, *333*(6042), 616 LP – 620.

Losert, D., Maurer, H. P., Leiser, W. L., & Würschum, T. (2017). Defeating the Warrior: genetic architecture of triticale resistance against a novel aggressive yellow rust race. *Theoretical and Applied Genetics*, *130*(4), 685–696.

Loutre, C., Wicker, T., Travella, S., Galli, P., Scofield, S., Fahima, T., Feuillet, C., & Keller, B. (2009). Two different CC-NBS-LRR genes are required for Lr10-mediated leaf rust resistance in tetraploid and hexaploid wheat. *The Plant Journal*, *60*(6), 1043–1054.

Lowe, I., Cantu, D., & Dubcovsky, J. (2011). Durable resistance to the wheat rusts: integrating systems biology and traditional phenotype-based research methods to guide the deployment of resistance genes. *Euphytica*, *179*(1), 69–79.

Loyce, C., Meynard, J. M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., … Doussinault, G. (2008). Interaction between cultivar and crop management effects on winter wheat diseases, lodging, and yield. *Crop Protection*, *27*(7), 1131–1142.

Luo, P., Hu, X., Zhang, H., & Ren, Z. (2009). Genes for resistance to stripe rust on chromosome 2B and their application in wheat breeding. *Progress in Natural Science*, *19*(1), 9–15.

Lupton, F. G. H., & Macer, R. C. F. (1962). Inheritance of resistance to yellow rust (Puccinia glumarum Erikss. & Henn.) in seven varieties of wheat. *Transactions of the British Mycological Society*, *45*(1), 21–45.

Maccaferri, M., Zhang, J., Bulli, P., Abate, Z., Chao, S., Cantu, D., Bossolini, E., Chen, X., Pumphrey, M., & Dubcovsky, J. (2015). A Genome-Wide Association Study of Resistance to Stripe Rust (Puccinia striiformis f. sp. tritici) in a Worldwide Collection of Hexaploid Spring Wheat (Triticum aestivum L.). *G3: Genes|Genomes|Genetics*, *5*(3), 449–465.

Mackay, I. J., Bansept-Basler, P., Barber, T., Bentley, A. R., Cockram, J., Gosman, N., Greenland, A. J., Horsnell, R., Howells, R., O'Sullivan, D. M., Rose, G. A., & Howell, P. J. (2014). An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation. *G3 (Bethesda, Md.)*, *4*(9), 1603–1610.

Mackay, I., & Powell, W. (2007). Methods for linkage disequilibrium mapping in crops. *Trends in Plant Science*, *12*(2), 57–63.

Mackay, T. F. (2001). The genetic architecture of quantitative traits. *Annual Review of Genetics*, *35*, 303–339.

Mago, R., Zhang, P., Vautrin, S., Šimková, H., Bansal, U., Luo, M.-C., … Dodds, P. N. (2015). The wheat Sr50 gene reveals rich diversity at a cereal disease resistance locus. *Nature Plants*, *1*(12), 15186.

Mains, E. B. (1933). Studies concerning Heteroecious Rusts. *Mycologia*, *25*(5), 407–417.

Mallard, S., Gaudet, D., Aldeia, A., Abelard, C., Besnard, A. L., Sourdille, P., & Dedryver, F. (2005). Genetic analysis of durable resistance to yellow rust in bread wheat. *Theoretical and Applied*

Genetics, *110*(8), 1401–1409.

Marchal, C., Zhang, J., Zhang, P., Fenwick, P., Steuernagel, B., Adamski, N. M., Boyd, L., McIntosh, R., Wulff, B. B. H., Berry, S., Lagudah, E., & Uauy, C. (2018). BED-domain-containing immune receptors confer diverse resistance spectra to yellow rust. *Nature Plants*, *4*(9), 662–668.

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K. S., … Imkova, H. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science*, *345*(6194), 1250092–1250092.

Markell, S. G., & Milus, E. A. (2008). Emergence of a Novel Population of Puccinia striiformis f. sp. tritici in Eastern United States. *Phytopathology*, *98*(6), 632–639.

Maydup, M., Antonietta, M., Guiamet, J., Graciano, C., López, J. R., & Tambussi, E. A. (2010). The contribution of ear photosynthesis to grain filling in bread wheat (Triticum aestivum L.). *Field Crops Research*, *119*, 48–58.

Mboup, M., Bahri, B., Leconte, M., de Vallavieille-Pope, C., Kaltz, O., & Enjalbert, J. (2012). Genetic structure and local adaptation of European wheat yellow rust populations: the role of temperature-specific adaptation. *Evolutionary Applications*, *5*(4), 341–352.

Mboup, M., Leconte, M., Gautier, A., Wan, A. M., Chen, W., de Vallavieille-Pope, C., & Enjalbert, J. (2009). Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. *Fungal Genetics and Biology*, *46*(4), 299–307.

McDonald, B. A., & Linde, C. (2002). Pathogen Population Genetics, Evolutionary Potential, and Durable Resistance. *Annual Review of Phytopathology*, *40*(1), 349–379.

McIntosh, R. A. (2009). History and status of the wheat rusts. In R. A. McIntosh (Ed.), *Proceedings of the Borlaug Global Rust Initiative Technical Workshop* (pp. 11–24). Ciudad Obregon.

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., & Xia, X. C. (2017). *Catalogue of Gene Symbols for Wheat: 2017 Supplement*.

McIntosh, R. A., Wellings, C. R., & Park, R. F. (1995). *Wheat rusts: An atlas of resistance genes*. East Melbourne, Australia and Dordrecht and Boston: CSIRO and {Kluwer Academic Publishers}.

McIntosh, R. A., Yamazaki, Y., Devos, K. M., Dubcovsky, J., Rogers, W. J., & Appels, R. (2003). Catalogue of Gene Symbols for Wheat. *International Wheat Genetics Symposium*, (4). Rome.

McIntosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, J., Morris, C., & Somers, D. J. (2008). Catalogue of gene symbols for wheat. *Proceedings of the 11th International Wheat Genetics Symposium*, (September), 24–29. San Diego, CA.

Melichar, J. P. E., Berry, S., Newell, C., MacCormack, R., & Boyd, L. A. (2008). QTL identification and microphenotype characterisation of the developmentally regulated yellow rust resistance in the UK wheat cultivar Guardian. *Theoretical and Applied Genetics*, *117*(3), 391–399.

Mendieta, J. R., Pagano, M. R., Muñoz, F. F., Daleo, G. R., & Guevara, M. G. (2006). Antimicrobial activity of potato aspartic proteases (StAPs) involves membrane permeabilization. *Microbiology, 152*(7), 2039–2047.

Mhamdi, A., Mauve, C., Gouia, H., Saindrenan, P., Hodges, M., & Noctor, G. (2010). Cytosolic NADP-dependent isocitrate dehydrogenase contributes to redox homeostasis and the regulation of pathogen responses in Arabidopsis leaves. *Plant, Cell & Environment, 33*(7), 1112–1123.

Michelmore, R. W., & Meyers, B. C. (1998). Clusters of Resistance Genes in Plants Evolve by Divergent Selection and a Birth-and-Death Process. *Genome Research , 8*(11), 1113–1130.

Milus, E. A., Kristensen, K., & Hovmøller, M. S. (2009). Evidence for increased aggressiveness in a recent widespread strain of Puccinia striiformis f. sp. tritici causing stripe rust of wheat. *Phytopathology, 99*(1), 89–94.

Mondal, S., Rutkoski, J. E., Velu, G., Singh, P. K., Crespo-Herrera, L. A., Guzmán, C., Bhavani, S., Lan, C., He, X., & Singh, R. P. (2016). Harnessing Diversity in Wheat to Enhance Grain Yield, Climate Resilience, Disease and Insect Pest Resistance and Nutrition Through Conventional and Modern Breeding Approaches. *Frontiers in Plant Science, 7*, 991.

Monteiro, F., & Nishimura, M. T. (2018). Structural, Functional, and Genomic Diversity of Plant NLR Proteins: An Evolved Resource for Rational Engineering of Plant Immunity. *Annual Review of Phytopathology, 56*(1), 243–267.

Moore, J. W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., Lillemo, M., Viccars, L., Milne, R., Periyannan, S., Kong, X., Spielmeyer, W., Talbot, M., Bariana, H., Patrick, J. W., Dodds, P., Singh, R., & Lagudah, E. (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nature Genetics, 47*, 1494–1498.

Mundt, C. C. (2002). Use of Multiline Cultivars and Cultivar Mixtures for Disease Management. *Annual Review of Phytopathology, 40*(1), 381–410.

Mundt, C. C. (2014). Durable resistance: a key to sustainable management of pathogens and pests. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 27*, 446–455.

Muñoz, F. F., Mendieta, J. R., Pagano, M. R., Paggi, R. A., Daleo, G. R., & Guevara, M. G. (2010). The swaposin-like domain of potato aspartic protease (StAsp-PSI) exerts antimicrobial activity on plant and human pathogens. *Peptides, 31*(5), 777–785.

Murray, G., & Brennan, J. (2009). Estimating disease losses to the Australian wheat industry. *Australasian Plant Pathology, 38*(6), 558–570.

Nelleman, C., MacDevette, C., Manders, M., Eickhout, B., Svihus, B., Prins, B., & Kaltenborn, B.

(Eds.). (2009). *The Environmental Food Crisis: The Environment's Role in Averting Future Food Crises* (GRID-Arend). United Nations Environment Programme.

Nelson, L. R., & Gates, C. E. (1982). Genetics of Host Plant Resistance of Wheat to Septoria nodoruml. *Crop Science, 22*, 771–773.

Noël, L., Moores, T. L., van der Biezen, E. A., Parniske, M., Daniels, M. J., Parker, J. E., & Jones, J. D. G. (1999). Pronounced Intraspecific Haplotype Divergence at the RPP5 Complex Disease Resistance Locus of Arabidopsis. *The Plant Cell, 11*(11), 2099 LP – 2111.

Oerke, E.-C., Dehne, H., Schonbeck, F., & Weber, A. (1994). *Crop Production and Crop Protection*. Amsterdam: Elsevier.

Oerke, E. C. (2006). Crop losses to pests. *The Journal of Agricultural Science, 144*(01), 31.

Oliver, R. P. (2014). A reassessment of the risk of rust fungi developing resistance to fungicides. *Pest Management Science, 70*(11), 1641–1645.

Parry, M. A. J., Reynolds, M., Salvucci, M. E., Raines, C., Andralojc, P. J., Zhu, X.-G., Price, G. D., Condon, A. G., & Furbank, R. T. (2011). Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of Experimental Botany, 62*(2), 453–467.

Pathan, A. K., Wellings, C. R., Bariana, H. S., & Park, R. F. (2008). Evaluation of seedling and adult plant resistance in European wheat cultivars to Australian isolates of Puccinia striiformis f. sp. tritici. *Euphytica*.

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., … Feuillet, C. (2008). A Physical Map of the 1-Gigabase Bread Wheat Chromosome 3B. *Science, 322*(5898), 101 LP – 104.

Pearce, S., Vazquez-Gross, H., Herin, S. Y., Hane, D., Wang, Y., Gu, Y. Q., & Dubcovsky, J. (2015). WheatExp: an RNA-seq expression database for polyploid wheat. *BMC Plant Biology, 15*(1), 299.

Peng, J., Richards, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., Beales, J., Fish, L. J., Worland, A. J., Pelica, F., Sudhakar, D., Christou, P., Snape, J. W., Gale, M. D., & Harberd, N. P. (1999). 'Green revolution' genes encode mutant gibberellin response modulators. *Nature, 400*(6741), 256–261.

Periyannan, S. (2018). Sustaining global agriculture through rapid detection and deployment of genetic resistance to deadly crop diseases. *The New Phytologist, 219*(1), 45–51.

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., Deal, K., Luo, M., Kong, X., Bariana, H., Mago, R., McIntosh, R., Dodds, P., Dvorak, J., & Lagudah, E. (2013). The Gene Sr33, an Ortholog of Barley Mla Genes, Encodes Resistance to Wheat Stem Rust Race Ug99. *Science, 341*(6147), 786–788.

Peterson, R. F., Campbell, A. B., & Hannah, A. E. (1948). A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Canadian Journal of Research, 26c*(5), 496–500.

Piepho, H.-P., & Möhring, J. (2007). Computing Heritability and Selection Response From Unbalanced Plant Breeding Trials. *Genetics, 177*(3), 1881–1888.

Pingali, P. (2012). Green Revolution: Impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences, 109*(31), 12302–12308.

Pingali, P., & Rosegrant, M. (1994). *Confronting the environmental consequences of the Green Revolution in Asia.*

Powell, N. M., Lewis, C. M., Berry, S. T., MacCormack, R., & Boyd, L. A. (2013). Stripe rust resistance genes in the UK winter wheat cultivar Claire. *Theoretical and Applied Genetics, 126*(6), 1599–1612.

Prasad, B. D., Creissen, G., Lamb, C., & Chattoo, B. B. (2009). Overexpression of Rice (Oryza sativa L.) OsCDR1 Leads to Constitutive Activation of Defense Responses in Rice and Arabidopsis. *Molecular Plant-Microbe Interactions, 22*(12), 1635–1644.

Pretty, J. (2008). Agricultural sustainability: concepts, principles and evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 363*(1491), 447–465.

Priestley, R. H., Bayles, R. A., & Crofts, J. (1982). *Yellow rust of wheat. UK cereal pathogen virulence survey 1981 annual report* (pp. 18–28). pp. 18–28. National Institute of Agricultural Botany.

Purdy, L. H., & Allan, R. E. (1965). Stripe rust head infection in five Pacific northwest wheats. *Plant Disease Reporter, 49*, 335–338.

R Core Team. (2017). *R: A language and environment for statistical computing.* Retrieved from www.R-project.org

R Core Team. (2019). *R: A Language and Environment for Statistical Computing.* Retrieved from https://www.r-project.org/

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., … Uauy, C. (2018). The transcriptional landscape of polyploid wheat. *Science, 361*(6403), eaar6089.

Rani, M., Mahato, A. K., Sinha, S. K., Dalal, M., Singh, N. K., & Mandal, P. K. (2019). Homeologue Specific Gene Expression Analysis of Two Vital Carbon Metabolizing Enzymes—Citrate Synthase and NADP-Isocitrate Dehydrogenase—from Wheat (Triticum aestivum L.) Under Nitrogen Stress. *Applied Biochemistry and Biotechnology, 188*(3), 569–584.

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., & Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications, 3*(1), 1293.

Rebetzke, G. J., Verbyla, A. P., Verbyla, K. L., Morell, M. K., & Cavanagh, C. R. (2014). Use of a large multiparent wheat mapping population in genomic dissection of coleoptile and seedling growth. *Plant Biotechnology Journal, 12*(2), 219–230.

Roelfs, A. P., Singh, R. P., & Saari, E. E. (1992). *Rust diseases of wheat: Concepts and methods of*

disease management. Mexico, D.F.: CIMMYT.

Rosewarne, G. M., Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., Lan, C. X., & He, Z. H. (2013). Quantitative trait loci of stripe rust resistance in wheat. *Theoretical and Applied Genetics*, *126*(10), 2427–2449.

