Host range expansion of wheat stem rust resistance genes into barley

Muhammad Asyraf Bin Md Hatta

Thesis submitted to the University of East Anglia

for the degree of Doctor of Philosophy

2017

Abstract

Puccinia graminis f. sp. *tritici* (*Pgt*) is the causal agent of the devastating stem rust disease in wheat. In recent years, new super virulent races of the fungus have emerged causing large scale epidemics. In an attempt to clone the stem rust resistance gene *Sr44*, we generated and screened an ethyl methane sulphonate (EMS) mutant population of an *Sr44* wheat-alien introgression line. We identified twelve independent susceptible mutants from 1171 M_2 families and sequenced the nucleotide binding leucine-rich repeat (NLR) genes in ten of the mutants and the wild-type. However, sequence comparison did not reveal a clear candidate. To investigate meristem cell fate in wheat, we phenotyped the sister spikes of ten *Sr44* M_2 families which segregated for susceptibility in the main spike. Ninety-two percent of the tested spikes were found to be resistant suggesting that they are genetically distinct from the main tiller. To improve the immunity of barley against wheat stem rust, we transformed the previously cloned wheat *Sr22*, *Sr33*, and *Sr45* genes into barley. The resultant transgenic lines expressed high-level resistance to *Pgt* indicating wheat *Sr* genes can be transferred into barley. Nucleotide sequence analysis of the *Sr22* locus revealed that some alleles have undergone historical sequence exchange in the LRR region. We also generated and phenotyped wheat transgenics to confirm the gene postulation of two previously identified *Sr22* alleles. Stacking of multiple *Sr* genes at a single transgene locus is expected to result in more durable resistance. We have attempted to use CRISPR/Cas9 to repair the hygromycin phosphotransferase II (*HPTII*) gene as a proof-of-concept to *in vivo* sequential stacking of multiple *Sr* genes. Super transformation of barley T_0 and T_1 transgenics containing a landing pad did not yield positive transformants. However, we identified one deletion event out of twenty-four calli of T_0 transgenics, indicating functional CRISPR/Cas9 activity.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, all praise is due to Allah, The Almighty for giving me the opportunity, determination and capability to complete my PhD journey.

I would like to express my greatest gratitude and thanks to Dr. Brande Wulff for his excellent supervision, endless guidance, motivation, encouragements, and suggestions throughout the entire research period. Thank you for your time and patience in reading my thesis, correcting grammar and sentences until it became a complete PhD thesis. I am extremely grateful to have you as my supervisor, a humble and an understanding person. You will be an example for the rest of my scientific career. It had been a great pleasure as well to play football alongside with you and team up for several occasions in which we became the champions for the Fit4Work tournament and a runner-up for the NBI annual football tournament. These moments will always be in my memories.

I would like to extend my gratitude towards my secondary supervisor, Professor Wendy Harwood for giving me the opportunity to carry out barley transformation in her laboratory and providing me with critical inputs to my project. Many thanks for your support and guidance.

Also, I would like to thank my supervisory team member Dr. Nicola Patron for guiding me in structuring the CRISPR/Cas9-gene addition experiment and teaching me the Golden Gate cloning. Special thanks also to Mark Smedley, the go-to person when it came to cloning problems. I am thankful for your valuable suggestions and guidance in regards to construct generation, PCR, and bacterial transformation. Many thanks to both of you for the fruitful discussions and suggestions during the meeting.

I would like to acknowledge members of the group, particularly Dr. Guotai Yu for his time and assistance throughout my PhD, Dr. Burkhard Steuernagel for RenSeq analysis, GBS analysis, and other bioinformatics support, Sreya Ghosh for helping me with the crosses and for giving me the opportunity to contribute to the Speed Breeding project, and Dr. Sanu Arora for fruitful discussions and guidance in *Sr22* alleles sequence analysis. My sincere thanks also to Yajuan Yue and her team for helping me in managing the wheat plants from sowing, transplanting, harvesting, threshing to counting the seed. To my other group members, Clemence Marchal, Ngoni Kangara, Agathe Jouet and also former members Dr. Lucy McCann, Deepika Narang and Wenjuan He, thank you for your support. It had been a great pleasure working with all of you.

I would like also to acknowledge our collaborator in the University of Minnesota, Professor Brian Steffenson and his team for their precious help in phenotyping the wheat and barley materials. Special thanks to Ryan Johnson and Oadi Matny and also former members Jamie Kaufman and Cole Welchlin for the great job.

Many thanks also to our collaborators in the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, Dr. Evans Lagudah, Dr. Michael Ayliffe, Dr. Sambasivam Periyannan and Dr. Rohit Mago for their contributions in wheat transformation and *Sr* barley transgenic project.

Not to forget, many thanks to Dr. Ian Dundas from The University of Adelaide, Australia for providing me with the *Sr44* introgression line materials.

I would like to express my sincere acknowledgement to members of the BRACT group especially Dr. Sadiye Hayta, Alison Hinchliffe and Maria Amador for their support in barley and wheat transformation, Tom Lawrenson for suggestions and advices in cloning and CRISPR/Cas9, and Monika Chhetry for her helps in harvesting barley and wheat materials. Thank you also to other group members, Dr. Penny Hundleby, Nicola Atkinson, and Jemma Taylor for their support in the lab.

Not to forget, my former 2Blades group in The Sainsbury Laboratory (TSL) especially Dr. Inmaculada Hernandez-Pinzon for her support in cloning and PCR and also other group members Dr. Matthew Moscou, Dr. Peter Van Esse, Dr. Nadine Ilk, Dr. Cintia Kawashima, Dr. Jan Bettgenhaeuser, Dr. Phon Green, Dr. Andrew Dawson, and Sue Banfield for helping me going through my first year. I would like also to acknowledge Mark Youles from the Synthetic Biology TSL for his guidance and support in Golden Gate cloning.

I would like to thank Dr. Oleg Raitskin from Nicola Patron's laboratory, Earlham Institute for his support and suggestions in setting up barley protoplast assays.

My sincere acknowledgement also goes to all staff in horticultural services for their kind support in looking after my plants and all technicians in the media kitchen for making the whole experiment easier. Not to forget, the lab manager Dr. Roy Dunford who was always there to help me.

I have a kind thought about my Malaysian company at JIC, Dr. FeiYian Yoong and Dr. Nadiatul Radzman for their laugh and support.

I am thankful to the Ministry of Higher Education Malaysia and Universiti Putra Malaysia for the sponsorship of my PhD degree. Special thanks to the Dean of Faculty of Agriculture, Professor Dr. Abdul Shukor Juraimi, Deputy Dean Associate Professor Dr. Halimi Mohd Saud and Head of Department Associate Professor Dr. Norsida Man for the support and opportunity to pursue my study abroad. Not to forget, Puan Zauyah Yusof and Puan Zuraida Safni Dzulkarnain for administrative support.

I would like also to thank all Malaysian families in Norwich for their moral support, helps, foods and potlucks in creating a little Malaysia in Norwich.

My deepest love and sincere gratitude goes to my family especially my father, Bapak and my mother, mama for their support and countless blessings and continuous prayer. My sincere thanks also to my father-in-law, Ayah and my mother-in-law, Mak who have been very understanding and thoughtful despite being far away from their daughter and grandchildren. Thank you for your pray and blessings.

I dedicate this PhD manuscript to my beloved family - my wife, Khalilah Ghazali and our children Afnan Wafi and Nur Amni. My sincere gratitude to my wife for her endless love, support and patience throughout this journey. Thank you dear, for helping make my dreams come true.

1 General introduction

1.1 Plant disease resistance

1.1.1 Plant immune receptors

Biotrophic or hemi-biotrophic plant pathogens including fungi, oomycete, bacteria and viruses invade host plant cells and cause disease. In an agricultural setting where a single cultivar of a given crop is often grown on a large area, this can lead to devastating epidemics and complete crop failure. A well-known epidemic was the late blight Irish potato famine caused by the oomycete *Phytophthora infestans* in the 1840s (Yoshida et al., 2013). More recently, the emergence of new virulent races of wheat stem rust in Africa caused by *Puccinia graminis* f. sp. *tritici* (Singh et al., 2015) and wheat blast in Bangladesh caused by *Magnaporthe oryzae* have destroyed few of the local cultivated wheat within a short period of time (Islam et al., 2016; Sadat and Choi, 2017). These reports demonstrate a continuing threat of plant pathogens to global food security.

With the long term aim of achieving durable resistance against crop pathogens, molecular plant pathologist have subjected the structural and functional features of the plant immune system to intense scrutiny over the past 25 years. Unlike mammals that have evolved an adaptive and circulating immune system, plants rely on the innate immunity of each individual cell to respond to specific pathogens. The plant immune system involves extracellular and intracellular receptors which represent the two layers of defence **(Figure 1.1)**. In the extracellular space, pattern recognition receptors (PRR) (receptor kinases and receptor-like proteins) perceive evolutionarily conserved molecules among pathogens known as pathogen-associated molecular patterns (PAMPs) or microbe-associated molecular patterns (MAMPs) (Dangl et al., 2013). The activation of the PRR typically leads to intracellular immune signalling and elicitation of a complex output response including reactive oxygen species (ROS) and the release or synthesis of antimicrobial compounds to halt pathogen proliferation, designated as PAMP-triggered immunity (PTI) (Jones and Dangl, 2006) **(Figure 1.1)**. Since the PAMPs are highly conserved among pathogens, activation of PTI is generally effective to ward off non-adapted pathogens, thus acting as an important front line defence upon infection.

Adapted biotroph or hemi-biotroph pathogens, in turn, subvert the PAMP/MAMP perception by deploying a set of virulence factors known as effectors to manipulate the host's cell physiology and development for nutrient acquisition and survival. The mechanism of effector delivery into host cells varies among plant pathogens, and includes the Type III

secretion system (TTSS) by extracellular bacteria (Baltrus et al., 2011), specialised feeding structures such as haustoria by oomycetes and fungi (O'Connell and Panstruga, 2006), or the stylet by aphids and nematodes (Bos et al., 2010) **(Figure 1.1)**. In the intracellular space, the second class of immune receptors encoded by resistance (*R*) genes recognise these effectors and initiate the second layer of defence known as effector-triggered immunity (ETI) (Jones and Dangl, 2006). The ETI is typically associated with a localised plant cell death, termed the hypersensitive response (HR), which blocks pathogen growth **(Figure 1.1)**. Most *R* genes encode modular proteins containing a nucleotide-binding (NB) domain and a leucine-rich repeat (LRR) domain (known as NLRs) in which the NB domain is crucial for regulation of *R* gene activity and the LRR domain is responsible for recognition specificity (Krasileva et al., 2010; Ravensdale et al., 2012). NLRs can be subdivided into two main groups based on their different N-terminal domains; those that possess a coiled-coil (CC) domain are referred to as CNLs, while those that possess a Toll-interleukin 1 receptor (TIR) domain are referred to as TNLs (Jupe et al., 2013). CNLs can be found in both dicot and monocot plants while TNLs are largely absent from monocot genomes (Cui et al., 2015; Sarris et al., 2016).