Rosewarne, G. M., Singh, R. P., Huerta-Espino, J., & Rebetzke, G. J. (2008). Quantitative trait loci for slow-rusting resistance in wheat to leaf rust and stripe rust identified with multi-environment analysis. *Theoretical and Applied Genetics*, *116*(7), 1027–1034.

Rosewarne, G. M., Singh, R. P., Huerta-Espino, J., William, H. M., Bouchet, S., Cloutier, S., McFadden, H., & Lagudah, E. S. (2006). Leaf tip necrosis, molecular markers and β1-proteasome subunits associated with the slow rusting resistance genes Lr46/Yr29. *Theoretical and Applied Genetics*, *112*(3), 500–508.

Saintenac, C., Zhang, W., Salcedo, A., Rouse, M. N., Trick, H. N., Akhunov, E., & Dubcovsky, J. (2013). Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science*, *341*(6147), 783–786.

Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., Vrána, J., Kubaláková, M., Krattinger, S. G., Wicker, T., Doležel, J., Keller, B., & Wulff, Brande B H. (2016). Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biology*, *17*(1), 221.

Sannemann, W., Lisker, A., Maurer, A., Léon, J., Kazman, E., Cöster, H., Holzapfel, J., Kempf, H., Korzun, V., Ebmeyer, E., & Pillen, K. (2018). Adaptive selection of founder segments and epistatic control of plant height in the MAGIC winter wheat population WM-800. *BMC Genomics*, *19*(1), 559.

Sarris, P. F., Cevik, V., Dagdas, G., Jones, J. D. G., & Krasileva, K. V. (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biology*, *14*(1), 8.

Schnurbusch, T., Paillard, S., Fossati, D., Messmer, M., Schachermayr, G., Winzeler, M., & Keller, B. (2003). Detection of QTLs for Stagonospora glume blotch resistance in Swiss winter wheat. *Theoretical and Applied Genetics*, *107*(7), 1226–1234.

Schürmann, P., & Jacquot, J.-P. (2000). Plant Thioredoxin Systems Revisited. *Annual Review of Plant Physiology and Plant Molecular Biology*, *51*(1), 371–400.

Scofield, S. R., Huang, L., Brandt, A. S., & Gill, B. S. (2005). Development of a Virus-Induced Gene-Silencing System for Hexaploid Wheat and Its Use in Functional Analysis of the Lr21-Mediated Leaf Rust Resistance Pathway. *Plant Physiology*, *138*(4), 2165–2173.

Scott, P. R., Benedikz, P. W., & Cox, C. J. (1982). A genetic study of the relationship between height, time of ear emergence and resistance to Septoria nodorum in wheat. *Plant Pathology*, *31*(1),

45–60.

Sears, E. R. (1954). The aneuploids of common wheat. *Missouri Agriculture Experimental Station Research Bulletin*, (572), 1–59.

Seeholzer, S., Tsuchimatsu, T., Jordan, T., Bieri, S., Pajonk, S., Yang, W., Jahoor, A., Shimizu, K. K., Keller, B., & Schulze-Lefert, P. (2010). Diversity at the Mla Powdery Mildew Resistance Locus from Cultivated Barley Reveals Sites of Positive Selection. *Molecular Plant-Microbe Interactions*, *23*(4), 497–509.

Seo, E., Kim, S., Yeom, S.-I., & Choi, D. (2016). Genome-Wide Comparative Analyses Reveal the Dynamic Evolution of Nucleotide-Binding Leucine-Rich Repeat Gene Family among Solanaceae Plants. *Frontiers in Plant Science*, *7*, 1205.

Shah, R., Huang, B. E., Whan, A., Newberry, M., Verbyla, K., Morell, M. K., & Cavanagh, C. R. (2019). The complex genetic architecture of recombination and structural variation in wheat uncovered using a large 8-founder MAGIC population. *BioRxiv*.

Shankar, M., Walker, E., Golzar, H., Loughman, R., Wilson, R. E., & Francki, M. G. (2008). Quantitative Trait Loci for Seedling and Adult Plant Resistance to Stagonospora nodorum in Wheat. *Phytopathology*, *98*(8), 886–893.

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany*, *60*(6), 1537–1553.

Shewry, P. R., & Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food and Energy Security*, *4*(3), 178–202.

Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., & Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, *5*(3), 291–317.

Simões, I., & Faro, C. (2004). Structure and function of plant aspartic proteinases. *European Journal of Biochemistry*, *271*(11), 2067–2075.

Singh, A., Knox, R. E., DePauw, R. M., Singh, A. K., Cuthbert, R. D., Campbell, H. L., Shorter, S., & Bhavani, S. (2014). Stripe rust and leaf rust resistance QTL mapping, epistatic interactions, and co-localization with stem rust resistance loci in spring wheat evaluated over three continents. *Theoretical and Applied Genetics*, *127*(11), 2465–2477.

Singh, D., Park, R. F., McIntosh, R. A., & Bariana, H. S. (2008). Characterisation of stem rust and stripe rust seedling resistance genes in selected wheat cultivars from the United Kingdom. *Journal of Plant Pathology*.

Singh, R. P. (1992). Genetic Association of Leaf Rust Resistance Gene Lr34 with Adult Plant Resistance to Stripe Rust in Bread Wheat . *Phytopathology*, *82*(8), 835.

Singh, R. P., Herrera-Foessel, S., Huerta-Espino, J., Singh, S., Bhavani, S., Lan, C., & Basnet, B. R. (2014). Progress Towards Genetics and Breeding for Minor Genes Based Resistance to Ug99

and Other Rusts in CIMMYT High-Yielding Spring Wheat. *Journal of Integrative Agriculture*, *13*(2), 255–261.

Singh, R. P., Huerta-Espino, J., & Rajaram, S. (2000). Achieving near-immunity to leaf and stripe rusts in wheat by combining slow rusting resistance genes. *Acta Phytopathologica et Entomologica Hungarica*, *35*(1/4), 133–139.

Singh, R. P., Huerta-Espino, J., & William, H. M. (2005). Genetics and Breeding for Durable Resistance to Leaf and Stripe Rusts in Wheat. *Turkish Journal of Agriculture and Forestry*, *29*, 121–127.

Singh, R. P., Singh, P. K., Rutkoski, J., Hodson, D. P., He, X., Jørgensen, L., Hovmøller, M. S., & Huerta-Espino, J. (2016). Disease Impact on Wheat Yield Potential and Prospects of Genetic Control. *Annual Review of Phytopathology*, *54*.

Singh, R., William, H., Huerta-Espino, J., & Rosewarne, G. (2004). Wheat Rust in Asia: Meeting the Challenges with Old and New Technologies. *Proceedings of the 4th International Crop Science Congress*. Brisbane, Australia.

Singh, S., Hernandez, M. V, Crossa, J., Singh, P. K., Bains, N. S., Singh, K., & Sharma, I. (2012). Multi-trait and multi-environment QTL analyses for resistance to wheat diseases. *PLoS ONE*, *7*(6), e38008.

Solh, M., Nazari, K., Tadesse, W., & Wellings, C. R. (2012). The growing threat of stripe rust worldwide. *Borlaug Global Rust Initiative (BGRI) Conference*. Beijing, China.

Sørensen, C. K., Hovmøller, M. S., Leconte, M., Dedryver, F., & Vallavieille-Pope, C. de. (2014). New Races of Puccinia striiformis Found in Europe Reveal Race Specificity of Long-Term Effective Adult Plant Resistance in Wheat. *Phytopathology*, *104*(10), 1042–1051.

Spielmeyer, W., McIntosh, R. A., Kolmer, J., & Lagudah, E. S. (2005). Powdery mildew resistance and Lr34/Yr18 genes for durable resistance to leaf and stripe rust cosegregate at a locus on the short arm of chromosome 7D of wheat. *Theoretical and Applied Genetics*, *111*(4), 731–735.

Stadlmeier, M., Hartl, L., & Mohler, V. (2018). Usefulness of a Multiparent Advanced Generation Intercross Population With a Greatly Reduced Mating Design for Genetic Studies in Winter Wheat. *Frontiers in Plant Science*, *9*, 1825.

Stadlmeier, M., Jørgensen, L. N., Corsi, B., Cockram, J., Hartl, L., & Mohler, V. (2019). Genetic Dissection of Resistance to the Three Fungal Plant Pathogens Blumeria graminis, Zymoseptoria tritici, and Pyrenophora tritici-repentis Using a Multiparental Winter Wheat Population. *G3: Genes, Genomes, Genetics*, *9*(5), 1745–1757.

Steele, K. A., Humphreys, E., Wellings, C. R., & Dickinson, M. J. (2001). Support for a stepwise mutation model for pathogen evolution in Australasian Puccinia striiformis f.sp. tritici by use of molecular markers. *Plant Pathology*, *50*(2), 174–180.

Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E., & Keller, B. (2000). Subgenome chromosome walking in wheat: a 450-kb physical contig in Triticum monococcum L. spans the Lr10 resistance locus in hexaploid wheat (Triticum aestivum L.). *Proceedings of the National Academy of Sciences of the United States of America*, *97*(24), 13436–13441.

Steuernagel, B., Jupe, F., Witek, K., Jones, J. D. G., & Wulff, B. B. H. (2015). NLR-parser: rapid annotation of plant NLR complements. *Bioinformatics (Oxford, England)*, *31*(10), 1665–1667.

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

Steuernagel, B., Witek, K., Krattinger, S. G., Ramirez-Gonzalez, R. H., Schoonbeek, H., Yu, G., Baggs, E., Witek, A. I., Yadav, I., Krasileva, K. V, Jones, J. D. G., Uauy, C., Keller, B., Ridout, C. J., & Wulff, B. B. H. (2018). Physical and transcriptional organisation of the bread wheat intracellular immune receptor repertoire. *BioRxiv*, 339424.

Sthapit, J., Gbur, E., Brown-Guedira, G., Marshall, D., & Milus, E. (2012). Characterization of Resistance to Stripe Rust in Contemporary Cultivars and Lines of Winter Wheat from the Eastern United States. *Plant Disease, 96*, 737–745.

Storey, J. D. (2015). *qvalue: Q-value estimation for false discovery rate control*. Retrieved from http://github.com/jdstorey/qvalue

Strange, R. N., & Scott, P. R. (2005). Plant disease: a threat to global food security. *Annual Review of Phytopathology, 43*, 83–116.

Summers, R. W., & Brown, J. K. M. (2013). Constraints on breeding for disease resistance in commercially competitive wheat cultivars. *Plant Pathology, 62*, 115–121.

Szabo, L. J., & Bushnell, W. R. (2001). Hidden robbers: The role of fungal haustoria in parasitism of plants. *Proceedings of the National Academy of Sciences, 98*(14), 7654–7655.

Tambussi, E. A., Bort, J., Guiamet, J. J., Nogués, S., & Araus, J. L. (2007). The Photosynthetic Role of Ears in C3 Cereals: Metabolism, Water Use Efficiency and Contribution to Grain Yield. *Critical Reviews in Plant Sciences, 26*(1), 1–16.

Tamura, T., Terauchi, K., Kiyosaki, T., Asakura, T., Funaki, J., Matsumoto, I., Misaka, T., & Abe, K. (2007). Differential expression of wheat aspartic proteinases, WAP1 and WAP2, in germinating and maturing seeds. *Journal of Plant Physiology, 164*(4), 470–477.

Tan, K.-C., Oliver, R., Solomon, P., & Moffat, C. (2010). Proteinaceous necrotrophic effectors in fungal virulence. *Functional Plant Biology, 37*, 907–912.

Thind, A. K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., Vrána, J., Doležel, J., & Krattinger, S. G. (2017). Rapid cloning of genes in hexaploid wheat using cultivar-specific

long-range chromosome assembly. *Nature Biotechnology*, *35*, 793.

Thomas, E. L., & van der Hoorn, R. A. L. (2018). Ten Prominent Host Proteases in Plant-Pathogen Interactions. *International Journal of Molecular Sciences*, *19*(2), 639.

Tuinstra, M. R., Ejeta, G., & Goldsbrough, P. B. (1997). Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *TAG Theoretical and Applied Genetics*, *95*(5–6), 1005–1011.

Uauy, C., Brevis, J. C., Chen, X., Khan, I., Jackson, L., Chicaiza, O., Distelfeld, A., Fahima, T., & Dubcovsky, J. (2005). High-temperature adult-plant (HTAP) stripe rust resistance gene Yr36 from Triticum turgidum ssp. dicoccoides is closely linked to the grain protein content locus Gpc-B1. *Theoretical and Applied Genetics*, *112*(1), 97–105.

Verbyla, A. P., George, A. W., Cavanagh, C. R., & Verbyla, K. L. (2014). Whole-genome QTL analysis for MAGIC. *Theoretical and Applied Genetics*, *127*(8), 1753–1770.

Vieira Dos Santos, C., & Rey, P. (2006). Plant thioredoxins are key actors in the oxidative stress response. *Trends in Plant Science*, *11*(7), 329–334.

Vleeshouwers, V. G. A. A., & Oliver, R. P. (2014). Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. *Molecular Plant-Microbe Interactions : MPMI*, *27*(3), 196–206.

Wang, M. N., & Chen, X. M. (2013). First Report of Oregon Grape (Mahonia aquifolium) as an Alternate Host for the Wheat Stripe Rust Pathogen (Puccinia striiformis f. sp. tritici) Under Artificial Inoculation. *Plant Disease*, *97*(6), 839.

Wang, M. N., & Chen, X. M. (2015). Barberry Does Not Function as an Alternate Host for Puccinia striiformis f. sp. tritici in the U. S. Pacific Northwest Due to Teliospore Degradation and Barberry Phenology. *Plant Disease*, *99*(11), 1500–1506.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., … 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.

Wei, F., Gobelman-Werner, K., Morroll, S. M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P., & Wise, R. P. (1999). The Mla (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics*, *153*(4), 1929–1948.

Wei, J., & Xu, S. (2016). A Random-Model Approach to QTL Mapping in Multiparent Advanced Generation Intercross (MAGIC) Populations. *Genetics*, *202*(2), 471–486.

Wellings, C. R. (2003). Stripe Rust Head Infection. *Cereal Rust Report*, *1*(1), 1–2.

Wellings, C. R. (2007). Puccinia striiformis in Australia: A review of the incursion, evolution, and adaptation of stripe rust in the period 1979–2006. *Australian Journal of Agricultural Research*,

58(6), 567.

Wellings, C. R. (2009). Late Season Stripe Rust Damage: Head Infection and Unpredicted Variety Responses. *Cereal Rust Report*, *7*(10), 1–4.

Wellings, C. R. (2011). Global status of stripe rust: a review of historical and current threats. *Euphytica*, *179*(1), 129–141.

Wellings, C. R., Singh, R. P., Yahyaoui, A. H., Nazari, K., & McIntosh, R. A. (2009). The development and application of near-isogenic lines for monitoring cereal rust pathogens. *Proceedings of the Borlaug Global Rust Initiative Technical Workshop*, 77–87.

Wellings, C. R., Wright, D., Keiper, F., & Loughman, R. (2003). First detection of wheat stripe rust in Western Australia: Evidence for a foreign incursion. *Australasian Plant Pathology*, *32*, 321–322.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis. *Springer*.

Wu, C.-H., Krasileva, K., Banfield, M., Terauchi, R., & Kamoun, S. (2015). The "sensor domains" of plant NLR proteins: more than decoys? *Frontiers in Plant Science*, *6*, 134.