In triggering resistance immune signalling, the mechanism underlying recognition specificity of NLR-effector interaction can be described either via direct binding as receptor and ligand or indirect binding through sensing of effector modifications on host components (Cui et al., 2015) **(Figure 1.1)**. In indirect recognition, the 'guard model' states that the NLRs are activated upon effector-triggered modification of the host cellular targets or guardees associated with the NLRs (Dodds and Rathjen, 2010). Another indirect recognition mode proposes that the guarded host components have no significant resistance function, instead they serve as structural mimics (or decoys) to trap effectors and trigger NLR activation upon effector-mediated modification (van der Hoorn and Kamoun, 2008). In addition, recent studies suggest that NLRs can contain additional domains besides NB and LRR domains. These additional domains may be fused to NLRs to form 'integrated decoys' or 'integrated sensors' as effector targets, which trigger activation of the NLR upon effector-mediated modifications (Cesari et al., 2014a; Kroj et al., 2016; Sarris et al., 2016).

Based on the gene-for-gene model, the recognition of effectors (or avirulence gene products) by an *R* protein is highly specific (Cui et al., 2015). Over evolutionary time, the more adapted pathogens may have accumulated sequence changes in a given effector which through adaptive selection has allowed them to avert detection by the host. In turn, new *R* gene alleles can appear through mutation or recombination that can recognise the newly evolved effectors. This repeated cycle of attack and defence over time has led to a competitive host-

pathogen coevolution race. Two popular models of co-evolutionary dynamics between host and pathogen at the population level have been proposed; the arms race model and the trench warfare model. The major distinction between these two models is the expected co-evolutionary pattern of *R* genes and effector. The arms race model proposes that although diversifying selection may act on both *R* gene and effector in a continuous cycle, the genetic variation at the population level is reduced and becomes temporarily fixed through recurrent selective sweeps (Dawkins and Krebs, 1979). By contrast, in the trench warfare model balancing variation within both *R* gene and effector are maintained in the host and pathogen populations (Stahl et al., 1999).

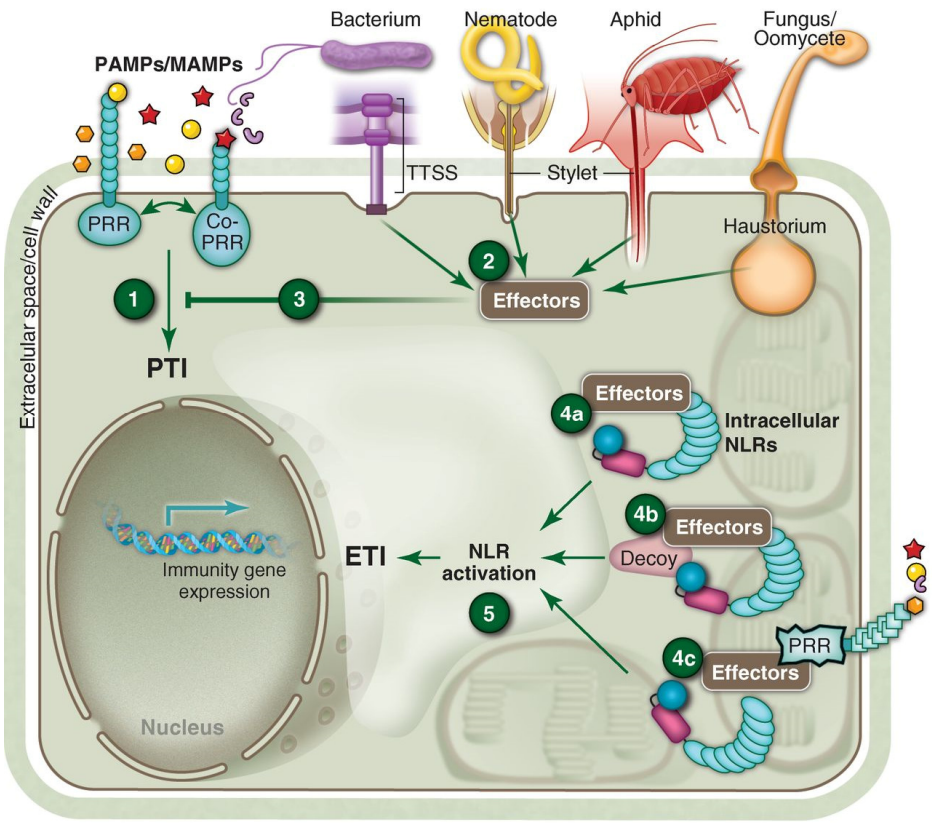


Figure 1.1 The principles of the plant immune system.

Pathogens (colour coded and labelled) express PAMPs and MAMPs upon host colonisation which are recognised by plant pattern recognition receptors (PRRs) that activate PAMP-triggered immunity (PTI; step 1). Pathogens secrete effectors to interfere with the PTI response (steps 2 and 3). Plants can detect the presence of these effectors by intracellular NB-LRR (nucleotide-binding leucine-rich repeat) proteins through their direct effector recognition (step 4a), indirect recognition of effector-triggered modification of a structural mimic (or decoy) (step 4b), or indirect recognition of effector-triggered modification of a host cellular target (or guardee) (step 4c), leading to NLR activation (step 5) and effector-triggered immunity (ETI) (Dangl et al., 2013).

1.1.2 Pathogen effectors

Effectors are considered as key factors in the establishment of a compatible interaction between host and pathogen. Genome sequencing of many plant pathogens has revealed large numbers of genes that encode putative effectors – i.e. secreted, small (<300 amino acids), and cysteine-rich proteins (Stergiopoulos et al., 2013). The virulence function and underlying molecular mechanism of some effectors has been revealed such as the *Ustilago maydis* effector Tin2 that targets anthocyanin biosynthesis in maize to promote virulence (Tanaka et al., 2014) and the *Cladosporium fulvum* effector Tom1 that degrades the antifungal glycoalkaloid α-tomatine in tomato for full virulence (Okmen et al., 2013). To date, however, the mechanism of most effectors in disease establishment remains obscure.

Effectors can reside in the apoplast such as those secreted by the tomato pathogen *Cladosporium fulvum* (Joosten and de Wit, 1999; Rooney et al., 2005; Shabab et al., 2008) and Pep1 secreted by *Ustilago maydis* (Doehlemann et al., 2009), or they can be translocated into plant cells such as those effectors injected by the type III secretion system of bacteria (Tseng et al., 2009). Inside the cell, they may target various compartments like the chloroplast or the nucleus (Khang et al., 2010). Bacterial and oomycete effectors are the most functionally characterised effectors. In contrast, relatively few fungal effectors have been characterised, most likely because most of the devastating fungal pathogens are obligate biotrophs, which complicates their genetic manipulation. The mechanism of effector translocation has been best characterised in oomycetes such as *Phytophthora infestans* that produce haustoria during infection. A short conserved amino acid sequence was identified near the signal peptide of certain oomycetes secreted effectors. This so-called RXLR motif resembles a motif (RXLXE/D/Q) found in virulence proteins of *Plasmodium* (the malaria parasite), which are required for effector translocation into erythrocytes (Hiller et al., 2004; Marti et al., 2004). However, RXLR and RXLR-like motifs are not commonly found in fungi and how fungal effectors are translocated remains unclear.

Core effectors are a set of effectors that are widely distributed across the population of a particular pathogen and broadly contribute to pathogen virulence. Identification of core effectors using computational biology followed by molecular characterisation may facilitate cloning of the corresponding *R* genes that they activate in diverse wild germplasm. Genome sequencing of several *P. infestans* species, the causal agent of potato late blight has revealed a set of core effectors which can now facilitate cloning of novel *R* genes in potato (Vleeshouwers et al., 2011). Stacking and deployment of multiple *R* genes that recognise defined core effectors could provide a more durable resistance.

1.1.3 Engineering disease resistance in crops

Chemical applications and various agronomic practices including crop rotation, planting density, eradication of the alternate pathogen host, disease-free seeds and clean equipment have been used by farmers to control crop diseases in addition to deployment of resistant varieties. Despite all these practices, estimates suggest that approximately 15% of global crop production is lost due to crop diseases (Popp and Hantos, 2011).

Multiple strategies could be pursued to engineer resistance to pathogens in crops. One strategy involves the deployment of PRRs that recognise conserved microbial molecules into species in which those receptors are absent. The Arabidopsis Ef-Tu receptor (EFR) has been transferred into several species that are unable to perceive the bacterial PAMP elongation factor Ef-Tu including *Nicotiana benthamiana* and *Solanum lycopersicum* (tomato) resulting in enhanced resistance against a wide range of bacterial pathogens (Lacombe et al., 2010). A recent study demonstrated that Arabidopsis EFR can be transferred successfully to distant taxa such as *Triticum aestivum* (wheat) to confer resistance to the cereal bacterial pathogen *Pseudomonas syringe* pv. *oryzae* (Schoonbeek et al., 2015a). The PRR from tomato, *Verticilium 1* (*Ve1*) can also be transferred to Arabidopsis to confer resistance to *Verticilium* race 1 strain (Fradin et al., 2011).

The second strategy aims at deployment of *R* genes, a class of immune receptor that recognises pathogen effectors. The transfer of an *R* gene from one species to another is normally achieved by introgression breeding between a crop and its wild or closely related relative. This traditional breeding is often a lengthy procedure which involves multiple rounds of crossing and backcrossing to reduce the size of the introgressed segment (Olivera et al., 2007). Despite this, the introgressed segment often contains undesirable alleles, a problem termed linkage drag. Historically, breeders have tended to introduce a new variety that carries only one or two new *R* genes. The deployment of a single *R* gene over large areas promotes the natural selection of pathogen strains with increased virulence, which often leads to breakdown of resistance within a few growing seasons – the so-called “boom and bust” cycle (Roelfs et al., 1992). In this cycle, breeders need to continually select and develop new resistant varieties in order to keep pace with pathogen evolution. For example, in 1970, the resistance of the bread wheat cultivar Yecora 70 with race-specific *R* genes against leaf rust was defeated after only three years of deployment in North-western Mexico, requiring a cultivar replacement (Singh, 2012). A more recent example concerns the resistance breakdown of the bread wheat cultivar Digalu, carrying *SrTmp*, to a new stem rust race, TKTTF (named as the Digalu isolate) in Southern Ethiopia during the 2013-2014 cropping

season (Singh et al., 2015). The host jump of the rice blast pathogen, *Magnaporthe oryzae*, from rice to wheat in Brazil in the mid-1980s has also been attributed to the defeat of major *R* genes in wheat (Inoue et al., 2017). From first principle, the deployment of multiple *R* genes in a single cultivar should delay the breakdown of resistance by the pathogen. This is because in the face of multiple *R* genes, there would be no selection for pathogen variants with loss of a single effector. A selective advantage would require the simultaneous loss of all corresponding effectors in the pathogen which would be expected to be a very rare event (REX-Consortium, 2016).

An alternative strategy exploits another class of plant genes, known as executor *R* genes that have evolved in response to pathogen reprogramming of host gene expression patterns during infection by transcription activator-like effectors (TALEs). TALEs such as from *Xanthomonas* and *Ralstonia* are injected into plant cells through the Type III secretion system and directly bind to DNA sequence elements in host promoters to upregulate host genes and promote disease. In turn, some plants have evolved TALE-mediated resistance by triggering induced HR upon TALEs binding to hosts promoters, thus restricting pathogen growth. Examples of executor *R* genes include *Bs2* and *Bs4c* from pepper (Römer et al., 2007; Strauss et al., 2012) and *Xa27*, *Xa10* and *Xa23* from rice (Gu et al., 2005; Tian et al., 2014; Wang et al., 2015). The expression of each executor *R* gene is induced by a specific TALE. This provides a useful strategy for durable resistance by designing executor *R* genes with engineered promoters that recognise multiple TALEs that are all expressed in the pathogen population. The pepper *Bs3* promoter was engineered to confer specific induction to three distinct TALEs from *Xanthomonas campestris* pv. *vesicatori*, the causal agent of bacterial leaf spot (Romer et al., 2009). Another example is the rice *Xa27* gene in which the promoter was engineered to contain three TALE binding sites from *Xanthomonas oryzae* pv. *oryzae* and three from *Xanthomonas oryzae* pv. *oryzicola* to confer resistance against both bacterial blight and bacterial leaf streak, respectively (Hummel et al., 2012). Likewise, rice transgenics carrying the modified $Xa10^{E5}$ gene with binding sites to five TALEs were resistant to 27 of the 28 selected *Xanthomonas oryzae* pv. *oryzae* collected from 11 countries (Zeng et al., 2015).

The host genes that contribute to pathogen colonisation (i.e. through the TALE mechanism) can be considered as disease susceptibility (*S*) genes. Impairment of *S* gene function by mutation could limit the ability of the pathogen to cause disease and lead to pathogen-specific resistance, thus providing a useful strategy for potential durable resistance. In contrast to *R* genes in which the resistances are typically dominant especially in polyploid wheat, most of the *S* genes are recessive. A notable example of an *S* gene's long durability in

the field is the *Mlo* gene, in which a recessive mutant has conferred powdery mildew resistance in barley for the past 70 years (Jørgensen, 1992). With the advent of genome-editing tools such as transcription activator-like effector nuclease (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR/Cas9), disabling of *S* gene function could be relatively straightforward. Examples include the TALEN-mediated simultaneous mutation of all three *MILDEW-RESISTANCE LOCUS* (*Mlo*) homoeoalles in wheat which led to broad-spectrum resistance to powdery mildew (Wang et al., 2014b), and the TALEN-induced mutation of the rice bacterial blight susceptibility gene *Os11N3* (also called *OsSWEET14*) which led to blight-resistant rice (Li et al., 2012). More recently, a transgene-free powdery mildew resistant tomato has been generated by knocking out the *SlMlo1* gene using CRISPR/Cas9 (Nekrasov et al., 2017), while CRISPR/Cas9-induced mutation of both alleles of *CsLOB1* in Duncan grapefruit conferred resistance to citrus canker (Jia et al., 2016). These reports demonstrate that the alteration of *S* gene function has opened up an alternative pathway to generate more durably resistant crop varieties.

1.2 The wheat *Puccinia graminis* f. sp. *tritici* pathosystem

Wheat rust diseases are amongst the most important diseases in agriculture worldwide. There are three rust pathogens of wheat, namely leaf or brown rust (*P. triticina*), stripe or yellow rust (*P. striiformis* f. sp. *tritici*) and stem or black rust (*P. graminis* f. sp. *tritici*). Stem rust can cause severe damage to the stems and other aerial parts of the wheat plant, leading to lodging and shrunken grain, and has been associated with dramatic, widespread crop failures throughout recorded history. This pathogen favours a warm temperature, up to 35 ˚C (Roelfs et al., 1992), and is able to develop and multiply rapidly in favourable conditions, changing a green wheat cultivar into a black tangle of broken stems within weeks of the appearance of first symptoms (Singh et al., 2008a). *Pgt* is heteroecious, requiring two unrelated host plants to complete its complex life cycle, and macrocyclic, with five different spore stages **(Figure 1.2)** (Singh et al., 2008a). In addition to wheat, barley and triticale are also the primary hosts of this pathogen. The common barberry, *Berberis vulgaris* is the most important alternate host of stem rust whereon it can reproduce sexually to give rise to new strains with novel virulence combinations **(Figure 1.2)**.

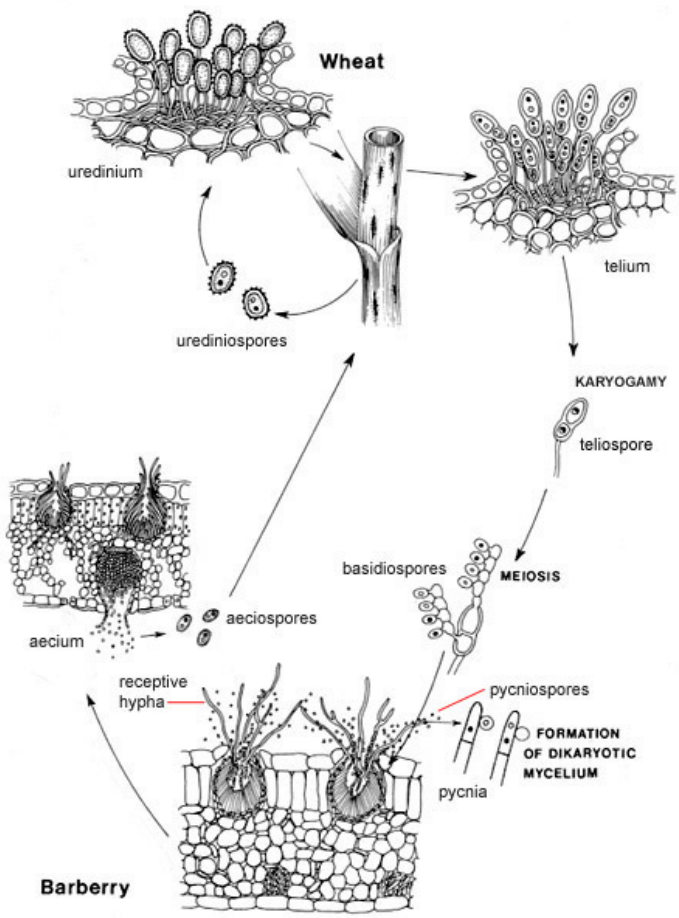


Figure 1.2 Life cycle of *Puccina graminis* f. sp. *tritici.*

Asexual cycle involves production of uridiniospores on the primary host wheat. Production of teliospores initiates sexual cycle, followed by production of basidiospores on wheat and pycniospores and aeciospores on the alternate host barberry, before re-infecting the primary host wheat (Schumann and Leonard, 2000).

1.3 The Ug99 and Digalu wheat stem rust epidemics

Despite its threatening potential, wheat stem rust has largely been kept under control for many decades due to the use of resistant cultivars developed by Edgar McFadden, Norman Borlaug and many others. McFadden was able to transfer resistance to stem rust (*Sr2*) from *Triticum turgidum* L. ssp. *dicoccum* (cv Yaroslav) into hexaploid wheat, producing the variety Hope (McFadden, 1930b). *Sr2* has provided durable, broad-spectrum rust resistance in wheat for more than 80 years. Through crossbreeding of various varieties of wheat, some 80 additional *Sr* genes have been identified (McIntosh et al., 2017). One very effective gene, *Sr31*, introduced into wheat from rye, also provided a 5% yield increase, thus in part contributing to its wide deployment (Sharma et al., 2013). *Sr31*-containing stem rust resistant cultivars have protected wheat production for several decades until the 1998/1999

growing season, when a new race of wheat stem rust called Ug99 was detected in Africa that was virulent on *Sr31* containing plants (Pretorius et al., 2010).

Ug99 (or TTKSK, according to the North American race nomenclature) was first detected in Uganda **(Figure 1.3)**. Since wheat stem rust produces wind-borne spores, this facilitated the rapid spread to neighboring countries including Kenya, Ethiopia and Sudan (Sharma et al., 2013). In 2006, it crossed the Red Sea and reached Yemen in the Arabian peninsula **(Figure 1.3).** A year later, it was detected in Iran, causing a major concern due to the likelihood of further migration to the Punjab, one of the most important wheat growing regions in Pakistan and India. This prediction is based on the previous long distance spread of a stripe rust race virulent on *Yr9*, from Africa to India (Singh et al., 2008a). To make matters worse, Ug99 has mutated and expanded its virulence resulting in two new derivatives of Ug99, TTKST and TTTSK, first identified in Kenya, that overcome *Sr24* and *Sr36*, respectively (Jin et al., 2009). In 2009, yet another new race in the Ug99 lineage was identified in South Africa with virulence on *Sr24* (Pretorius et al., 2010). To date, Ug99 and ten Ug99 variant races have been detected in 13 countries including most recently Egypt where virulence to *Sr31* was reported in the 2014 cropping season (Patpour et al., 2015) **(Figure 1.3)**. The rapid emergence of new virulent races of stem rust in Africa is most likely linked to the prevalence and functionality of the alternate host, barberry (Glen and Institute, 2002; Keet et al., 2016).

In Europe, although no Ug99 race group has been reported so far, stem rust outbreaks caused by other stem rust races were detected in some countries **(Figure 1.3)**. In the summer of 2013, the first German stem rust outbreak in the last five decades occurred (Olivera Firpo et al., 2017). Several outbreaks have also been reported in recent years in Russia (Shamanin et al., 2016). More recently, Sicily in Italy was severely affected with thousands of hectares of both bread and durum wheat being destroyed (Bhattacharya, 2017). This outbreak has been potrayed as the largest European stem rust outbreak in more than 50 years.

Another widely distributed stem rust race which is not part of the Ug99 race group is the Digalu race (race TKTTF), which caused a severe stem rust epidemic in southern Ethiopia at the end of 2013 **(Figure 1.3)**. This race destroyed almost 100% of the most widely grown wheat cultivar Digalu, which carries the *SrTmp* resistance gene (Olivera et al., 2015). *SrTmp* was deployed in that cultivar due to its effectiveness against the Ug99 race group and the good resistance of Digalu to the prevalent races of stripe rust (Singh et al., 2015).