Wulff, B. B. H., & Moscou, M. J. (2014). Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. *Frontiers in Plant Science*, *5*, 174.

Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., Dixon, R. A., & Lamb, C. (2004). An extracellular aspartic protease functions in Arabidopsis disease resistance signaling. *The EMBO Journal*, *23*(4), 980–988.

Yahiaoui, N., Srichumpa, P., Dudler, R., & Keller, B. (2004). Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene Pm3b from hexaploid wheat. *The Plant Journal*, *37*(4), 528–538.

Yuan, C., Wu, J., Yan, B., Hao, Q., Zhang, C., Lyu, B., Ni, F., Caplan, A., Wu, J., & Fu, D. (2018). Remapping of the stripe rust resistance gene Yr10 in common wheat. *Theoretical and Applied Genetics*, *131*(6), 1253–1262.

Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, *14*(6), 415–421.

Zegeye, H., Rasheed, A., Makdis, F., Badebo, A., & Ogbonnaya, F. C. (2014). Genome-Wide Association Mapping for Seedling and Adult Plant Resistance to Stripe Rust in Synthetic Hexaploid Wheat. *PLoS ONE*, *9*(8), e105593.

Zeng, S.-M., & Luo, Y. (2006). Long-Distance Spread and Interregional Epidemics of Wheat Stripe Rust in China. *Plant Disease*, *90*(8), 980–988.

Zhang, P., McIntosh, R. A., Hoxha, S., & Dong, C. (2009). Wheat stripe rust resistance genes Yr5 and Yr7 are allelic. *Theoretical and Applied Genetics*, *120*(1), 25–29.

Zhang, W., Chen, S., Abate, Z., Nirmala, J., Rouse, M. N., & Dubcovsky, J. (2017). Identification and characterization of Sr13, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. *Proceedings of the National Academy of Sciences of the United States of America*.

Zhao, J., Wang, L., Wang, Z., Chen, X., Zhang, H., Yao, J., Zhan, G., Chen, W., Huang, L., & Kang, Z. (2013). Identification of eighteen Berberis species as alternate hosts of Puccinia striiformis f. sp. tritici and virulence variation in the pathogen isolates from natural infection of barberry plants in China. *Phytopathology*, *103*(9), 927–934.

Zhao, J., Zhang, H., Yao, J., Huang, L., & Kang, Z. (2011). Confirmation of Berberis spp. as alternate hosts of Puccinia striiformis f. sp. tritici on wheat in China. *Mygosystema*, *30*(6), 895.

Zhou, X. L., Wang, W. L., Wang, L. L., Hou, D. Y., Jing, J. X., Wang, Y., Xu, Z. Q., Yao, Q., Yin, J. L., & Ma, D. F. (2011). Genetics and molecular mapping of genes for high-temperature resistance to stripe rust in wheat cultivar Xiaoyan 54. *Theoretical and Applied Genetics*, *123*(3), 431–438.

Zimin, A. V, Puiu, D., Hall, R., Kingan, S., Clavijo, B. J., & Salzberg, S. L. (2017). The first near-complete assembly of the hexaploid bread wheat genome, Triticum aestivum. *GigaScience*.

8 APPENDICES

APPENDIX A: SUPPLEMENTARY MATERIAL FOR CHAPTER 1

Race name	Virulence profile
Madrigal_Lynx,v4	*Yr1,2,3,4,9,17,25*
Robigus	*Yr1,2,3,4,9,17,25,32*
Solstice_Oakley	*Yr1,2,3,4,6,9,17,25,32*
Solstice_Oakley,v7	*Yr1,2,3,4,6,7,9,17,25,32*
Tulsa	*Yr3.4.6.25.32*
PstS2,v27	*Yr2,6,7,8,9,25*
Triticale2006	*Yr2,6,7,8,10*
Warrior	*Yr1,2,3,4,6,7,9,17,25,32,Sp,Amb*
Kranich	*Yr1,2,3,6,7,8,9,17,25,32,Amb*
Warrior(-)	*Yr1,2,3,4,6,7,9,17,25,32,Sp*
Hereford	*Yr2,3,6,7,8,25,32*
Triticale2015	*Yr2,6,7,9,Sp*
PstS14	*Yr2,3,6,7,8,9,17,25,32,Sp*
Other	

Figure 8.1 Colour key for race changes in Europe as illustrated in Figure 1.2

APPENDIX B: SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Table 8.1. Summary of Recommended List yellow rust resistance ratings (RL yellow rust scores min-max), postulated *Yr* genes and Adult Plant Resistance (APR) reported to date for the eight founders of the NIAB Elite MAGIC population.

MAGIC founder	Years on RL	RL yellow rust scores (min-max)	All-stage resistance	APR
Alchemy	2006-2016	6-9	unknown	unknown
Brompton	2005-2009	8-9	unknown	unknown
Claire	1998-2017	5-9	*Yr2,Yr3,Yr4,Yr25* (d)	*QYr.niab-2D.1, QYr.niab-2D.2, QYr.niab-2B, QYr.niab-7B* (e)
Hereward	1991-2011	4-9	Yr3,Yr32(c) *Yr3,Yr4,Yr25,Yr32,YrSD*(d)	undetermined (c)
Rialto	1995-2001	4-6	*Yr6,Yr9* (c) *Yr9* (a&b)	unknown
Robigus	2003-2012	2-3	Unknown but postulated *Yr2,Yr32* (c)	unknown
Soissons	1995-2009	5-8	*Yr3a + Yr4a* (f)	undetermined (g)
Xi19	2002-2010	9-9	unknown	unknown

(a) Pathan et al., 2008, (b) Singh et al., 2008, (c) Hovmøller, 2007, (d) Lewis, 2006, (e) Powell et al., 2013, (f) de Vallavieille-Pope et al., 1990), (g) de Vallavieille-Pope et al., 2011. Unknown means that there have been no reports of yellow rust resistance. Undetermined means that resistance has been reported but the underlying factors conferring this resistance remain unknown.
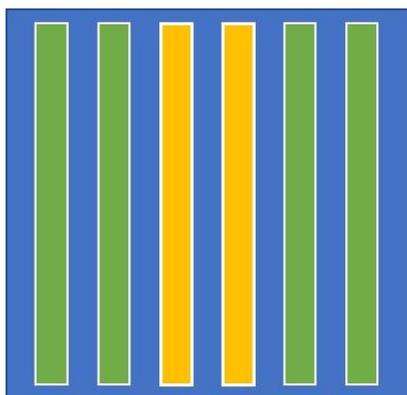
Figure 8.2 Layout of NIAB15 pathology trial plots

In a 1 x 1 m plot, each treatment (green rows) is exposed to a spreader (yellow rows). The spreader consists of a mix of the wheat varieties Victo and Vuka.

Table 8.2 MAGIC line controls for OSG15 and ROTH15

MAGIC line	Check type	2012 nursery YR score[a]	2014 nursery YR score[b]
MEL_007-2c	Susceptible	5	5
MEL_005-3	Susceptible	4	5
MEL_018-2	Susceptible	5	4
MEL_023-4	Resistant	2	0
MEL_027-1e	Resistant	1	0
MEL_029-2b	Susceptible	5	4
MEL_038-1a	Resistant	1	0
MEL_040-1c	Resistant	1	0
MEL_042-1b	Resistant	1	0
MEL_045-1a	Resistant	2	0
MEL_046-7	Susceptible	4	4
MEL_052-7	Susceptible	4	5
MEL_084-4	Susceptible	7	5
MEL_092-1	Susceptible	4	4
MEL_093-8	Resistant	3	1
MEL_104-3	Susceptible	4	4

MEL_115-3	Susceptible	4	4
MEL_121-1f	Resistant	1	0
MEL_151-6	Resistant	1	1
MEL_195-1d	Resistant	1	0

Selection of MAGIC lines was based on yellow rust infection data from the 2012 and 2014 multiplication nurseries. a: 1-9 Breeders scale; b: 1-5 scale. See Table 8.3 for 2012 assessment scale and Figures 8.3 and 8.4 for 2014 assessment scale.

Table 8.3 Ungenotyped F_8 MAGIC lines used in 2016 trials

MAGIC line	Type	2015 YR infection score[a]
MEL_001-2a	Susceptible	95
MEL_005-1	Resistant	0
MEL_007-2c	Susceptible	95
MEL_008-7	Susceptible	99
MEL_020-2	Resistant	0
MEL_020-6	Resistant	0.5
MEL_020-8	Resistant	0.25
MEL_024-6	Susceptible	98
MEL_038-1c	Resistant	0.01
MEL_045-1a	Resistant	0
MEL_046-5	Susceptible	99
MEL_058-6	Susceptible	100
MEL_070-1f	Susceptible	100
MEL_073-1	Resistant	0
MEL_078-1cB	Susceptible	99

MAGIC line	Type	2015 YR infection score[a]
MEL_083-4	Susceptible	99
MEL_084-4	Susceptible	100
MEL_086-1	Susceptible	100
MEL_087-1f	Resistant	0.75
MEL_094-2a	Susceptible	100
MEL_095-5	Susceptible	99
MEL_098-8	Susceptible	100
MEL_107-2a	Susceptible	100
MEL_115-3	Susceptible	99
MEL_118-1a	Resistant	1
MEL_126-1	Resistant	0.5
MEL_128-2	Susceptible	98
MEL_132-3	Susceptible	100
MEL_133-8	Susceptible	100
MEL_134a-1	Resistant	0
MEL_135-1f	Susceptible	99
MEL_136-3aA	Resistant	0.01
MEL_138-2	Resistant	1
MEL_138-8	Susceptible	99
MEL_141-3	Resistant	0.75
MEL_147-2	Resistant	0
MEL_151-6	Resistant	0
MEL_162-8	Resistant	0.01
MEL_167-2	Susceptible	100

MAGIC line	Type	2015 YR infection score[a]
MEL_174-5	Susceptible	99
MEL_181-6	Resistant	0
MEL_194a-2	Resistant	0
MEL_195-1b	Resistant	0.5
MEL_201-8	Susceptible	95
MEL_203-1	Resistant	0.5
MEL_206a-1b	Resistant	0
MEL_208-8	Resistant	0
MEL_210-1a	Resistant	0.5

a: The 2015 YR assessment scores are the mean of all final infection scores of NIAB15, OSG15 and ROTH15.

Table 8.4 Breeder's scale for 2012 YR assessments

Score	Disease symptoms
1	No infection observed
2	1 stripe per tiller
3	2 stripes per tiller
4	Most tillers infected but some top leaves uninfected
5	All leaves infected but leaves appear green overall
6	Leaves appear half infected half green
7	Leaves appear more infected than green
8	Very little green leaf tissue left
9	Leaves dead – no green tissue left

The top 4 leaves of all plants are inspected. Adapted from (HGCA, 2011)

Figure 8.3 Scoring scale (0-5) used to assess YR leaf infection in the 2014 MAGIC multiplication nursery: scores 0-2.

A:0 – No infection observed; B:1 – Some tillers infected; C:2 – All tillers infected with several stripes. Observations were made over the whole plot, ignoring any naturally senescent tissue.

Figure 8.4 Scoring scale used to assess YR leaf infection in the 2014 MAGIC multiplication nursery: scores 3-5.

A:3 – All leaves infected at 25-50 %; B:4 – all leaves infected at 51-75 %; C:5 – Little to no green tissue left. Observations were made over the whole plot, ignoring any naturally senescent tissue.
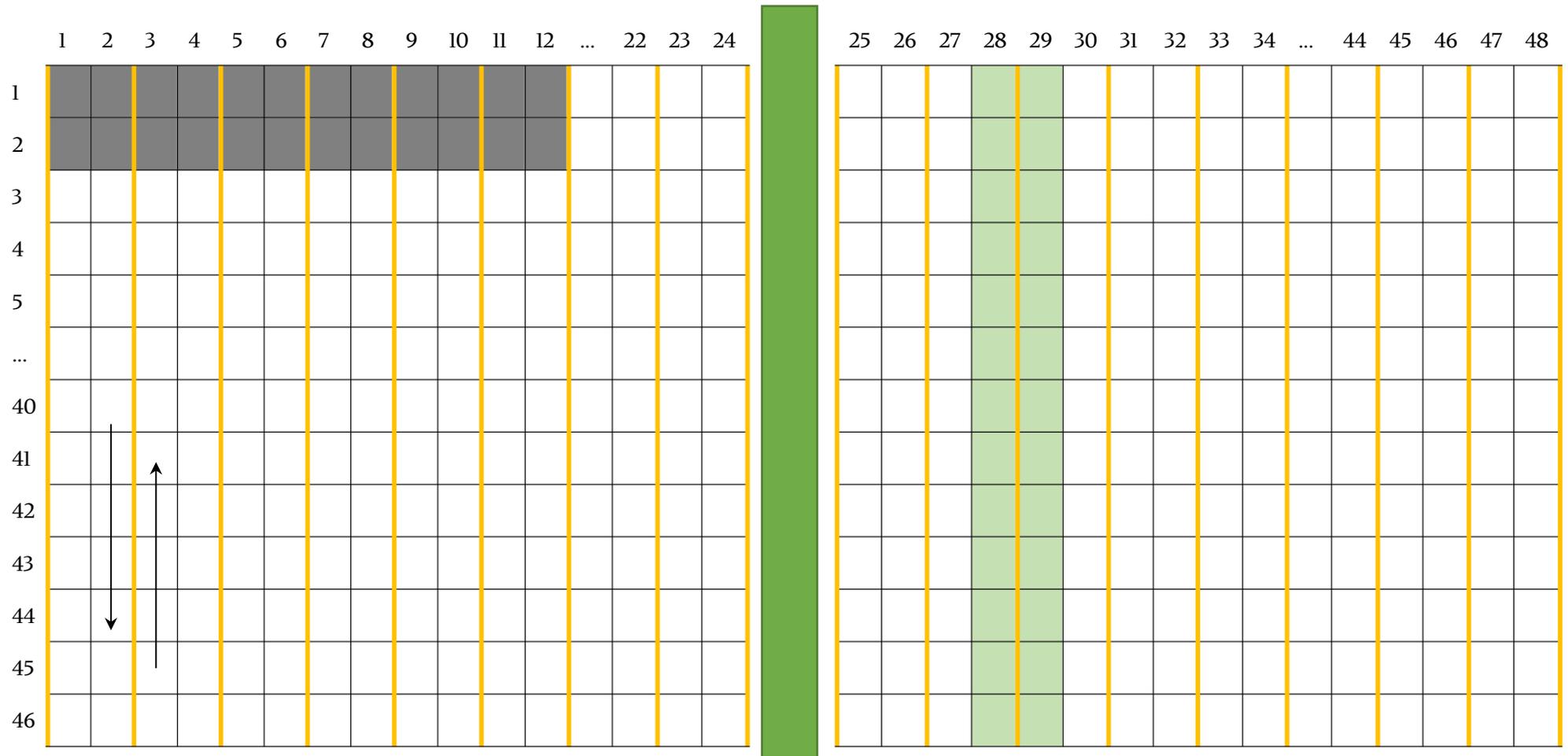
Figure 8.5 NIAB15 pathology trial layout

The trial is 48 columns wide and 46 rows long. Replicates are divided by a discard traverse (dark green rectangle). Rep 1 is columns 1 to 24. Rep 2 is columns 25 to 48. Thick yellow lines: spreader double rows. Light green plots: whole traverse with a double spreader row running in the middle and treatment plots either side of it. Grey plots: block size of 12 x 2 plots. Blank plots: treatment plots. Arrows indicate the direction in which the trial was walked when scoring.
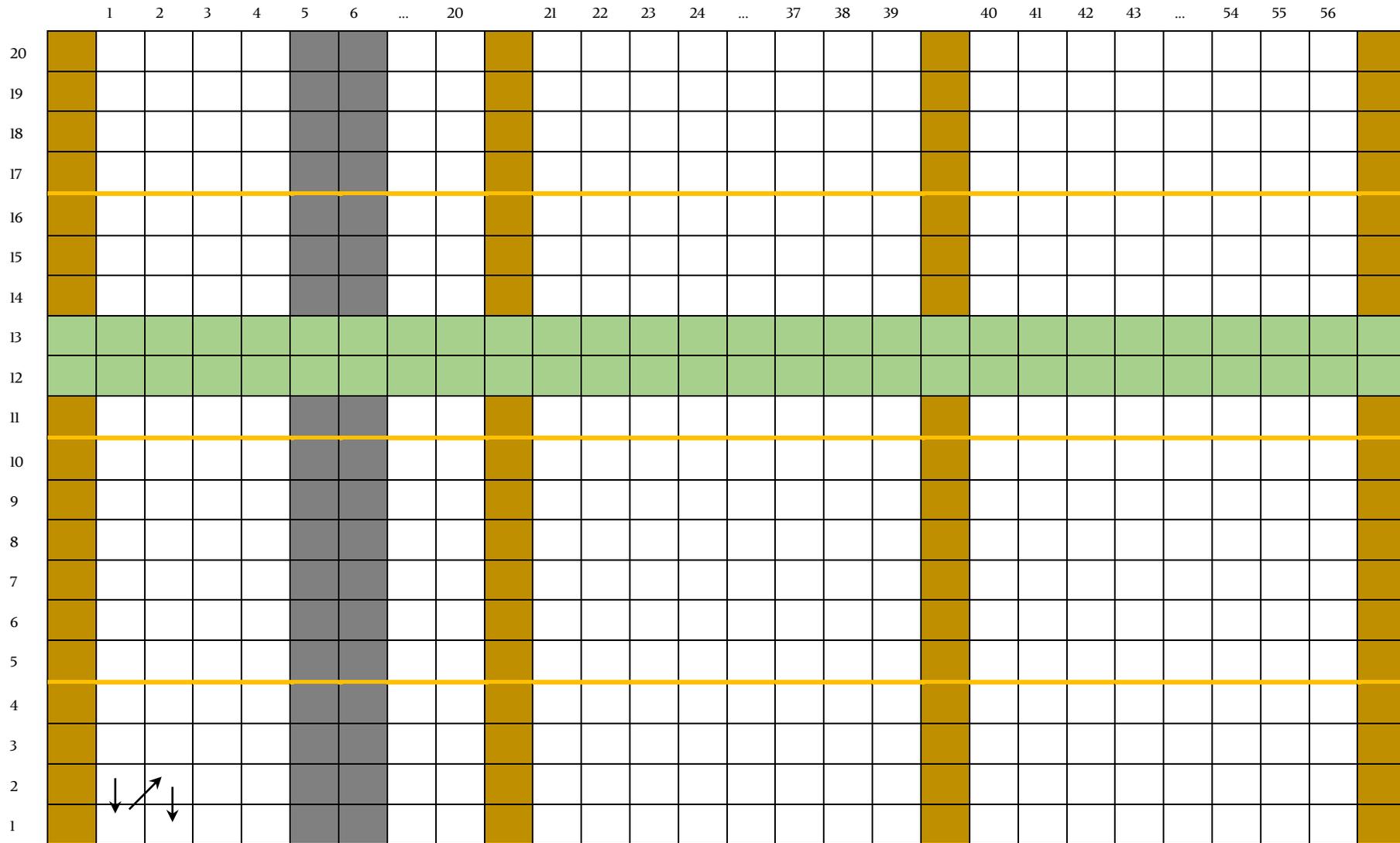
Figure 8.6 ROTH15 pathology trial layout

The trial is 20 columns wide and 60 rows long. Column numbers specify treatment columns. Thick yellow lines represent the spreader traverses. Brown plots indicate Vuka plots. Grey plots indicate block size of 20 x 2. Light green plots represent a whole traverse. Blank plots indicate treatment plots. Arrows indicate the direction in which the trial was walked when scoring. The trial was walked from left to right within the traverse, as indicated by the arrows, and in a serpentine manner between traverses.
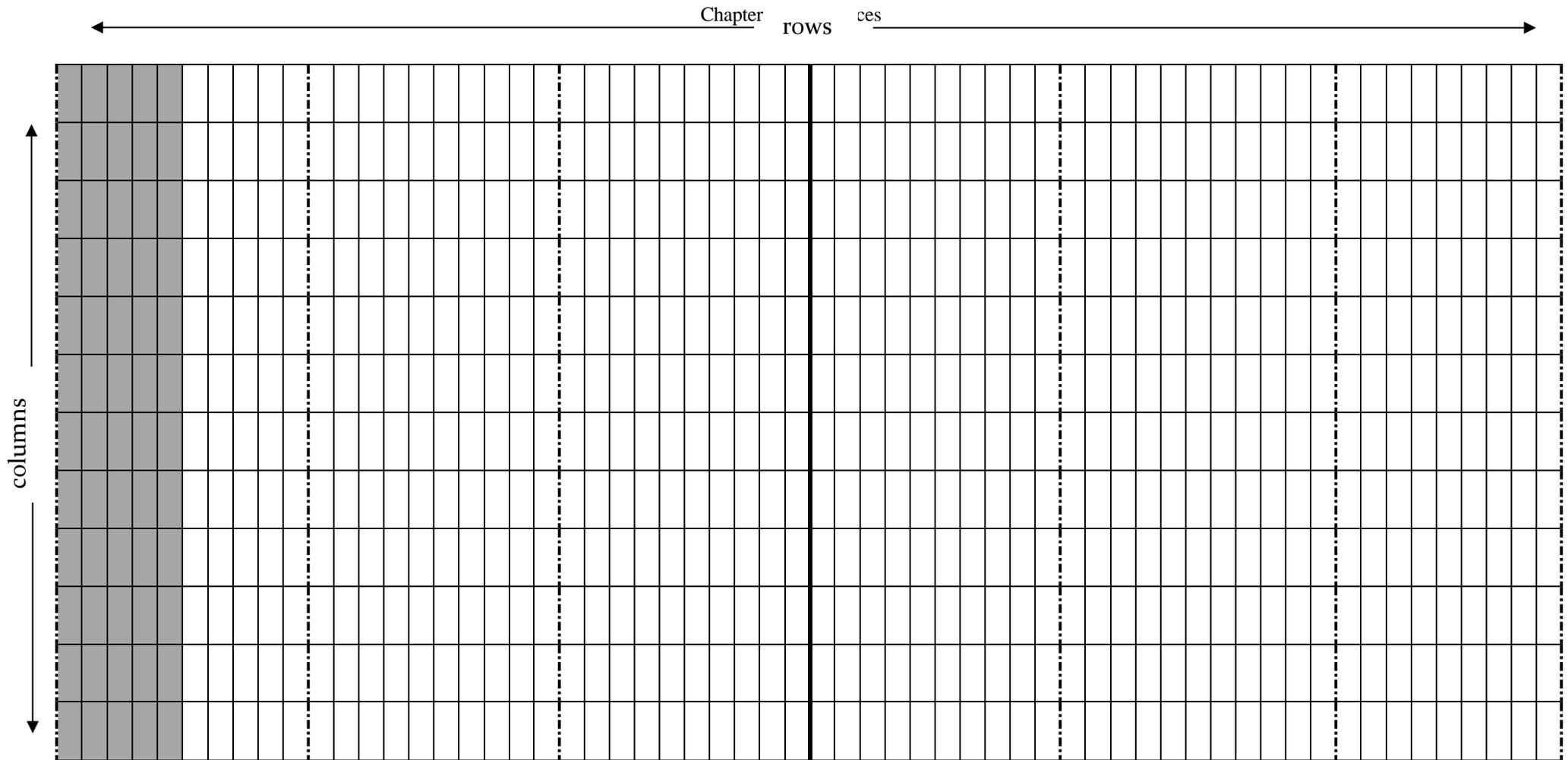
Figure 8.7 2016 pathology trials layout

The trial is 20 columns wide and 60 rows long and is divided into two reps, as indicated by the thick black line. Three blocks, nested within each replicate, are outlined by dashed lines. Two 20x5 plot sub-blocks, nested within each block, are highlighted in grey (not to scale). NIAB16 and OSG16 have different spreader layouts. These were excluded from this figure to highlight the blocking structure and are specified in Table 2.3, Section 2.1.2. Blank plots indicate treatment plots.

APPENDIX C: SUPPLEMENTARY MATERIAL FOR CHAPTER 3

8.1 Residual diagnostic plots for NIAB15, NIAB16 and OSG16
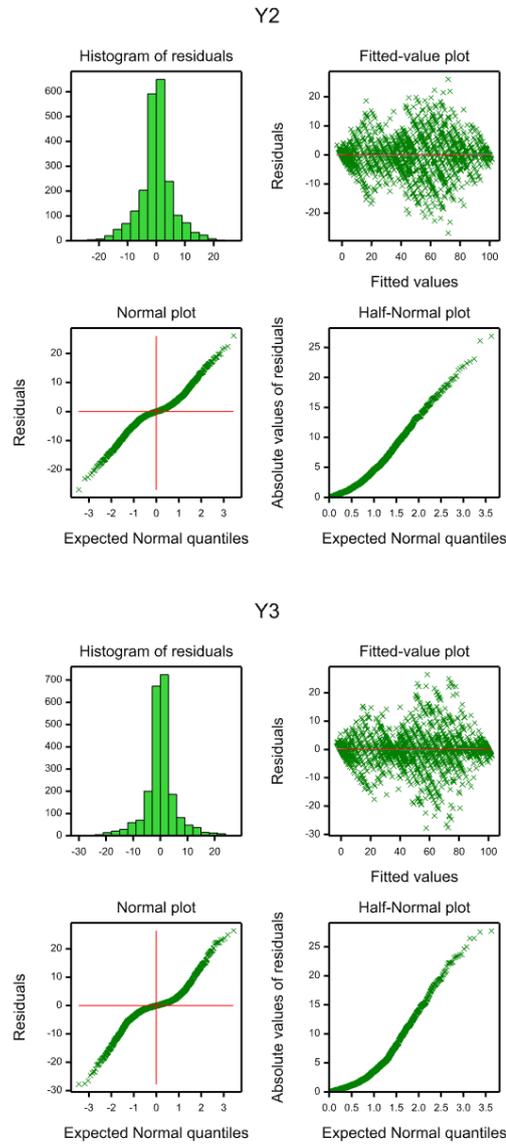
Figure 8.8 Residual diagnostic plots for yellow rust disease severity data from NIAB15

Y2: second score; Y3: third and final score.
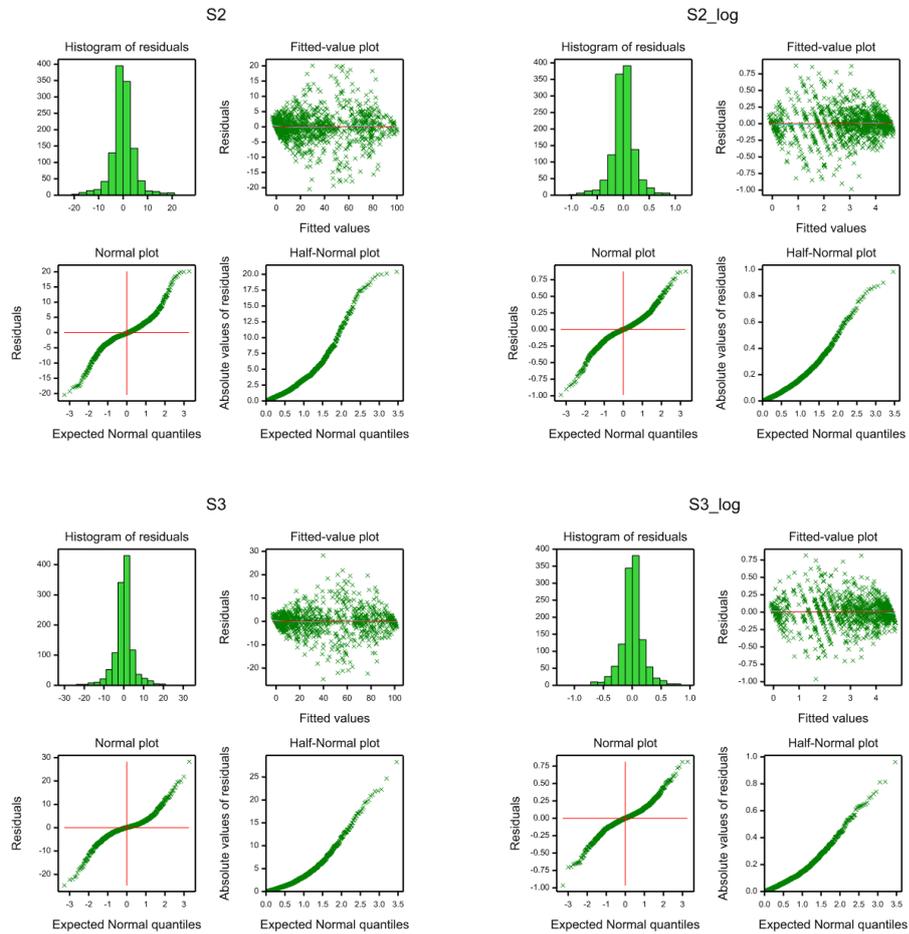
Figure 8.9 Residual diagnostic plots for yellow rust disease severity data from OSG16, for original (S2 and S3) and log-transformed (S2_log) and S3_log) score.