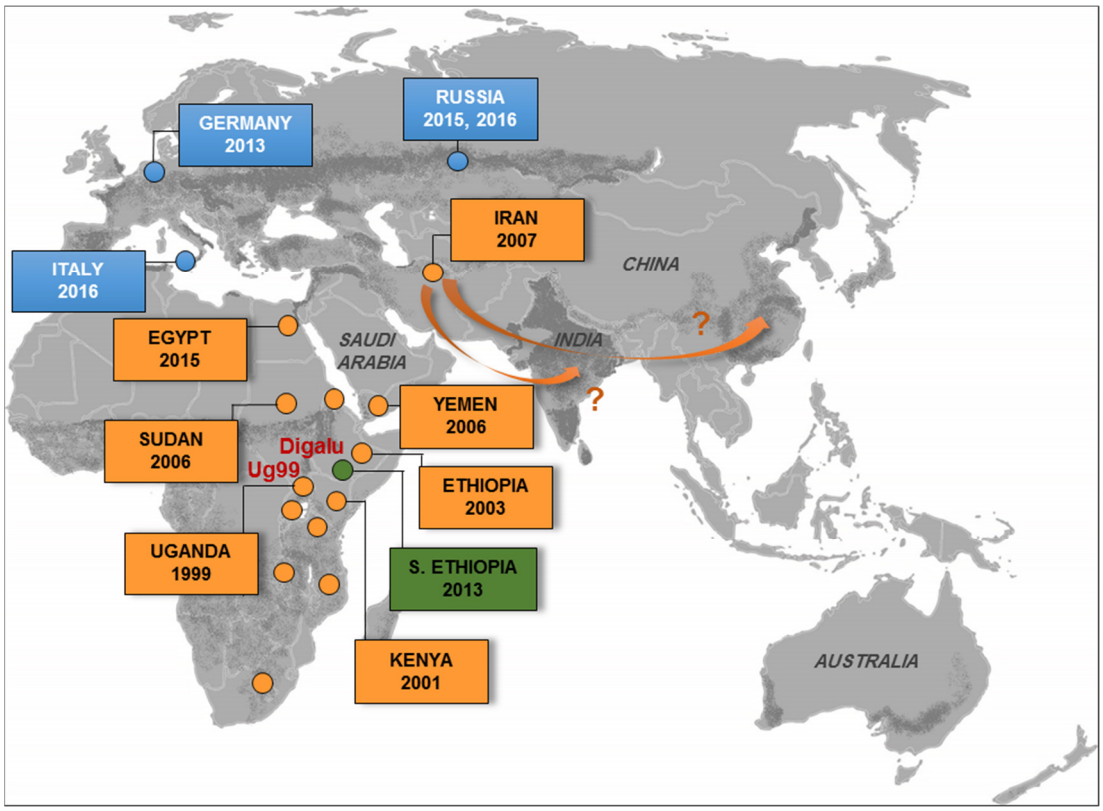


Figure 1.3 The spread of wheat stem rust Ug99 and wheat stem rust outbreaks in Europe.

In 1999, Ug99 (or TTKSK) was first identified in Uganda and a few years later it had spread to neighboring countries in Africa (Kenya, Ethiopia, Sudan, and Egypt), the Arabian peninsula (Yemen) and Asia (Iran). The rapid spread of Ug99 is a concern for the major wheat producing areas (dark grey in the map) such as Pakistan and India. In late 2013, Digalu (or TKTTF) caused a severe epidemic in southern Ethiopia. In the same year, a major stem rust outbreak was reported in Germany followed by a major outbreak in Italy three years later. Yellow shows the spread of Ug99. Blue shows the wheat stem rust outbreaks in Europe. Green shows emergence of a non-Ug99 race group, TKTTF.

1.4 A novel approach to disease resistance gene deployment

To reduce the risk of boom-and-bust cycles, multiple *R* genes could be stacked in a wheat cultivar before releasing it into the field. Stacking or pyramiding can be achieved either by marker-assisted selection or by genetic modification (GM) and should provide a more durable resistance as the pathogen would require mutations in multiple corresponding effector genes to evolve virulence. Since the mid-1950s, losses to stem rust disease have been reported to be minimal in the Northern Great Plains of North America due to the deployment of wheat cultivars with multiple stem rust resistance genes pyramided using marker-assisted selection in addition to the eradication of barberry (Leonard and Szabo, 2005). High levels of wheat rust resistance have also been observed using a similar approach in Australia and other wheat production areas in Asia (Bariana et al., 2007; Singh et al., 2004).

However, the conventional stacking method is challenging and time-consuming, especially when combining multiple loci within crops with long generation times, such as wheat, while linkage drag often impedes the deployment of otherwise excellent genes from wild relatives (Mago et al., 2011). Also, multiple resistance genes are unlikely to be inherited as a single unit. There is therefore a risk that they will be separated by outcrossing with other wheat cultivars in the field or when being bred into other locally adapted cultivars by breeders who may not have the facilities to track every gene in the pyramid. In turn, single genes may again be exposed to the pathogen, which could then more easily break down, thereby eroding the pyramid and putting the crop at risk. Another problem is that multiple loosely linked genes could create a "linkage block" making it difficult for the breeder to introgress novel traits into that region without breaking up the stack. A GM approach could overcome all these limitations by introducing multiple genes simultaneously at a single locus.

To date, five *R* genes against stem rust have been cloned in wheat, namely *Sr22* (Steuernagel et al., 2016), *Sr33* (Periyannan et al., 2013), *Sr35* (Saintenac et al., 2013), *Sr45* (Steuernagel et al., 2016) and *Sr50* (Mago et al., 2015). The continuing drop in the price of sequencing over the last few years has contributed to the development of novel strategies in gene cloning including mutagenesis coupled to Resistance gene enrichment sequencing (MutRenSeq) (Steuernagel et al., 2016), or chromosome sequencing (MutChromSeq) (Sánchez-Martín et al., 2016) and Targeted Chromosome-based cloning via long-range assembly (TACCA) (Thind et al., 2017). These new technologies have allowed scientists to overcome some of the challenges of cloning *R* genes by map-based cloning. This suite of *R* genes can now be utilised to engineer a stack at a single locus, which will allow the facile shuttling of the pyramid from one elite cultivar to another.

1.5 Map-based cloning of genes in Triticeae

Map-based cloning has been widely used by researchers to identify the molecular nature of genes underlying important agronomic traits in Triticeae. This traditional way of gene cloning is based on the genetic structuring of germplasm resulting from recombination between two genetically distinct accessions or cultivars. However, the application of this approach in wheat and barley has been limited by low marker density, suppressed recombination (Choulet et al., 2014; Kunzel et al., 2000) and the high cost of generating a physical contig across a genetically defined map interval due to the enormous genome size and high repeat content from transposable elements (TEs). A large, freely recombining population saturated with molecular markers is required to generate a high-resolution genetic map. The advent of

next-generation sequencing (NGS) has accelerated marker discovery such as the development of high-density genotyping arrays on a fixed platform in wheat (Cavanagh et al., 2013; Wang et al., 2014a). However, the inherent lack of recombination across large tracks of Triticeae genomes can make it very difficult to physically delimit short map intervals. This in turn increases the cost of construction and screening of BAC libraries to generate a physical map across the genetically defined map interval. Despite these challenges, at least 20 wheat genes have been successfully cloned by forward genetics using map-based approximation in the last two decades **(Table 1.1)**.

Table 1.1 Examples of genes cloned in wheat by forward genetics.

Gene	Gene function	Class	Cloning method	Reference
Sr33	Stem rust resistance	NLR	Mapping	(Periyannan et al., 2013)
Sr35	Stem rust resistance	NLR	Mapping	(Saintenac et al., 2013)
Sr22	Stem rust resistance	NLR	MutRenSeq	(Steuernagel et al., 2016)
Sr45	Stem rust resistance	NLR	MutRenSeq	(Steuernagel et al., 2016)
Sr50	Stem rust resistance	NLR	Mapping	(Mago et al., 2015)
Lr34/Yr18	Leaf rust resistance	ABC transporter	Mapping	(Krattinger et al., 2009)
Lr1	Leaf rust resistance	NLR	Mapping	(Cloutier et al., 2007; Qiu et al., 2007)
Lr10	Leaf rust resistance	NLR	Mapping	(Feuillet et al., 2003; Stein et al., 2000; Wicker et al., 2001)
Lr21	Leaf rust resistance	NLR	Mapping	(Huang et al., 2003)
Lr22a	Leaf rust resistance	NLR	Mapping and TACCA	(Thind et al., 2017)
Lr67	Leaf rust resistance	Hexose transporter	Mapping	(Moore et al., 2015)
Yr10	Stripe rust resistance	NLR	Mapping	(Liu et al., 2014)
Yr36	Stripe rust resistance	START Kinase	Mapping	(Fu et al., 2009)
Pm2	Powdery mildew resistance	NLR	MutChromSeq	(Sánchez-Martín et al., 2016)
Pm3	Powdery mildew resistance	NLR	Mapping	(Srichumpa et al., 2005; Yahiaoui et al.,

				2006; Yahiaoui et al., 2004)
Fhb1	Fusarium head blight resistance	Pore-forming toxin-like (*PFT*) gene (Agglutinin superfamily domains and ETX/MTX2 superfamily domain)	Mapping	(Bernardo et al., 2012; Cuthbert et al., 2006; Rawat et al., 2016)
Gpc-B1	Senescence and grain protein, zinc, and iron content	NAC transcription factor	Mapping	(Uauy et al., 2006b)
Q	Threshing character and spike phenotype	AP2 transcription factor	Mapping	(Faris et al., 2003)
VRN1	Vernalization response	AP1 like MADS-box transcription factor	Mapping	(Yan et al., 2003)
VRN2	Vernalization response	Putative Zinc finger and CCT domain	Mapping	(Yan et al., 2004)
VRN3	Vernalization response	RAF kinase inhibitor like protein	Mapping	(Yan et al., 2006)
Ph1	Chromosome pairing locus	Cyclin-dependent kinase (cdk)-like genes	Mapping	(Al-Kaff et al., 2008; Griffiths et al., 2006)
Tsn-1	Host-selective toxin *Ptr ToxA*	NLR and S/TPK domains	Mapping	(Faris et al., 2010)
TaPHS1	Preharvest sprouting resistance	TaMFT-like gene	Mapping	(Liu et al., 2013a)

1.6 Gene cloning using whole-genome sequencing in plants

The costs of sequencing DNA has significantly dropped over the past decade, as has our ability to computationally analyse large quantities of DNA sequence. This has enabled researchers to use next-generation sequencing technologies more routinely and accelerate the gene cloning process especially in plants with small genomes such as Arabidopsis (135 Mb) and rice (430 Mb). Several genes in these species have now been cloned by applying whole genome sequencing to mutant populations. For example, a method called SHOREmap developed by Schneeberger and colleagues allowed simultaneous mapping and identification of a gene responsible for slow growth in *Arabidopsis thaliana.* A recessive EMS-derived mutant with a growth defect was crossed to a genetically diverged ecotype followed

by one round of self-crossing. Five hundred BC_2 mutant plants were pooled and deep sequenced which resulted in identification of a map interval and candidate gene with causative mutation based on the high mutant to parent allele frequency (Schneeberger et al., 2009).

One of the limitations of SHOREMap is the requirement to cross the mutant to a genetically diverged cultivar. This can complicate the accurate phenotyping of F_2 progeny due to multiple segregating minor QTLs with potential for transgressive phenotypes, which in turn can give rise to false positive and false negative individuals in the bulks. More recently, this limitation was overcome by backcrossing the mutant plant to the non-mutagenized wildtype parent, and using the allele frequency of EMS mutations segregating in the mutant bulks to genetically map and clone the gene. The resulting application, dubbed MutMap, was used to rapidly identify mutations in rice genes responsible for causing pale green leaves and semi-dwarfism (Abe et al., 2012), and high salinity tolerance (Takagi et al., 2015).

SHOREMap and MutMap are both limited by the requirement for recombination. As mentioned above, the Triticeae genomes contain large areas of suppressed recombination (Choulet et al., 2014). The observation of low recombination rate has been well documented in barley, wheat, and some wild relatives of wheat. A detailed study of chromosome 3B of wheat revealed a patchy, highly variable recombination landscape across the chromosome. An average of 0.60 cM/Mb and 0.96 cM/Mb was observed towards the telomeres of the short and long arms, respectively, while the genetic distance dropped to 0.05 cM/Mb across the centromere and pericentromeric region (Choulet et al., 2014). This study resolved twenty-two chromosome recombination breakpoints within 1 Mb bins, and found that these were confined to only 13% of the physical chromosome (Choulet et al., 2014). The heterogeneous distribution of recombination rates was also observed in barley where 50% of the recombination occurred in only 5% of the genome (Kunzel et al., 2000).

Whilst being an important source of genetic variation in wheat breeding, alien introgressions have fallen into disfavour by many breeders due to the co-introduction of undesirable alleles of genes on the alien introgression and the difficulty of separating the target trait from these undesirable genes (also known as linkage drag). This is typically due to suppressed recombination between the alien chromatin and the domesticated chromatin. For instance, a low frequency of pairing and recombination between wheat and an alien introgression from *Haynaldia villosa* carrying the powdery mildew resistance gene *Pm26*, complicated the genetic characterisation of the locus including attempts to clone *Pm21* through map-based

approximation (Cao et al., 2011). Recombination-based cloning is further confounded by the strong divergence of copy number and sequence variation at many resistance loci (Noël et al., 1999) which further suppresses recombination or gives rise to uneven recombination (Parniske et al., 1997; Wulff et al., 2004).

1.7 Mutational genomics: cloning genes by sequence comparison of multiple mutant alleles

To circumvent the limitation imposed by uneven and patchy recombination, researchers have exploited mutational genomics in which causative mutations in a single candidate gene are identified by comparison of a wild type parent with multiple independently derived mutants. Mutagens that are commonly used in plants include ethyl methanesulfonate (EMS), sodium azide, and gamma rays. EMS and sodium azide typically produce point mutation (Olsen et al., 1993; Xu, 2010) whereas gamma rays mainly induce deletion (Morita et al., 2009). Since the mutations are relatively even and randomly distributed across a genome (Farrell et al., 2014; Krasileva et al., 2017; Shirasawa et al., 2016), the probability of obtaining multiple mutants by chance alone in a gene other than the target gene is therefore very low. The mutation density is largely dependent on ploidy level. Hexaploid and tetraploid plants tolerate a much higher mutation density than diploid plants. In diploid (2x), it is typically around 1 mutation every 300 to 500 kb, while in tetraploid (4x) and hexaploid (6x) it is around 1 mutation every 45 and 30 kb, respectively **(Table 1.2)**.

Numerous studies have used mutational genomics for gene cloning in plants with relatively small genomes as compared to wheat including *A. thaliana* (Allen et al., 2013; Austin et al., 2011; Schneeberger et al., 2009), rice (Abe et al., 2012; Takagi et al., 2015) and barley (Mascher et al., 2014). This approach can facilitate mutant identification by sequencing two or more independently derived allelic mutants and comparing the genome to search for homozygous causal mutations in the same gene as demonstrated in the cloning of the *PEP1* gene, an *A. alpina* ortholog of *FLC* responsible for flowering without vernalization (Nordstrom et al., 2013).

Mutational genomics has also been employed to clone disease resistance genes in plants with large genomes such as wheat. Disease susceptible loss-of-function mutants can easily be identified from *R* gene suppressor screens as in most of the cases, the mutations are recovered in the *R* gene itself rather than in second-site suppressors (Steuernagel et al., 2016). This property coupled with exome capture for resistance gene analogues of the NB-LRR class was recently exploited to speed up the cloning of *R* genes in wheat (Steuernagel et

al., 2016). However, this approach only captures a specific gene class. In a much less biased approach, chromosome flow sorting and sequencing of whole mutant chromosomes (MutChromSeq) allowed rapid gene isolation in barley and wheat (Sánchez-Martín et al., 2016).

Mutational genomics has also been extensively used for reverse genetics such as with TILLING (Targeting Induced Local Lesions in Genomes). This method allows identification of single base mutations in a target gene within a large mutant population using a sensitive DNA screening-technique. Introduced in the beginning of the twenty-first century (McCallum et al., 2000), the technique has since been widely adopted to various crops including wheat (Acevedo-García et al., 2017; Chen et al., 2014; Slade et al., 2005; Slade et al., 2012), rice (Tsai et al., 2011), sorghum (Blomstedt et al., 2012; Xin et al., 2008), soy bean (Dierking and Bilyeu, 2009) and potato (Muth et al., 2008). However, the laborious conventional screening for mutations in a target gene has hindered the extensive application of TILLING in crop improvement. The advent of NGS technologies has allowed the TILLING pipeline to be adopted for high-throughput screening. Recently, a wheat mutational genomics resources was developed based on exome capture and sequencing of two mutant wheat populations, one hexaploid bread wheat in the cultivar Cadenza and one tetraploid durum wheat in the cultivar Kronos (Krasileva et al., 2017). Sequencing of 2,735 mutant lines of hexaploid and tetraploid wheat revealed on average 5,351 and 2,705 mutations per hexaploid and tetraploid line, respectively, with mutation densities of 35-40 mutations per kb in each population. This resource allows the rapid *in silico* identification of mutations in different homoeologs which can then be combined through crossing to create double and triple knock-outs to overcome the effect of polyploid redundancy and permit phenotypic studies. The recent development of the speed breeding method can shorten the generation time of spring wheat to expedite the crossing process (Watson et al., 2017).

Table 1.2 Mutation density in wheat and barley.

Species	Ploidy	Mutagen	Dose (v/v)	Mutation density (mutation per kb)	Reference
Barley	2x	EMS[c]	0.2 - 0.3	1/1000	(Caldwell et al., 2004)
Barley	2x	EMS	0.2 - 0.63	1/500	(Gottwald et al., 2009)
Barley	2x	NaN_3[d]	0.1	1/374	(Talame et al., 2008)
Barley	2x	NaN_3	0.025	1/2500	(Lababidi et al., 2009)
Barley	2x	MNU[e]	0.05 - 0.015	1/504	(Kurowska et al., 2012)
Wheat	2x	EMS	0.24	1/1300	(Rothe, 2010)
Wheat	4x	EMS	nr[b]	1/44	(Krasileva et al., 2017)
Wheat	4x	EMS	0.75 - 0.75	1/51	(Uauy et al., 2009)
Wheat	4x	EMS	0.75	1/40	(Slade et al., 2005)
Wheat	6x	EMS	nr	1/30	(Krasileva et al., 2017)
Wheat	6x	EMS	0.5 - 0.7	1/37, 1/23	(Dong et al., 2009)
Wheat	6x	EMS	0.9 - 1.0	1/38	(Uauy et al., 2009)
Wheat	6x	EMS	0.75, 1.0	1/24	(Slade et al., 2005)

[a] Adapted from (Uauy et al., 2017).
[b] nr, not recorded.
[c] EMS, ethyl methanesulfonate.
[d] NaN_3, sodium azide.
[e] MNU, N-methyl-N-nitrosourea.

1.8 Genome complexity in Triticeae

In contrast to rice and Arabidopsis, the large genomes and suppressed recombination in Triticeae impose additional challenges. Even though the cost of NGS continues to fall, the genome size of 5.1 Gbp for barley and 17 Gbp for wheat **(Figure 1.4a)** would still make a whole genome shotgun-based approach costly, but also computationally challenging. Furthermore, on average ~24% and ~17% of the genes have been subject to intrachromosomal duplications in wheat (Mayer et al., 2014) and barley (Mayer et al., 2012), respectively, and the majority of both genomes (81% and 84% for wheat and barley, respectively) are also comprised of repetitive DNA, mainly long terminal repeat (LTR) retrotransposons (Mayer et al., 2012; Mayer et al., 2014) **(Figure 1.4b, c, and d)**. Moreover, the construction of a local, high-density genetic map can be troublesome if the target gene resides in pericentromeric and centromeric chromosomal regions, or an alien introgression,

all of which are characterized by low recombination rates **(Figure 1.4e)**. These genome characteristics of Triticeae complicates the bioinformatics analysis of NGS data.