S2: second score; S3: third score.
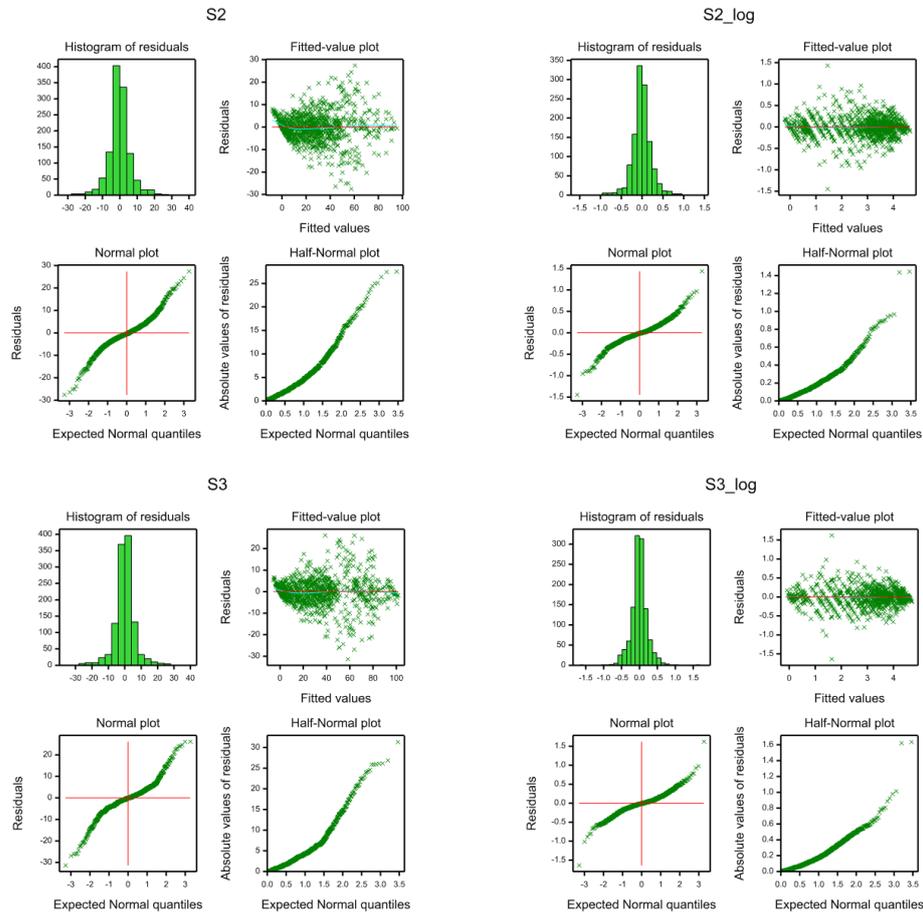
Figure 8.10 Residual diagnostic plots for yellow rust disease severity data from NIAB16, for original (S2 and S3) and log-transformed (S2_log) and S3_log) score.

S2: second score; S3: third score.

8.2 Results from statistical analyses for all MAGIC pathology trials

Table 8.5 Summary of REML estimates for NIABl5

YR score	Model	REML estimates						H$_2$	σ^2$_e$
		Random			Fixed				
		Term	Est.	SE	Term	Wald	*Chi pr*		
S2	Block + column	block	9.47	2.62	G	20055.15	<0.001	0.94	86.71
		column	1.36	1.04					
S3	Block	rep	1.3	2.2	G	21416.80	<0.001	0.94	104.8
		block	15.2	3.7					

Covariance estimates of the mixed linear model based on Restricted Maximum Likelihood (REML). S2-3: YR scores 2-3; Est.: Estimates; SE: Standard Error; B: Blocking; C: Column; H$_2$: broad sense heritability; σ^2$_e$: residual.

Table 8.6 Summary of spatial analysis and REML estimates for OSG15

YR score	Model	AR1 x AR1 estimates					REML estimates (fixed term for G)		H_2
		Row	SE	Column	SE	σ^2_e	Wald	*Chi pr*	
S2	Spatial	0.08453	0.39865	0.8090	0.0524	35.01	94706.66	<0.001	0.95
S3	Spatial	0.8073	0.0698	0.5659	0.0727	72.39	142371.54	<0.001	0.95

Spatial analysis estimates and covariance estimates of the mixed linear model based on Restricted Maximum Likelihood (REML). S2-3: YR scores 2-3; Est.: Estimates; SE: Standard Error; B: Blocking; C: Column; H_2: broad sense heritability; σ^2_e : residual.

Table 8.7 Summary of spatial analysis and REML estimated for ROTH15

YR score	Model	AR1 x AR1 estimates					REML estimates (fixed term for G)		H_2
		Row	SE	Column	SE	σ^2_e	Wald	*Chi pr*	
S2	Spatial	0.2017	0.3346	-0.5253	0.0963	44.64	40105.26	<0.001	0.95
S3	Spatial	-0.07908	0.40253	0.3368	0.1162	59.82	29291.30	<0.001	0.95

Spatial analysis estimates and covariance estimates of the mixed linear model based on Restricted Maximum Likelihood (REML). S2-3: YR scores 2-3; Est.: Estimates; SE: Standard Error; H_2: broad sense heritability; σ^2_e: residual.

Table 8.8 Summary of REML estimates for NIAB16

YR value	Model	REML estimates							H_2	σ^2_e
		Fixed			Random		*Chi pr*			
		Term	Est.	SE	Term	Est.				
S2	B + Wlk	R	2.12	4.45	G	8943.97	<0.001		0.90	61.40
		R.B	2.27	2.44						
		R.B.SB	0.50	1.26						
		Wlk	13.19	6.65						
S2_log	B + Wlk	R	0.0027	0.0054	G	16411.81	<0.001		0.94	0.117
		R.B	0.0025	0.0029						
		Wlk	0.0094	0.0058						
S3	B + Wlk	R	7.29	11.56	G	10745.39	<0.001		0.91	74.64
		R.B	2.10	2.19						
		Wlk	6.16	3.59						
S3_log	B + Wlk	R	0.0078	0.0125	G	15619.52	<0.001		0.93	0.116
		R.B	0.0022	0.0026						
		Wlk	0.0053	0.0038						

Covariance estimates of the mixed linear model based on Restricted Maximum Likelihood (REML). S2-3: YR scores 2-3; Est.: Estimates; SE: Standard Error; B: Blocking; SB: Sub-Block; C: Column; R: Rep; Wlk: Walking route taken during YR assessment.

Table 8.9 Summary of spatial analysis estimates for OSG16

YR score	Model	AR1 x AR1 estimates					REML						H$_2$
		row	SE	column	SE	σ^2$_e$	Fixed term	Est.	*Chi pr*	Random term	Est.	SE	
S2	Spatial + **Wlk** + lin(row)	0.14	0.07	0.08	0.07	52.87	G	18067.37	<0.001	Wlk	2.72	1.94	0.94
							Lin(row)	12.45					
S2_log	Spatial + **Wlk**	0.30	0.07	0.08	0.08	0.107	G	22373.64	<0.001	Wlk	0.0034	0.0031	0.94
S3	Spatial + **Blocking**	0.08	0.08	0.18	0.07	54.97	G	24249	<0.001	R.B	2.95	2.59	0.95
										R.B.SB	0.4319	1.26	
S3_log	Spatial + **Blocking** + **C**	0.28	0.07	0.14	0.08	0.077	G	34121.15	<0.001	B	0.0034	0.0030	0.95
										Column	0.0022	0.0018	

Covariance estimates of the mixed linear model based on Restricted Maximum Likelihood (REML). S2-3: YR scores 2-3; Random terms are highlighted in bold.

Est.: Estimates; SE: Standard Error; B: Blocking; SB: Sub-Block; C: Column; R: Rep; Wlk: Walking route taken during YR assessment; S: Spatial Analysis; lin(row):

linear trend across rows. σ^2$_e$: residual error.

8.3 Manhattan plots for Single Marker Analysis and Interval Mapping

Figure 8.11 Manhattan plots from NIAB15 Single Marker Analysis

P values expressed as log10(*p*) are uncorrected for False Discovery Rates
The dotted line represents a significance threshold of -log10(p)= 4 for indication. When P = 0, -log10(p) was arbitrarily set at 16 to avoid infinite values. S2: second score, S3: third and final score.
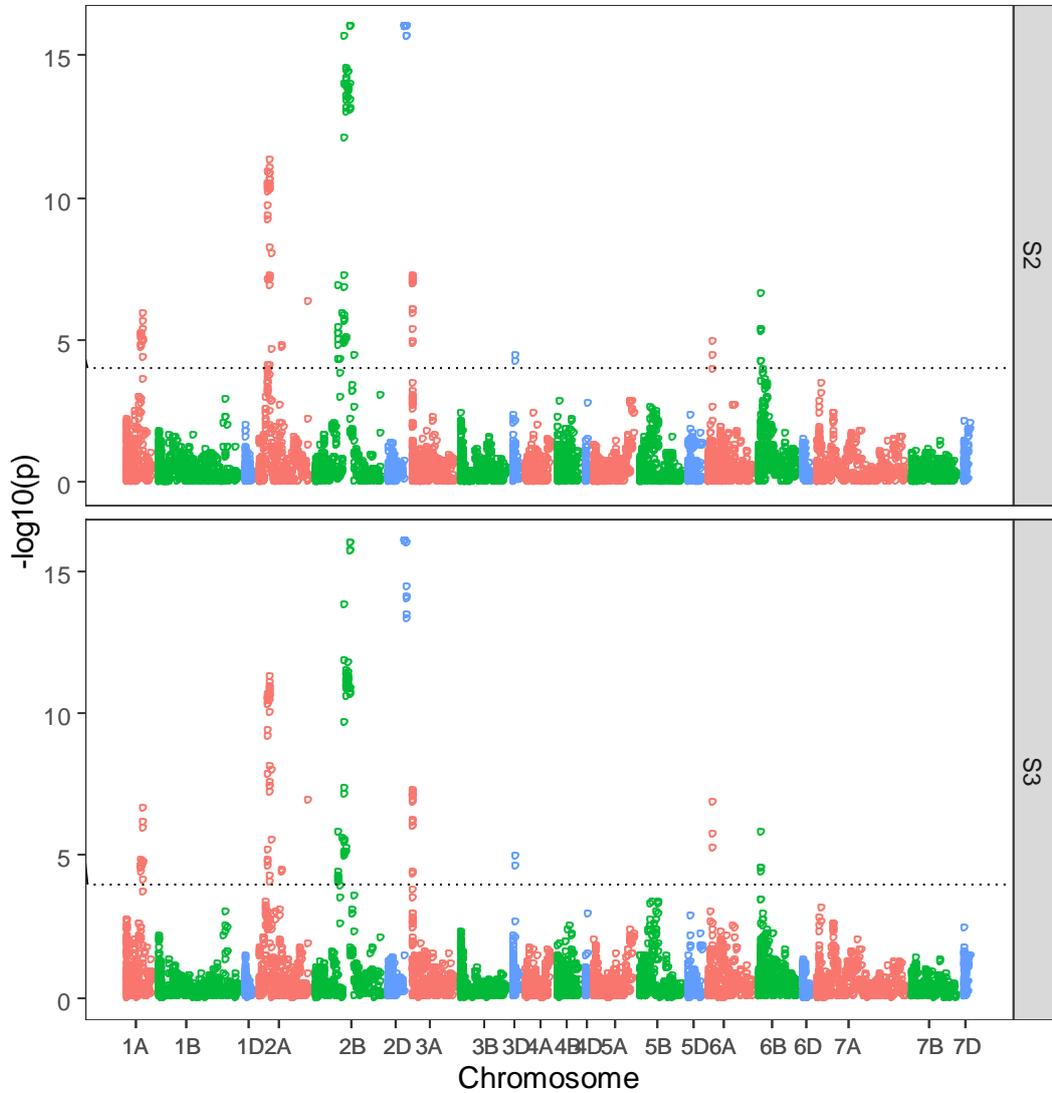
Figure 8.12 Manhattan plots from OSG15 Single Marker Analysis

P values expressed as log10(*p*) are uncorrected for False Discovery Rates The dotted line represents a significance threshold of -log10(p)= 4 for indication. When P = 0, -log10(p) was arbitrarily set at 16 to avoid infinite values. S2: second score, S3: third and final score.
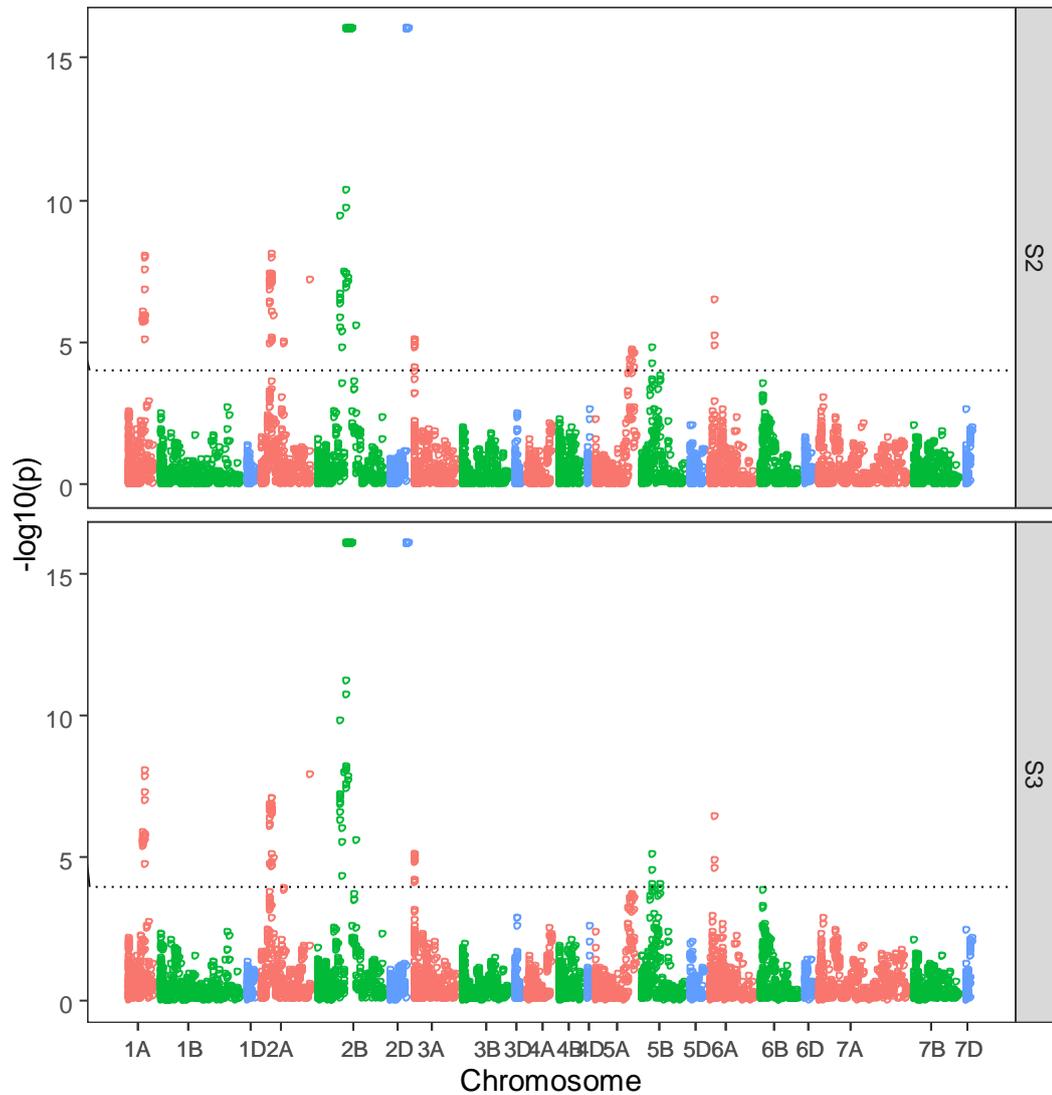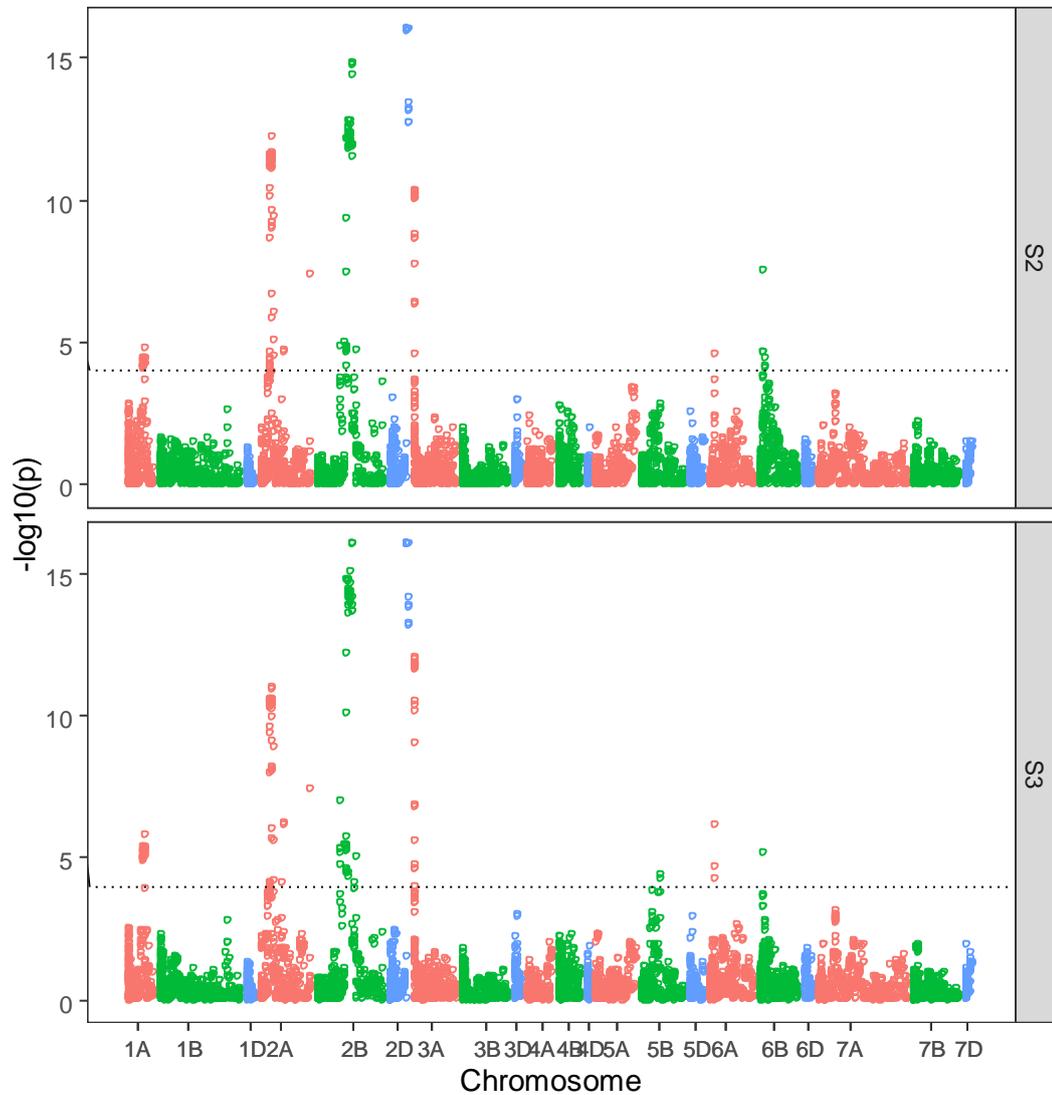
Figure 8.13 Manhattan plots from ROTH15 Single Marker Analysis