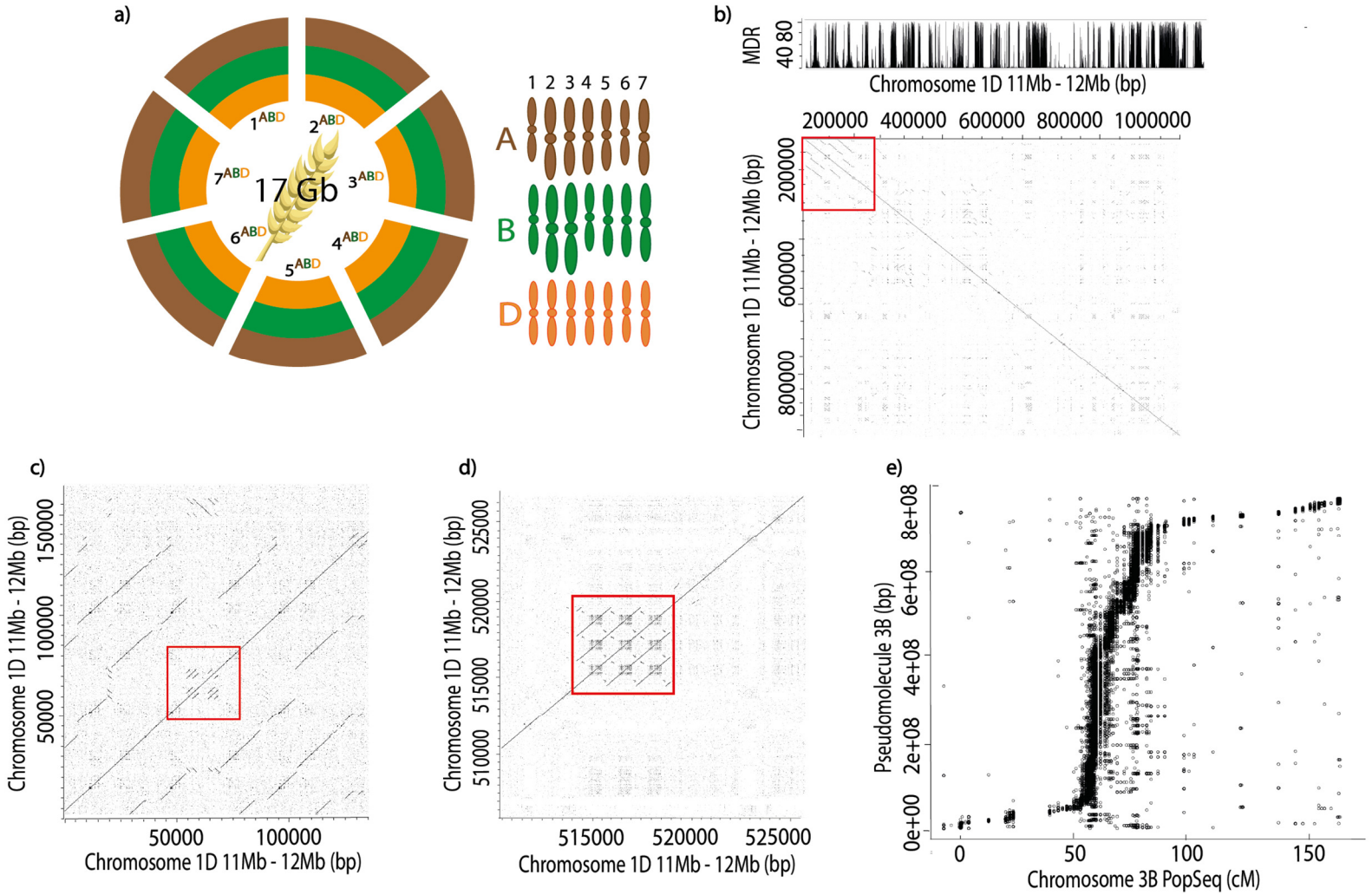


Figure 1.4 Levels of complexity within the wheat genome.

(a) The 17 Gb hexaploid wheat genome consists of three homoeologous sub genomes termed A, B, and D. (b, c and d) Dot plots of the wheat *Sr33* locus chromosome 1D (11 Mb to 12 Mb relative to the Chinese Spring IWGSC RefSeq v1.0 genome sequence (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Annotations). (b) Repetitive nature of the wheat genome. On top, a "Mathematically Defined Repeats" (MDRs) analysis (Wicker et al., 2008) of wheat chromosome 3B was carried out by counting 21-mers in the sequence and projecting the counts onto the wheat *Sr33* locus chromosome 1D (11 Mb to 12 Mb). Displayed is the average count every 10 bp using a cut-off of MDR = 100. The region outlined by a red box in (b) is shown enlarged in (c). The main diagonal represents the sequence's alignment with itself while lines parallel with and perpendicular to the main diagonal represent direct and inverted repeats, respectively, within the sequence. The region outlined by a red box in (c) is shown enlarged in (d) and represents a complex segmental duplication. The small boxes within the segmental duplication represent microsatellites. (e) Recombination landscape in wheat. From PopSeq data (Choulet et al., 2014), contigs anchored to the wheat chromosome 3B were aligned (by BLAST) against the pseudomolecule of the wheat chromosome 3B. Hits with an identity percentage of more than 98 were used and the display is the cM position from PopSeq map against the start position of the BLAST hit. A small genetic distance between 50 to 100 cM in X-axis against a large physical distance between 0e+00 and 8e+08 bp in Y-axis suggests a suppressed recombination region.

1.9 Gene cloning using genome complexity reduction

1.9.1 Exome capture

One way to facilitate NGS data analysis is by reducing the complexity of the genome prior to sequencing. Various technologies have been developed over the years which are primarily based on targeted enrichment and sequencing of a subset of the genome such as the exons or a specific gene class. Also known as reduced representation sequencing, this approach can generate simpler outputs with higher read depth to assist the identification of genetic variation in plants with large genomes and higher ploidy levels. Exome capture is a method that selectively captures and sequences the exonic regions of a genome. It involves the hybridization of an NGS library to short biotinylated nucleic acid baits (typically RNA) which are complementary to the target sequence. The hybrid molecules are captured using streptavidin-coated magnetic beads, and the purified target-enriched library is then sequenced using an NGS platform **(Figure 1.5)**. This method has been used to clone several genes in barley and wheat. For example, an exome capture of bulked early flowering segregants allowed the identification of the causal mutation in the barley *HvPHYTOCHROME C* (*HvPHYC*) gene underlying the *EARLY MATURITY 5* (*Eam5*) locus in barley (Pankin et al., 2014). Another study used exome capture on two phenotypic bulks of an F_2 mapping population segregating for a mutant phenotype to rapidly clone the *many-noded dwarf* (*mnd*) gene in barley (Mascher et al., 2014).

The exome can be further sub-divided by targeting for capture a preferential gene class within the exome **(Figure 1.5)**. One important gene class in plants are the nucleotide-binding and leucine-rich repeat (NLR) genes which encode intracellular immune receptors. Most disease resistance (*R*) genes which have been defined by genetics and subsequently cloned belong to this structural class (Meyers et al., 2003), of which a plant genome typically contains several hundred members. For example, in the barley and wheat genomes, 224 and 627 NLRs, respectively, have been defined by sequence homology searches (Sarris et al., 2016). Exome capture and sequencing for NLRs, also known as Resistance gene enrichment Sequencing (RenSeq) allows rapid and cost-effective sequencing and characterisation of the NLR complement of a plant genome. The first study to report this technology combined RenSeq with bulked segregant analysis (BSA) on two different biparental potato populations segregating for the major dominant *R* genes *Rpi-ber2* and *Rpi-rzc1* effective against *Phytophthora infestans.* This allowed identification of SNPs in NLRs linked to resistance (BSA) (Jupe et al., 2013). However, the *de novo* assembly of complete NLR genes from short reads is challenging owing to the high copy number and sequence similarity between paralogs. This

limitation was overcome in potato by combining RenSeq with the long-read single-molecule real-time (SMRT) sequencing platform to clone the broad-spectrum resistance gene against *P. infestans, Rpi-amr3i* (Witek et al., 2016). Another limitation when applying RenSeq to a segregating population derived from a bi-parental cross is the dependence on recombination which can be suppressed, or at best uneven between *R* gene haplotypes which are often diverged between two parents. To circumvent this limitation, RenSeq coupled with mutational genomics (MutRenSeq) permitted rapid cloning of the stem rust resistance genes *Sr22* and *Sr45* from hexaploid bread wheat without any positional fine mapping (Steuernagel et al., 2016). This was achieved in both cases by comparing the sequences of six ethyl methane sulfonate (EMS)-derived susceptible mutants with the wild type.

1.9.2 RNA-Seq

Exome capture and RenSeq require prior knowledge of the target gene. The capture of the target sequence by complementary nucleotide baits is biased by the accuracy of annotated reference sequence(s). Also, a single reference sequence, even if well annotated does not adequately represent the pan-genome of a species. For instances, 12,150 genes present in the pan-genome of 18 elite and commercial wheat cultivars (Montenegro et al., 2017) are missing from the set of 124,201 genes annotated in Chinese Spring by the International Wheat Genome Sequencing Consortium (IWGSC) (Mayer et al., 2014). In the case of RenSeq, successful target capture also relies on the assumption that the target gene is an NLR, which is not the case for all *R* genes. Examples of non-canonical *R* genes include the barley and wheat *Mlo* gene, which encodes seven-transmembrane (TM) proteins (Devoto et al., 2003; Konishi et al., 2010), the wheat *Lr34* gene, which encodes an ABC transporter (Krattinger et al., 2009), the wheat *Lr67* gene, which encodes a hexose transporter (Moore et al., 2015), and the wheat *Yr36* gene, which encodes a START kinase (Fu et al., 2009). Therefore, the inherent bias imposed by exome capture could in some cases result in a target *R* gene being absent from the resulting sequence data. A less biased approach can possibly be achieved by sequencing the transcriptome (RNA-Seq) **(Figure 1.5)**. Similar to exome capture, RNA-Seq substantially reduces the genome size. RNA-Seq in combination with BSA (BSR-Seq) was used in maize to clone the *glossy3* (*gl3*) gene involved in the accumulation of epicuticular wax (Liu et al., 2012). Although no gene has been cloned so far in Triticeae using this approach, the combination of RNA-Seq and BSA allowed the fine mapping of the previously cloned grain protein content (GPC) gene *GPC-B1* in tetraploid wheat (Trick et al., 2012), and the identification of SNPs closely linked to the *Yr15* gene in hexaploid wheat (Ramírez-Gonzàlez et al., 2015). In contrast to diploid maize, the *de novo* assembly of polyploid wheat

transcriptome samples is problematic due to the co-expression of highly similar homoeologous sequences. Therefore in wheat, a good reference sequence for mapping and comparison may be required. RNA-Seq is also limited by the requirement of the target gene to be expressed at the particular time points and at sufficient levels in the tissue sampled. This can be particularly problematic for genes which have a very low basal level of expression, such as some *R* genes in unchallenged hosts (MacQueen and Bergelson, 2016). Finally, RNAseq will not directly identify causative mutations in regulatory sequences.

1.9.3 Chromosome flow sorting

The biases and limitations imposed by exome capture and RNAseq can be overcome by exploiting nature's very own complexity reduction. The bread wheat and barley genomes are divided into 42 and 14 chromosomes, respectively. Differences in the physical size and sequence composition between chromosomes can be used to separate these in a technique known as chromosome flow sorting (Doležel et al., 2011). This technique provides a medium level of complexity reduction by 21-fold in hexaploid wheat and 7-fold in barley and the diploid progenitors of wheat. Moreover, the DNA sequence analysis in polyploid wheat is simplified since homoeologues, paralogues and pseudogenes from other chromosomes are excluded. Chromosome flow sorting was employed to deconvulate and sequence the 17-gigabase hexaploid wheat genome by sequencing each individual chromosome arm (except for 3B, which sequenced as a complete chromosome because it could not be flow sorted to high purity due to its relatively large size). A series of aneuploid wheat lines was used in which halves of each chromosome were missing, making them much smaller and allowing them to be easily purified. However, the success of flow sorting is determined by the ability to separate individual chromosomes based on size differences or base-pair composition. The high repetitive DNA of wheat and barley can limit the application. Recent advances in labelling repetitive DNA of isolated chromosomes, named Fluorescent *In situ* Hybridization In Suspension (FISHIS) can be used to obtain pure chromosome preparations of any chromosome from any cultivar of wheat or barley (Giorgi et al., 2013).

Mutational genomics was recently combined with chromosome flow sorting. The resulting application, MutChromSeq **(Figure 1.5)** was used in a proof-of-concept to re-clone the barley *Eceriferum-q* gene required for epicuticular aliphatic wax accumulation, and to clone the wheat *Pm2* gene conferring resistance to powdery mildew (Sánchez-Martín et al., 2016). Despite the large datasets generated by this method, in each case the identification of causal mutations in a single candidate gene assembled into a single contig with regulatory regions was achieved.