P values expressed as log10(*p*) are uncorrected for False Discovery Rates The dotted line represents a significance threshold of -log10(p)= 4 for indication. When P = 0, -log10(p) was arbitrarily set at 16 to avoid infinite values. S2: second score, S3: third and final score.
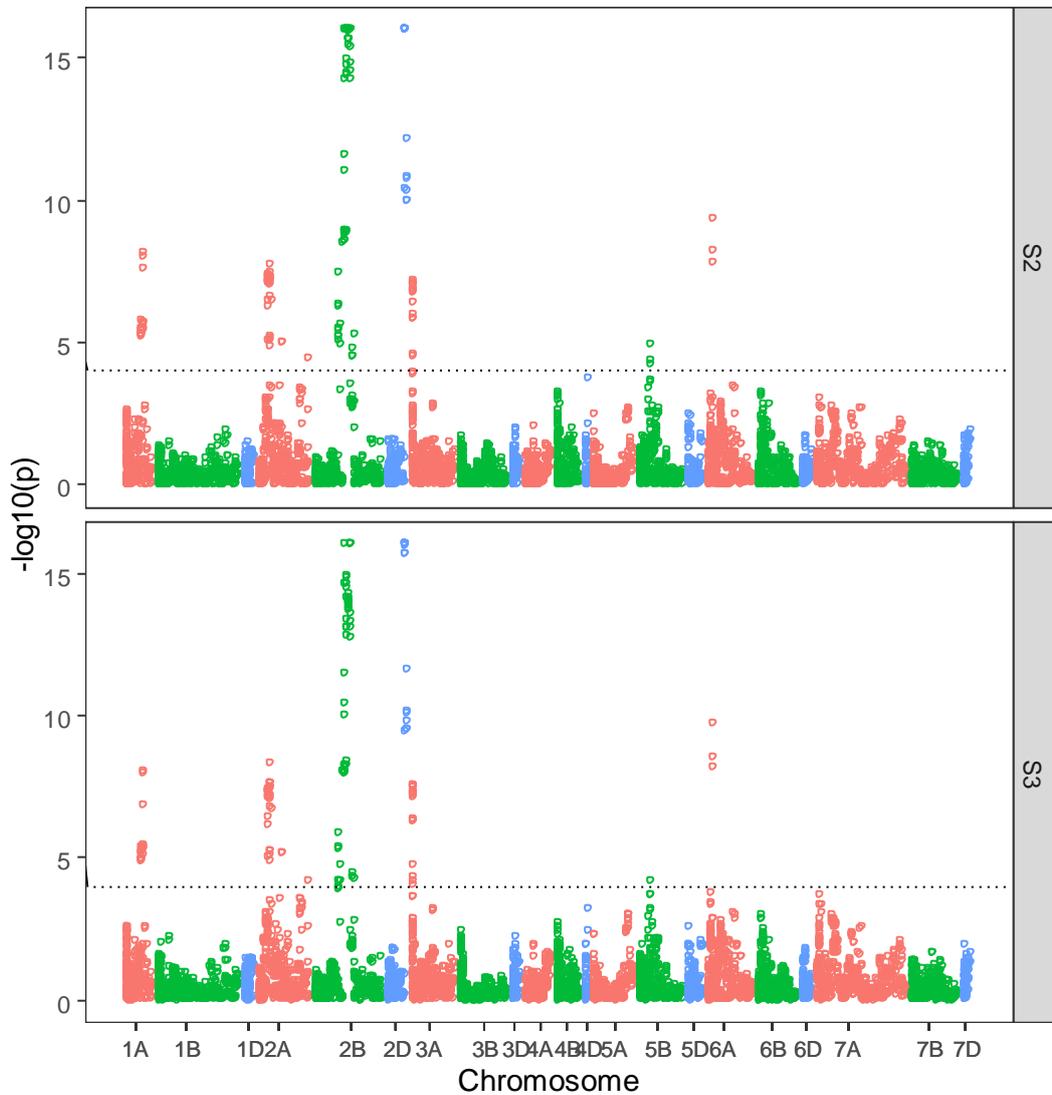
Figure 8.14 Manhattan plots from OSG16 Single Marker Analysis

P values expressed as log10(*p*) are uncorrected for False Discovery Rates and are based on log-transformed adjusted disease severity means. The dotted line represents a significance threshold of -log10(p)= 4 for indication. When P = 0, -log10(p) was arbitrarily set at 16 to avoid infinite values. S2: second score, S3: third and final score.
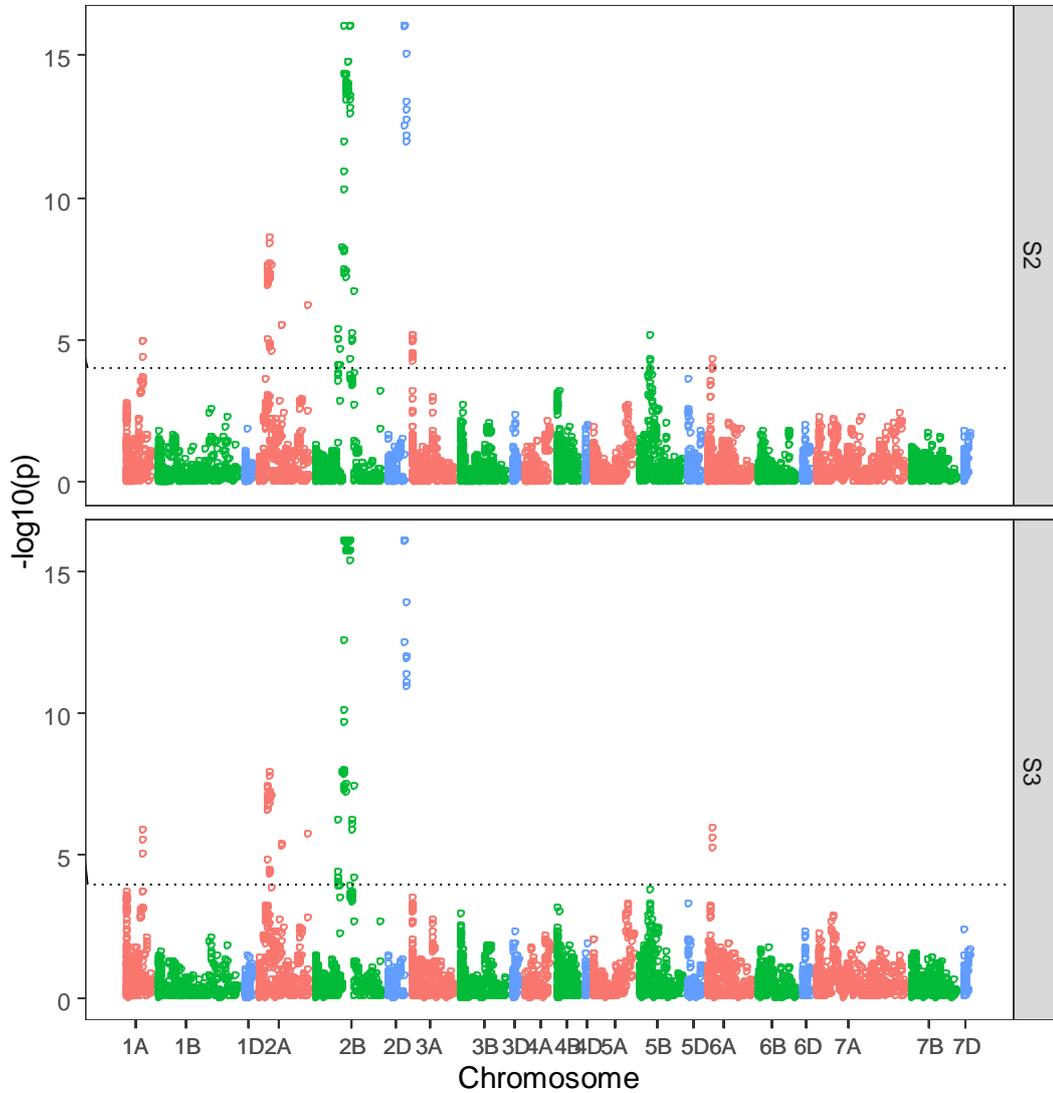
Figure 8.15 Manhattan plot from NIAB16 Single Marker Analysis

P values expressed as log10(*p*) are uncorrected for False Discovery Rates and are based on log-transformed adjusted disease severity means. The dotted line represents a significance threshold of - log10(p)= 4 for indication. When P = 0, -log10(p) was arbitrarily set at 16 to avoid infinite values. S2: second score, S3: third and final score

Figure 8.16 Manhattan plot from NIAB15 Interval Mapping based on founder haplotypes.

P values expressed as log10(*p*) are uncorrected for False Discovery Rates. Chromosome numbers correspond to chromosome name as follows; 1: 1A, 2: 1B, 3: 1D.....19: 7A, 20: 7B, 21: 7D.
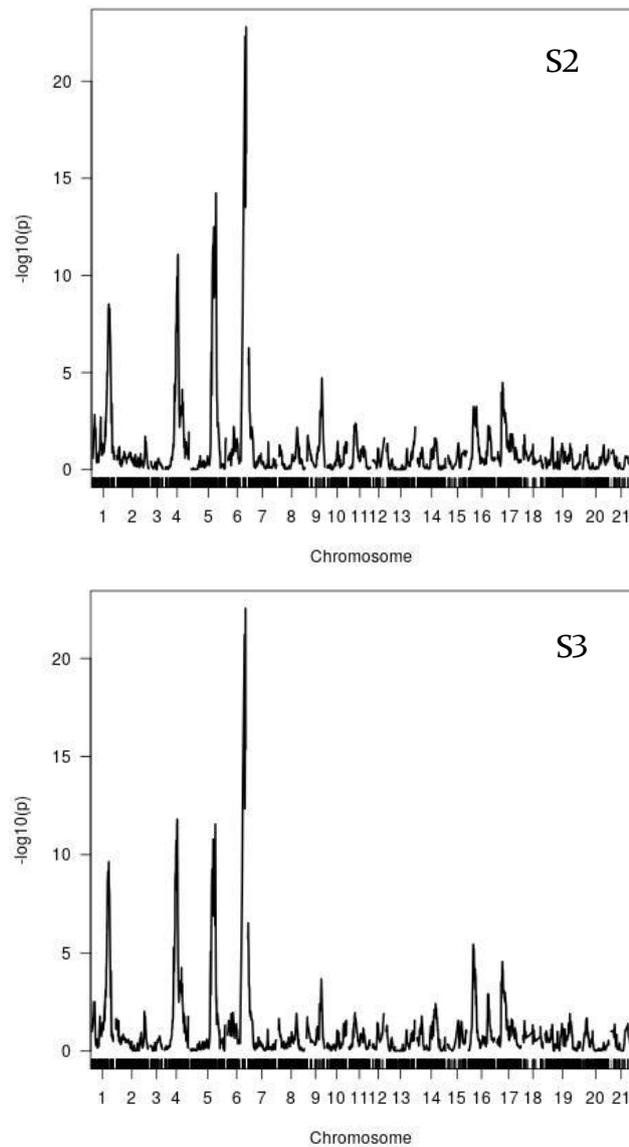
Figure 8.17 Manhattan plot from OSG15 Interval Mapping based on founder haplotypes.