More recently, the broad-spectrum leaf rust resistance gene *Lr22a* was cloned by exploiting chromosome flow sorting and long-range scaffolding (dubbed "targeted chromosome-based cloning via long-range assembly", TACCA). The long scaffolds with an N50 of 9.76 Mb were achieved by first isolating chromosome 2D from a cultivar containing *Lr22a* using flow cytometry followed by sequencing. A cultivar-specific *de novo* chromosome assembly was generated by combining the short-read Illumina sequences with proximity ligation, a chromatin-based method that relies on favourable ligation events between the cross-linked DNA fragments. The resultant long scaffold of 6.38 Mb was used to develop additional markers which allowed the map interval to be narrowed down and restrict the number of candidate genes (Thind et al., 2017). This shows that the technology is applicable even in reduced recombination rate regions and also potentially independent of loss-of-function mutants. However, in the case of *Lr22a*, the reduced recombination rate was due to the location of the gene in the pericentromeric region, which also contains a low gene density. If the target gene were in a gene dense and recombination-suppressed region such as an alien introgression which can contain several Mbs with hundreds, or even thousands of genes, then loss-of-function mutants may be required.

One drawback of chromosome flow sorting is the requirement for a high level of expertise and specialist equipment. Only a few specialised labs have the capacity to flow sort chromosomes to a high quantity and purity. A technically less demanding option would be to obtain single copies of chromosomes, amplify these through multiple displacement amplification (MDA) and then determine their identity retrospectively with molecular markers (Capal et al., 2015). However, MDA from single chromosomes generates DNA in the region of 3 to 25 kb, while most NGS technologies for generating long reads and scaffolds, e.g. PacBio, BioNano, 10x and Oxford Nanopore require large quantities of higher molecular weight input DNA for optimal performance. Finally, the sequencing of a whole chromosome (605 to 993 Mb (average 803 Mb) in wheat) is still costly although not as high as WGS of a Triticeae genome (~4 to 17 Gb).

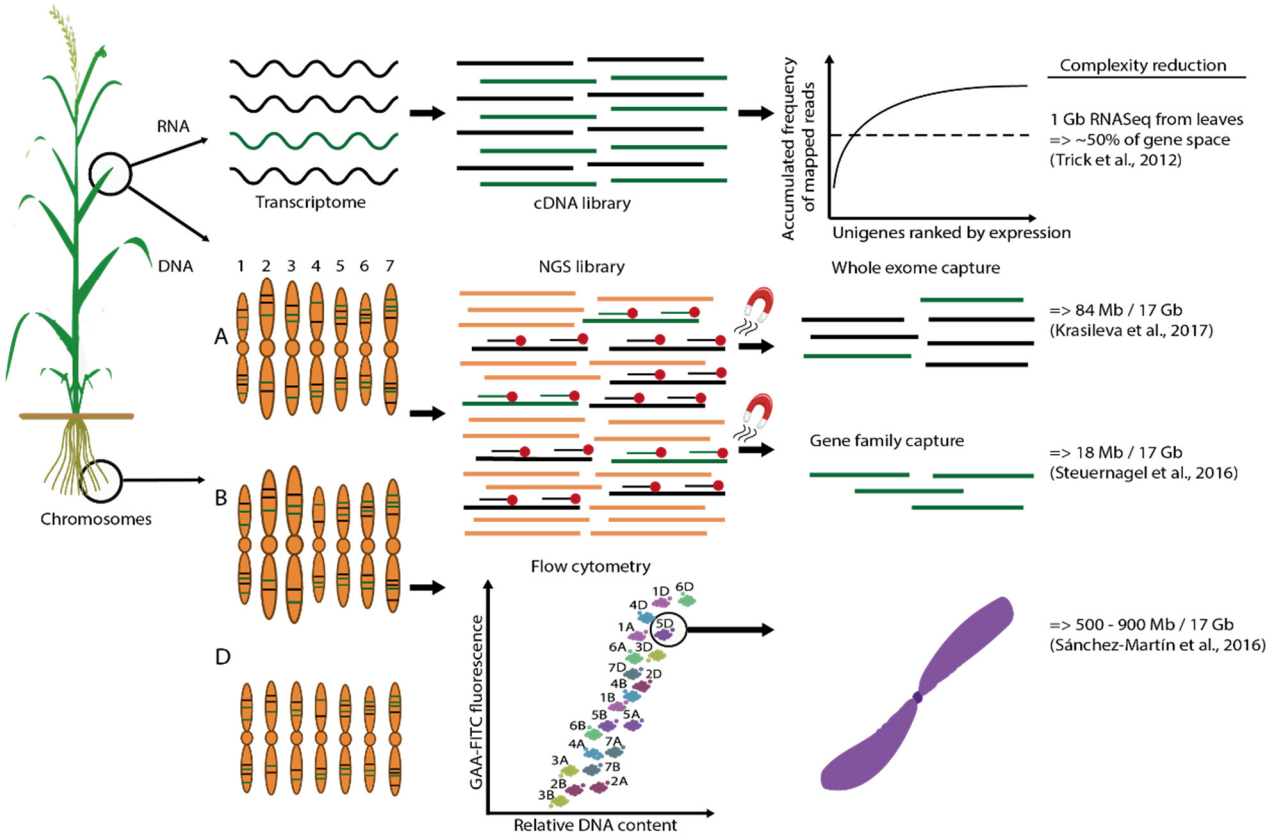


Figure 1.5 Genome complexity reduction methods.

From left to right, **(Top)** Transcriptome or RNA sequencing. RNA is isolated and purified followed by converting into complementary DNA (cDNA) for sequencing. (**Middle)** Exome sequencing. Short biotinylated RNA-baits complementary to the exonic region or a gene family sequence (e.g. NLR) are hybridized to the DNA library and purified to yield a highly enriched target sequence followed by sequencing. **(Bottom)** Chromosome flow sorting. An individual chromosome is flow sorted based on DNA content and fluorochrome signal prior to sequencing.

1.10 The poor agronomy of wild relatives of wheat and barley

The wild relatives of wheat and barley represent a tremendous source of genetic diversity that can be used to improve their domesticated brethren. The gene cloning procedures outlined above, however, require the generation of laboratory-controlled genetic population structures such as a bi-parental or mutant population, and this can pose a significant challenge in undomesticated grasses due to their poor agronomy. This includes features such as a long generation time, vernalisation requirement, seed shattering, difficulties in seed threshing, poor germination or dormancy, unruly growth habit, and a high potential for cross pollination or obligate outcrossing. The introgression of genetic diversity from wild relatives into the cultivated Triticeae species background overcomes the issue of poor agronomy. Classic examples in Triticeae include the introgression of the stem rust resistance gene, *Sr2* from *T. turgidum* subsp. *dicoccum* (McFadden, 1930a) and *Sr36* from *Triticum timopheevi* (McIntosh and Gyarfas, 1971) into wheat and *rpg6* from *Hordeum bulbosum* into barley (Fetch et al., 2009). Although some introgressions harbouring *R* gene are linked with genes that have positive impact on agronomic characteristic, many of them are associated with undesirable genes **(Table 1.3)**. The repeated failure to uncouple the deleterious traits on yield and quality linked to the target trait, also known as "linkage drag" limits the agricultural deployment. Thus, new breeding techniques and technologies need to be developed to fully access and exploit the rich source of genetic variation found in the wild relatives.

Table 1.3 Examples of linkage drag in wheat.

Origin	Introgression	Gene(s)	Impact on agronomic characteristics	References
Secale cereale	1RS.1BL translocation	*Pm8/Sr31 /Lr26/Yr9*	Improved root structure, higher grain yield	(Mago et al., 2005b; Schlegel and Meinel, 1994)
Lophopyrum elongatum	1RS translocation	*Lr19/Sr25*	Higher biomass, yellow pigmentation in flour	(Autrique et al., 1995; Sarma and Knott, 1966)
Triticum turgidum ssp. *dicoccoides*	6B substitution	*Yr36*	Higher grain protein content	(Uauy et al., 2005; Uauy et al., 2006a)

Secale cereale	1RS translocation	*Sr50*	Negative impact on bread making quality, yield reduction	(Mago et al., 2004; The et al., 1988)
Agropyron elongatum	7el2.7D translocation	*Sr43*	Distorted segregation, yellow pigmentation in flour	(Kibirige-Sebunya and Knott, 1983)
Aegilops ventricosa	7D translocation	*Pch1*	Yield penalty	(Doussinault et al., 1983; Groos et al., 2003; Worland et al., 1988)
Aegilops speltoides	4A.5BS translocation	*Pm16*	Yield penalty	(Chen et al., 2005)
Aegilops speltoides	2S.2B translocation	*Sr39*	Yield penalty, increased flour water absorption	(Labuschagne et al., 2002; McIntosh et al., 1995)

1.11 Future perspectives

A method that promises greater access to natural genetic diversity is association genetics. In contrast to map-based cloning which is most often limited by the genetic diversity and recombination rate in a biparental mapping population, genome-wide association studies (GWAS) allow detection of many allelic variations simultaneously in a diverse collection of a cultivar. This method exploits historical recombination between loci (linkage disequilibrium) in a diversity panel, often encompassing several hundreds of genetically diverse individuals, and analyses associations between marker polymorphisms and phenotypic variance across all chromosomes to search for QTL or genes associated with agronomic traits. Over the last few years numerous GWAS have been performed on crop species such as rice (Huang et al., 2012a; Huang et al., 2010; Huang et al., 2012b; Zhao et al., 2011), maize (Jiao et al., 2012; Li et al., 2013; Poland et al., 2011; Tian et al., 2011), tomato (Lin et al., 2014), barley (Cockram et al., 2010; Ramsay et al., 2011) and wheat (Ain et al., 2015; Aoun et al., 2016; Arruda et al., 2016; Maccaferri et al., 2015).

In a recent study in rice, WGS was used to improve the resolution of GWAS and allow identification of four new genes associated with agronomic traits (Yano et al., 2016). However, GWAS in polyploid crops requires a large number of markers and reference

genome sequences. Targeted association genetics on reduced representation sequencing data can be used to overcome these limitations by focussing the analysis on a particular fraction of the genome. Associative transcriptomics studies have been performed in *Brassica napus* using RNA-Seq data to associate trait variation with both the gene expression variation and transcript sequence variation (Harper et al., 2012; Lu et al., 2014). In a more recent study, exome sequencing of a barley diversity panel was used for association studies of environmental adaptation in which they observed a significant relationship between days to heading (flowering) and height with seasonal temperature and dryness (Russell et al., 2016). However, very few GWAS studies have led to direct gene identification. There are several factors that need to be considered in the experimental design for optimising association genetics including the composition of the diversity panel, exome capture design and sequencing, historical recombination (the extent of linkage disequilibrium), as well as practical considerations if working on wild relatives, and the requirement for additional experiments to test candidate genes. The functional testing of candidate genes, can be done with targeted knockouts (e.g. using CRISPR-Cas9), targeted knock-down (e.g. using RNAi), TILLING, or by generating transgenics.