P values expressed as log10(*p*) are uncorrected for False Discovery Rates. Chromosome numbers correspond to chromosome name as follows; 1: 1A, 2: 1B, 3: 1D.....19: 7A, 20: 7B, 21:
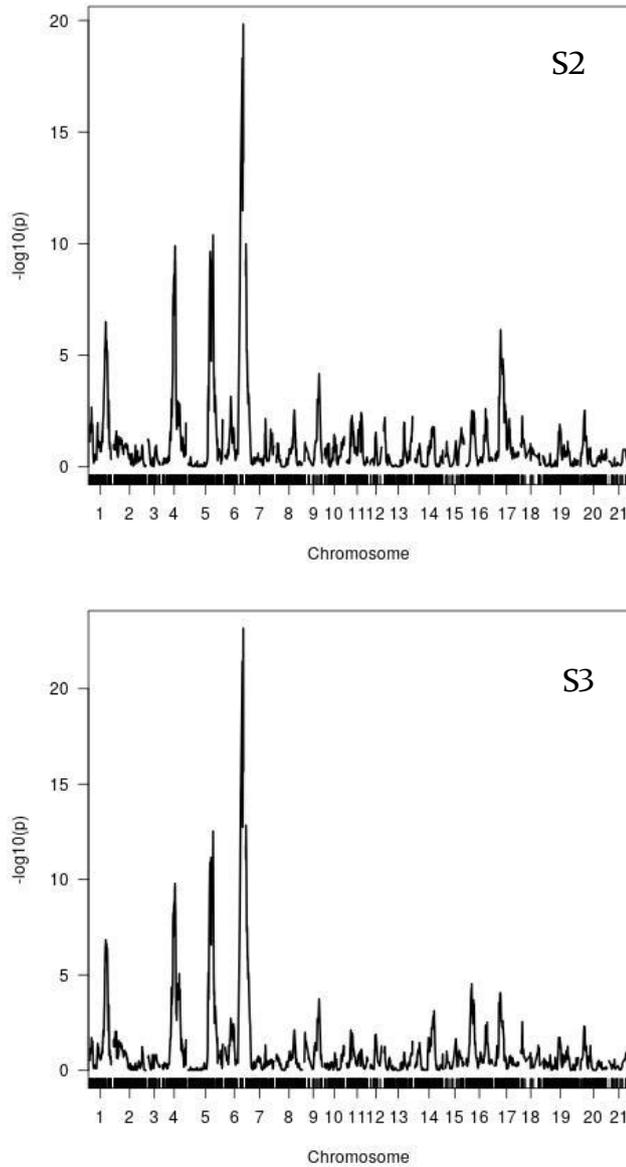
Figure 8.18 Manhattan plot from ROTH15 Interval Mapping based on founder haplotypes.

P values expressed as log10(*p*) are uncorrected for False Discovery Rates. Chromosome numbers correspond to chromosome name as follows; 1: 1A, 2: 1B, 3: 1D.....19: 7A, 20: 7B, 21: 7D.
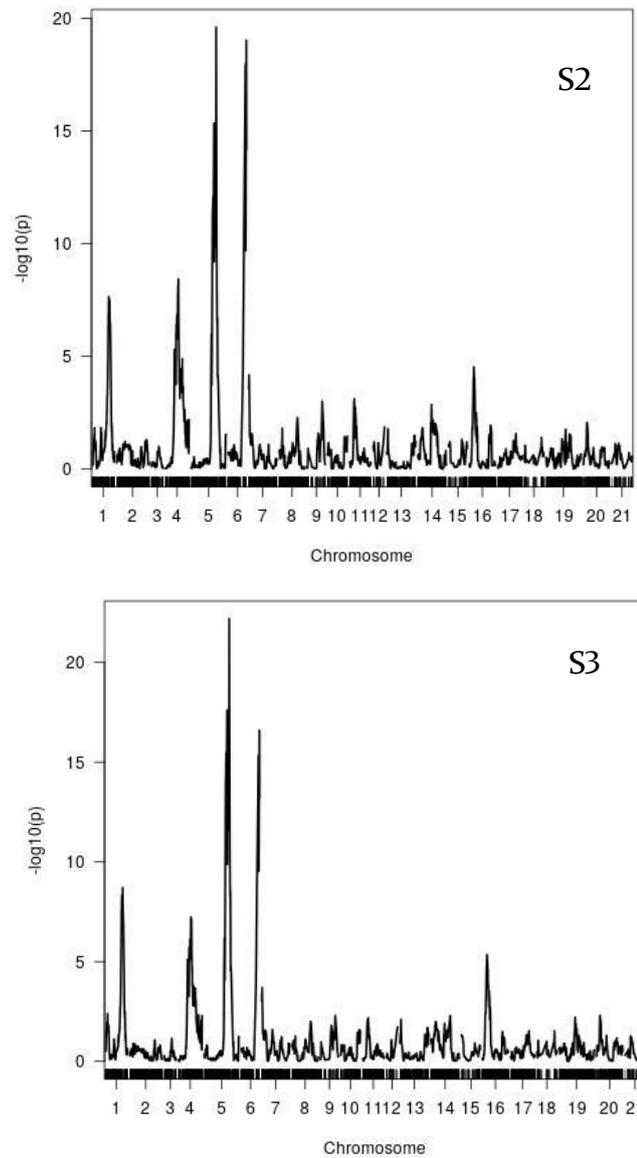
Figure 8.19 Manhattan plot from NIAB16 Interval Mapping based on founder haplotypes.

P values expressed as log10(*p*) are uncorrected for False Discovery Rates. Chromosome numbers correspond to chromosome name as follows; 1: 1A, 2: 1B, 3: 1D.....19: 7A, 20: 7B, 21: 7D.
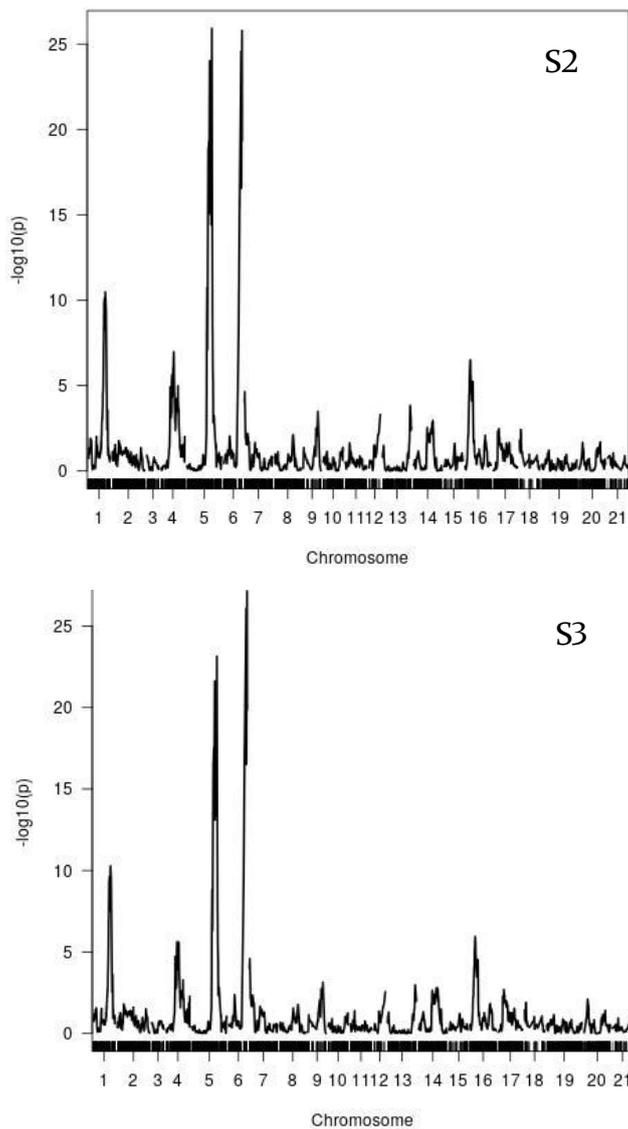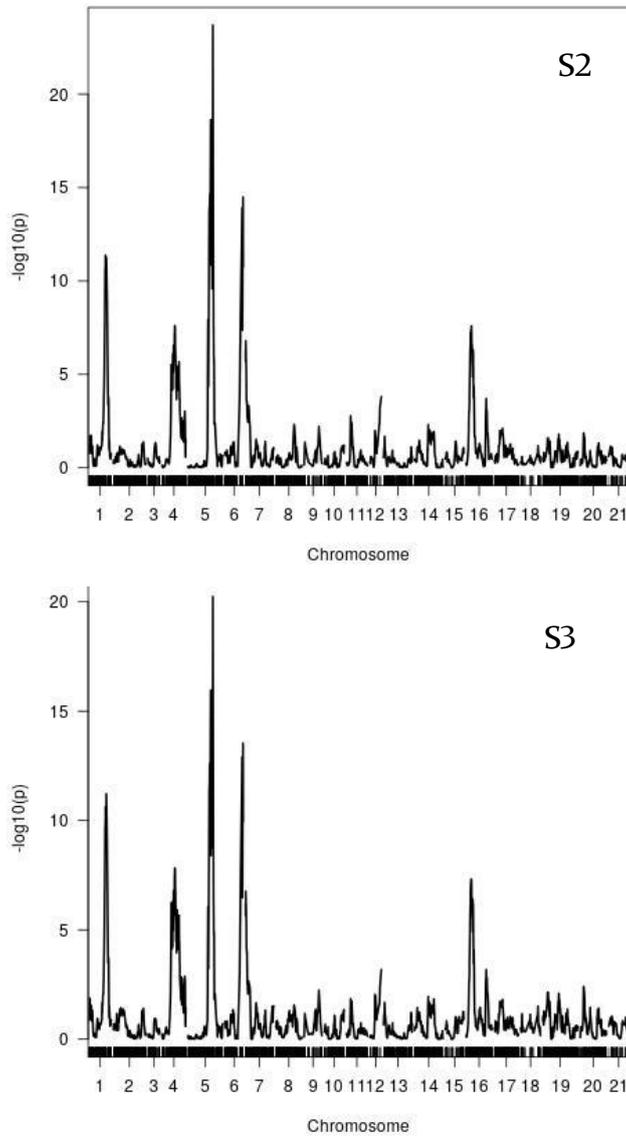
Figure 8.20 Manhattan plot from ROTH16 Interval Mapping based on founder haplotypes.

P values expressed as log10(*p*) are uncorrected for False Discovery Rates. Chromosome numbers correspond to chromosome name as follows; 1: 1A, 2: 1B, 3: 1D.....19: 7A, 20: 7B, 21: 7D.

Table 8.10 Flanking markers of the nine QTLs conferring YR resistance in the MAGIC population in 2015-16

QTL	Chr	Flanking markers		Genetic map position, cM		QTL interval, cM
		left	right	left	right	
QYr.niab-1A.1	1A	wsnp_Ex_c12399_19776420	BS00065268_51	176.51	193.67	17.17
QYr.niab-2A.1	2A	BS00055512_51	Kukri_c13006_731	133.18	152.47	19.29
QYr.niab-2A.2	2A	Kukri_c365_345	BS00064055_51	256.82	259.39	2.56
QYr.niab-2B.1	2B	wsnp_Ex_c1016_1943827	BS00043338_51	263.29	281.91	18.62
QYr.niab-2D.1	2D	RFL_Contig1128_620	BS00010685_51	188.01	198.86	10.85
QYr.niab-3A.1	3A	RAC875_c46403_277	BS00066230_51	0.00	16.64	16.64
QYr.niab-6A.1	6A	Kukri_c14679_913	BobWhite_c11652_421	46.86	65.10	18.24
QYr.niab-6A.2	6A	BobWhite_c24848_219	CAP7_rep_c8019_110	73.64	75.69	2.05
QYr.niab-6B.1	6B	IAAV8704	CAP12_c1784_424	56.49	69.86	13.37

Genetic map position and QTL interval are from the MAGIC genetic map (Gardner et al. 2016).

8.4 MAGIC founder contributions for *QYr.niab-2D.1*

Figure 8.21 MAGIC founder effects at the *QYr.niab-2D.1* locus contributing towards YR resistance in 2015

Grey bars: Founder contributions based on Interval Mapping (0 covariates) relative to Xi19. Black bars: Founder contributions based on Haplotype Analysis, relative to Alchemy. Al: Alchemy, Br: Brompton, Cl: Claire, He: Hereward, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19.

Figure 8.22 MAGIC founder effects at the *QYr.niab-2D.1* locus contributing towards YR resistance in 2016

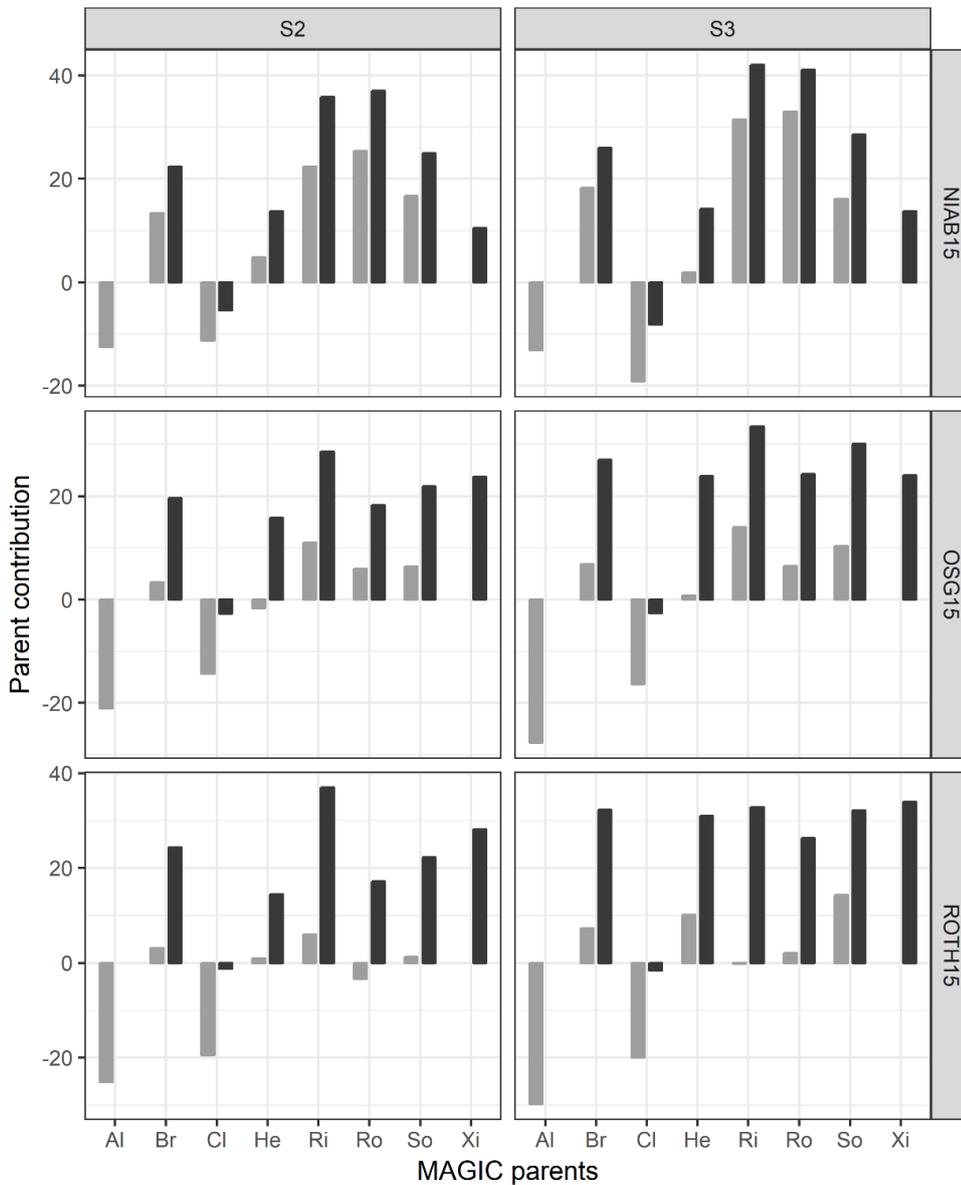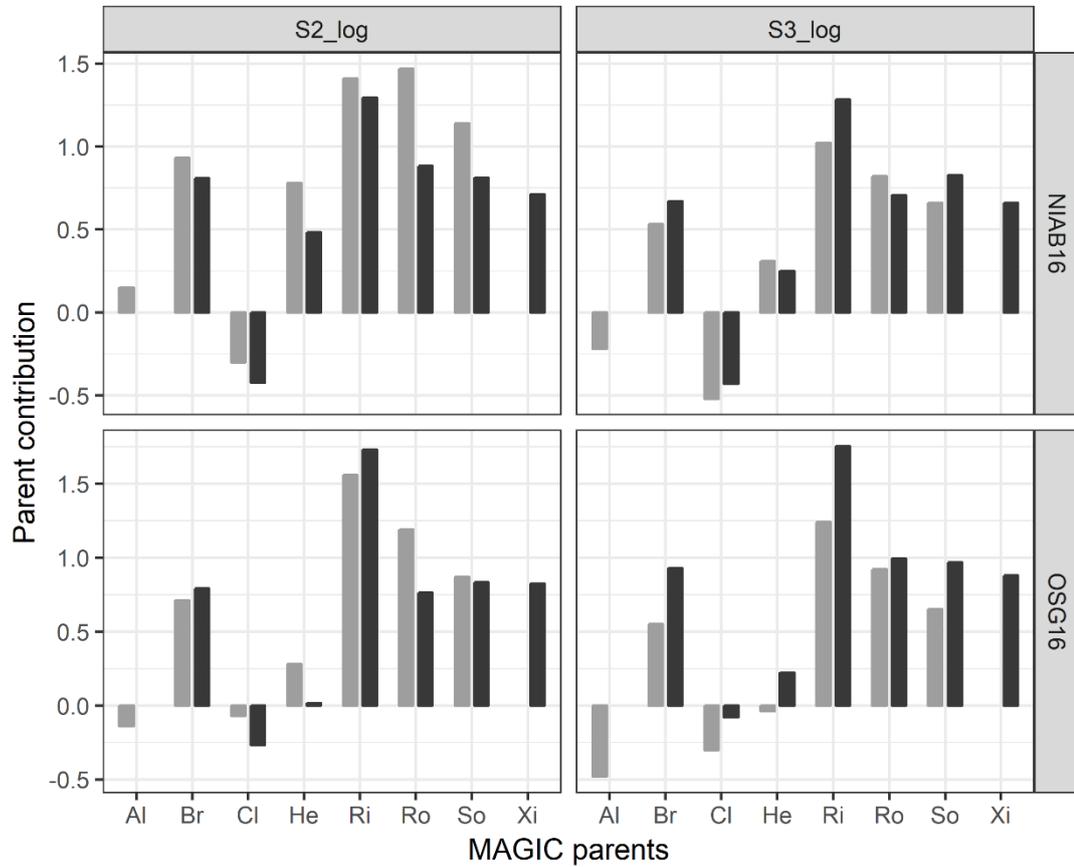Grey bars: Founder contributions based on Interval Mapping (0 covariates) relative to Xi19. Black bars: Founder contributions based on Haplotype Analysis, relative to Alchemy. Al: Alchemy, Br: Brompton, Cl: Claire, He: Hereward, Ri: Rialto, Ro: Robigus, So: Soissons, Xi: Xi19.

8.5 QTL-QTL interactions

Table 8.11 Pairwise comparison with t tests

	1a	1a_2a	1a_2a_2b	1a_2a_2b_2d	1a_2a_2d	1a_2b	1a_2b_2d	1a_2d	2a	2a_2b	2a_2b_2d	2a_2d	2b	2b_2d	2d
1a_2a	0.00388	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1a_2a_2b	2.00E-08	0.00106	-	-	-	-	-	-	-	-	-	-	-	-	-
1a_2a_2b_2d	1.30E-05	0.03589	0.40869	-	-	-	-	-	-	-	-	-	-	-	-
1a_2a_2d	3.20E-10	3.40E-05	0.30434	0.08882	-	-	-	-	-	-	-	-	-	-	-
1a_2b	6.70E-07	4.70E-12	< 2e-16	9.00E-14	< 2e-16	-	-	-	-	-	-	-	-	-	-
1a_2b_2d	3.50E-12	5.00E-07	0.0336	0.0086	0.25974	< 2E-16	-	-	-	-	-	-	-	-	-
1a_2d	1.90E-10	2.10E-05	0.24809	0.07032	0.89683	< 2E-16	0.31775	-	-	-	-	-	-	-	-
2a	0.00605	2.00E-07	2.50E-13	9.50E-10	4.40E-15	0.00913	< 2E-16	2.7E-15	-	-	-	-	-	-	-
2a_2b	1.00E-08	0.00062	0.86573	0.33144	0.38985	< 2E-16	0.04951	0.32304	1.3E-13	-	-	-	-	-	-
2a_2b_2d	5.00E-07	0.01138	0.41054	0.90936	0.06702	< 2E-16	0.00387	0.0505	6.5E-12	0.322	-	-	-	-	-
2a_2d	0.00189	0.80694	0.00222	0.05789	7.90E-05	1.8E-12	1.3E-06	0.00005	7.9E-08	0.00133	0.02127	-	-	-	-
2b	1.30E-05	0.08559	0.09601	0.52975	0.00832	5E-15	0.00027	0.00583	2E-10	0.06763	0.39217	0.13829	-	-	-
2b_2d	2.20E-08	0.00115	0.97782	0.42237	0.2916	< 2e-16	0.03146	0.23707	2.7E-13	0.84392	0.42641	0.00241	0.1015	-	-
2d	0.03703	0.38811	6.00E-05	0.00519	1.40E-06	1.5E-10	1.7E-08	8.5E-07	5.5E-06	0.000033	0.00093	0.26933	0.01116	0.000066	-
no_QTL	1.20E-10	9.00E-16	< 2E-16	< 2E-16	< 2E-16	0.03171	< 2E-16	< 2E-16	7.5E-06	< 2E-16	< 2E-16	3.6E-16	< 2E-16	< 2E-16	2.5E-14

P value outputs from pairwise comparisons of QTL combinations with t tests. Columns and rows running either side of the table represent the different QTL combinations e.g. 1_a_2a is *QYr.niab-1A.1* with *QYr.niab-2A.1*

APPENDIX D: SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Table 8.12 Summary of tBLASTn analysis output for candidate genes identified within QTL intervals

Candidate gene	Query gene	Genomic Location	Orientation	Query start	Query end	Query orientation	Length	Score	E-value	% ID
TraesCS2D02G573600	*Pm3b*	2D:638648716-638648955	Forward	600	679	Forward	80	125	3.40E-34	38.8
TraesCS2D02G573800	*Pm8*	2D:638778772-638778900	Forward	406	448	Forward	43	104	8.00E-23	44.2
TraesCS2D02G573900	*Sr45*	2D:638781450-638781644	Forward	1137	1201	Forward	65	82	1.00E-39	30.8
TraesCS2D02G574300	*Sr35*	2D:638863352-638863693	Reverse	410	523	Forward	114	401	1.20E-145	66.7
TraesCS2D02G574400	*Sr50*	2D:639092244-639092510	Reverse	1	89	Forward	89	261	9.60E-64	59.6
TraesCS2D02G574500	*Sr35*	2D:639147255-639147731	Reverse	120	278	Forward	159	307	2.80E-130	41.5
TraesCS2D02G574600	*Sr22*	2D:639165554-639165874	Reverse	278	384	Forward	107	262	1.30E-71	46.7
TraesCS2D02G574700	*Sr35*	2D:639277797-639278138	Reverse	410	523	Forward	114	357	2.20E-115	62.3
TraesCS2D02G574800	*Sr50*	2D:639295255-639295401	Reverse	205	253	Forward	49	116	1.60E-12	49
TraesCS2D02G575000	*Sr33*	2D:639363548-639364072	Reverse	1	175	Forward	175	358	0	42.9
TraesCS2D02G576100	*Sr22*	2D:639689249-639689983	Reverse	278	522	Forward	245	641	5.10E-114	48.6
TraesCS2D02G577300	*Sr35*	2D:640619420-640619569	Reverse	264	313	Forward	50	88	4.40E-63	32
TraesCS3A02G007400	*Sr22*	3A:7497925-7498215	Reverse	389	485	Forward	97	205	3.70E-61	42.3
TraesCS3A02G008000	*Sr22*	3A:7928204-7928401	Reverse	184	249	Forward	66	118	3.80E-24	39.4
TraesCS3A02G008100	*Sr22*	3A:7949861-7950181	Forward	414	520	Forward	107	241	5.20E-34	43.9
TraesCS3A02G009200	*Sr22*	3A:8325780-8326070	Reverse	429	525	Forward	97	173	7.90E-48	34
TraesCS3A02G009300	*Sr35*	3A:8550656-8550817	Forward	190	243	Forward	54	164	1.30E-64	55.6
TraesCS2A02G560600	*Lr1*	2A:762760389-762760706	Forward	393	498	Forward	106	258	5.20E-69	44.3
TraesCS2A02G560700	*Lr1*	2A:762771969-762772277	Forward	393	495	Forward	103	255	3.00E-60	40.8

Candidate gene	Query gene	Genomic Location	Orientation	Query start	Query end	Query orientation	Length	Score	E-value	% ID
TraesCS2A02G560900	*Lr1*	2A:762802610-762802912	Forward	395	495	Forward	101	269	3.80E-69	46.5
TraesCS2A02G561100	*Lr1*	2A:762985271-762985582	Forward	393	496	Forward	104	274	5.80E-52	47.1
TraesCS2A02G564200	*Sr22*	2A:764669485-764669811	Reverse	278	386	Forward	109	284	1.6E-66	49.5
TraesCS2A02G564300	*Lr10*	2A:764678386-764678625	Forward	562	641	Forward	80	127	3.9E-41	35
TraesCS2A02G564500	*Sr35*	2A:764724594-764724845	Reverse	583	666	Forward	84	237	8.4E-97	59.5
TraesCS2A02G564600	*Sr50*	2A:764728233-764728499	Reverse	1	89	Forward	89	270	1.2E-101	60.7
TraesCS2A02G564800	*Sr35*	2A:764786877-764787119	Reverse	306	386	Forward	81	175	5E-39	40.7
TraesCS2A02G564900	*Lr10*	2A:764820220-764820474	Forward	557	641	Forward	85	117	4.8E-32	30.6
TraesCS2A02G565800	*Sr22*	2A:765173651-765174385	Reverse	278	522	Forward	245	633	5.1E-115	47.8
TraesCS2B02G486100	*Lr1*	2B:683043889-683044182	Forward	294	391	Forward	98	214	8.00E-114	38.8
TraesCS2B02G486200	*Lr1*	2B:683055005-683055286	Forward	294	387	Forward	94	197	3.70E-66	37.2
TraesCS2B02G486300	*Lr1*	2B:683068599-683068892	Forward	294	391	Forward	98	217	2.80E-118	39.8
TraesCS2B02G486400	*Lr1*	2B:683129116-683129397	Forward	294	387	Forward	94	202	3.00E-77	38.3
TraesCS2B02G486700	*Lr1*	2B:683160473-683160766	Forward	294	391	Forward	98	203	6.40E-93	37.8
TraesCS2B02G487700	*Lr1*	2B:683753122-683753433	Forward	399	502	Forward	104	228	2.50E-43	39.4
TraesCS2B02G488000	*Lr1*	2B:685267626-685267925	Forward	396	495	Forward	100	249	3.70E-66	45
TraesCS2B02G488400	*Lr1*	2B:685743425-685743724	Forward	396	495	Forward	100	231	9.00E-60	40
TraesCS2B02G488700	*Lr1*	2B:686050526-686050804	Forward	1252	1344	Forward	93	181	9.00E-60	45.2
TraesCS2B02G489400	*Lr1*	2B:686812193-686812513	Forward	396	502	Forward	107	245	4.10E-64	39.3
TraesCS6A02G051800	*Pm3b*	6A:26952663-26952806	Forward	174	221	Forward	48	96	1.40E-13	35.4
TraesCS6A02G051900	*Sr22*	6A:26963026-26963196	Forward	408	464	Forward	57	134	1.30E-17	43.9

Candidate gene	Query gene	Genomic Location	Orientation	Query start	Query end	Query orientation	Length	Score	E-value	% ID
TraesCS6A02G052200	*Pm3b*	6A:27012034-27012267	Reverse	174	251	Forward	78	133	2.30E-15	29.5
TraesCS2A02G362300	*Lr1*	2A:606678426-606678791	Reverse	395	516	Forward	122	250	2.50E-54	41
TraesCS6A02G052400	*Pm3b*	6A:27087860-27088102	Reverse	174	254	Forward	81	149	2.20E-18	34.6
TraesCS6A02G052500	*Sr45*	6A:27108032-27108325	Reverse	149	246	Forward	98	136	3.70E-18	28.6

tBLASTn results based on searches using translated CDS of *Lr, Pm, Sr, Yr* cloned resistance genes as queries, referred to as 'Query gene' here. For each candidate gene, the result listed is that of the alignment with the highest score, using an e-value cut-off E-10.

APPENDIX E: SUPPLEMENTARY MATERIAL FOR CHAPTER 5

Table 8.13 Summary of REML estimates for glume YR resistance in the NIAB16 and OSG16 MAGIC pathology trials

Site	REML estimates						σ^2_e	H^2
	Random			Fixed				
	Term	Est.	SE	Term	Wald	*Chi pr*		
NIAB16	column	0.00137	0.0013	G	4227.1	<0.001	0.0603	0.80
OSG16	rep.block	0.00009	0.00043	G	4773.1	<0.001	0.0568	0.82

Covariance estimates of the mixed linear model based on Restricted Maximum Likelihood (REML). Est.: Estimates; SE: Standard error; G: genotype; H_2: broad sense heritability; σ^2_e residual error.

Table 8.14 Outputs of QTL mapping analysis, per peak marker for each identified QTL

QTL	Chr	Peak marker	Position, cM	Single Marker Analysis		Haplotype Analysis		Interval Mapping		Composite Interval Mapping	
				NIAB16	OSG16	NIAB16	OSG16	NIAB16	OSG16	NIAB16	OSG16
QYrg.niab-2D.Ppd1_D1	2D	Ppd_D1	60.67	.	.	.	.	6.44E-05	0.001979	4.61E-07	1.57E-05
QYrg.niab-2D.1	2D	Ra_c21099_1781	197.36	.	.	.	.	0.001684	0.00034	1.48E-05	6.7E-06
QYrg.niab-4A.1	4A	BS00067903_51	180.38	.	.	.	.	0.00038	.	3.73E-05	.
QYrg.niab-4D.1	4D	D_GDRF1KQ02H66WD_341	125.78	.	.	.	.	4.09E-06	3.28E-06	6.32E-08	1.32E-05
QYrg.niab-5A.1	5A	BS00065714_51	309.03	6.22E-05	.	6.22E-05	1.95E-05	5.7E-06	1.7E-08	9.12E-09	3.98E-09

Genetic map position and QTL interval are from the MAGIC genetic map (Gardner et al. 2016). *P* values shaded in grey correspond to values uncorrected for False Discovery Rates. Dots indicate the absence of a significant *p* value for the corresponding peak marker at that particular QTL.