It is anticipated that sequencing costs will continue to decrease, and that long read technologies will be more accurate. This will sustain a continued revolution in gene cloning from wheat and barley through both recombination- and mutation-based approaches. However, the generation of ever larger NGS data sets poses an analytical and computational challenge. Thus, development and access to more advanced computational tools and high power compute infrastructures will be required to take full advantage of the continued fall in sequencing price. This will allow routine WGS sequencing of the large and complex wheat and barley genomes, as well as their wild relatives. The availability of pan-genome reference sequences for wheat, barley and their relatives, along with access to sequence-configured diversity panels will also play an important role in accelerating gene cloning.

1.12 Research objectives

This research aimed to clone the wheat stem rust resistance gene, *Sr44* using mutagenesis and sequence capture (MutRenSeq). The generated *Sr44*-EMS mutant population was utilised to investigate meristem cell fate in wheat. We also hypothesised that the wheat NLR genes would work in barley. To test this, three of the previously cloned wheat stem rust resistance genes, *Sr22, Sr33,* and *Sr45* were transformed into barley and screened with stem rust. Twenty-two *Sr22* alleles were identified and sequence analysis of the *Sr22* locus was performed at nucleotide and amino acid level for structure and function studies. Wheat transgenics containing two of the alleles were also generated to confirm the gene postulation. Finally, as a proof-of-concept of *in vivo* sequential stacking of multiple *Sr* genes, CRISPR/Cas9 was used to repair the hygromycin phosphotransferase II (*HPTII*) in barley.

2 Study of meristem cell fate in wheat using loss-of-function mutants of the wheat stem rust resistance gene *Sr44*

2.1 Introduction

2.1.1 Targeting the wheat stem rust resistance gene *Sr44* using MutRenSeq

Wheat-alien introgression lines are a rich source of *R* genes. For many decades, breeders have been transferring resistance from wild relatives into wheat through a series of wide crossing. However, deployment of these introgression lines has been hindered by linkage drag, the co-introduction of linked, deleterious alleles of genes associated with the introgressed segment from the wild species. *Sr44* was introgressed into wheat chromosome 7A from chromosome 7Ai derived from Cauderon's (Cauderon et al., 1973) wheat-*Thinopyrum intermedium* addition L1 (Khan, 1996). Further characterisation of 36 fertile 7A/7Ai#1 recombinants enabled the mapping of the *Sr44* location on the distal segment of the short arm of chromosome 7Ai-1 of *A. intermedium* **(Figure 2.1)** (Khan, 2000). *Sr44* (previously known as *SrAgi*) confers resistance to the Ug99 race complex (Liu et al., 2013b), although virulence has been detected in isolates from Southern Africa, South America, Turkey, China (Huerta-Espino, 1992) and Australia (Ian Dundas, unpublished data). We are targeting *Sr44* for cloning by MutRenSeq. To this end, we have generated an ethyl methane sulphonate (EMS) mutant population of the *Sr44* introgression line IK1019 x Angas (Khan, 1996). We carried out an *Sr44*-mediated disease resistance suppressor screen in which we identified twelve independent susceptible mutants from 1171 M_2 families. NLRs of the ten mutants and the wild-type *Sr44* were captured and sequenced. However, sequence analysis of mutants and wildtype did not reveal a clear candidate.

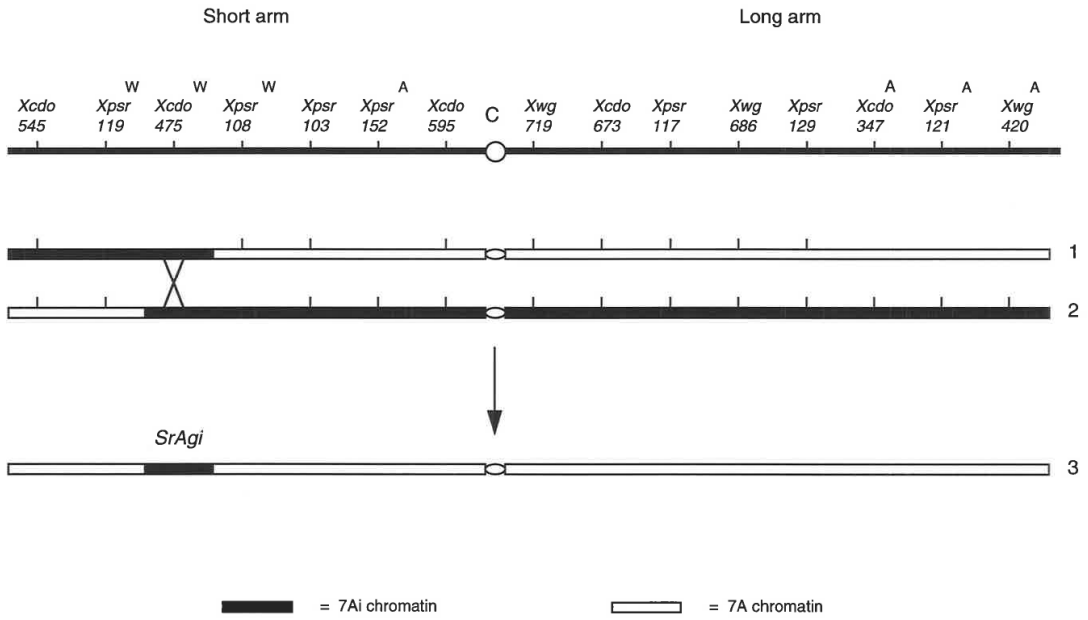


Figure 2.1 Introgression of wheat chromosome 7A from *Thinopyrum intermedium* chromosome 7Ai carrying the stem rust resistance gene *Sr44* (previously known as *SrAgi*).

On top is the linear order of the probe loci. C indicates the position of centromere. 1 and 2 are the resistant homoeologous cross-over product of chromotype B and D, respectively. 3 is the wheat chromosome having introgression segment of *Thinopyrum intermedium* carrying the stem rust resistance gene *Sr44* from chromosome 7Ai (Khan, 2000)**.**

2.1.2 Study of meristem cell fate in wheat

In plants, meristem cells proliferate and give rise to various differentiated tissues to ensure their continuous growth. Many studies have been conducted to understand the fate of meristem cells through clonal analysis such as in Arabidopsis (Irish and Sussex, 1992), maize (Johri and Coe, 1983) and barley (Döring et al., 1999). Within the meristem, a group of primordial cells act as the initial founder cells from which a specific organ is able to grow and develop. In the barley seed embryo, a tiller primordia gives rise to a tiller and, based on early observations from barley mutagenesis experiments by Stadler (Stadler, 1928), each tiller is thought to be genetically distinct (i.e. derived from a distinct meristem cell). Stadler observed that in the seed embryo the cells from which the tiller developed are already separated, thus genetically independent. Therefore, a mutation occurring in one of these cells will only affect the spike developing from this cell except on the favourable conditions when the axillary buds occurred later on in the development giving rise to clonal tillers (Stadler, 1928). Mutagenesis studies in maize have likewise demonstrated that the tassel (male) and ear (female) shoot (Coe Jr and Neuffer, 1977; Johri and Coe, 1983) and tiller (McDaniel and Poethig, 1988) are derived from independent lineages.

However, to our knowledge, meristem cell fate has not been experimentally addressed in bread wheat (*Triticum aestivum*). Knowledge of whether wheat tillers are genetically identical or independent of each other would be useful in the generation of mutant populations for suppressor screens. If independent tillers of the same plant are genetically distinct (i.e. derived from independent meristem cells), then the harvesting and screening of multiple individual spikes from the same plant would increase the effective size of an M_2 mutant population. This in turn reduces the number of seed required for mutagenesis, and it reduces space requirements for the mutant population, an important consideration when mutant populations are grown under controlled conditions in glasshouses to maximise survival and reduce the risk of pollen cross-contamination. To investigate meristem cell fate in wheat, we phenotyped the sister spikes of ten M_2 families which were identified and confirmed in the M_3 as susceptible mutants in the first screen of the primary spike. The majority (12 out of 13, 92%) of the tested sister spikes were found to be resistant suggesting that the tillers are indeed genetically distinct, while one sister spike may be derived from the same meristem cell.

Because bread wheat is a hexaploid, there is a high degree of genetic redundancy which makes it difficult to identify recessive mutations. We therefore decided to study a dominant gene. Disease resistance genes are typically dominant and obtaining loss-of-function mutants is straighforward. To study meristem cell fate in wheat, we screened for loss-of-function of a major dominant race specific disease resistance gene, namely *Sr44*.

2.2 Results

2.2.1 Targeting the wheat stem rust resistance gene *Sr44* using MutRenSeq

2.2.1.1 EMS mutagenesis of Sr44 and screening for stem rust susceptible mutants

I performed an EMS mutagenesis dose response curve by treating 100 seeds with 0.70%, 0.75%, and 0.80% EMS, respectively, in two independent experiments. Based on the number of germinated seed four weeks after planting, I estimated the EMS concentration to cause the optimal lethal dose of 50% seedling mortality or abortive growth (LD50) to be between 0.70% and 0.80% EMS **(Table 2.1)**. For each concentration, I transplanted the germinated seed into 1L pots and harvested the seed from the M_1 plants (i.e. M_2 families). In total, 264 M_1 plants were harvested for the three concentrations from two independent experiments **(Table 2.2)**. I then scaled up the mutagenesis by treating 900 seeds with 0.70% and 900 seeds with 0.80% EMS. From these two concentrations, 1079 M_1 plants were recovered. In total, 1343 M_1 plants were harvested from the pilot and scaled up experiments. M_2 families with

more than five seeds were selected for stem rust screening. 1171 M_2 families were sent to the lab of Brian Steffenson at the University of Minnesota, USA, for stem rust infection assays with *Pgt* race TTKSK to identify loss-of-function (susceptible) mutants.

From the segregating M_2 families, we identified 24 putative susceptible mutants **(Table 2.2, Supplementary Table 1 and Supplementary Figure 1)**. The progeny of these M_2 families (M_3 families) were re-tested with the same wheat stem rust race and this yielded 12 *bona fide* susceptible mutants **(Table 2.3 and Supplementary Table 1)**. However, based on genotype-by-sequencing (GBS) analysis, mutant M1 is likely to be contaminated due to undetectable introgression segment on chromosome 7A **(Figure 2.2 and Supplementary Figure 3b),** thus was excluded for further analysis. M11 and M12 were also excluded for further analysis after progeny testing of M_3 families. Assuming that the mutations are intragenic, a close to 3:1 segregation of resistant to susceptible in these 11 M_2 families strongly indicates that this gene is a dominant *R* gene **(Supplementary Table 1)**. However, the phenotype of F_1 plants derived from *Sr44* (IK1019) crossed to the recurrent parent Angas is needed to confirm this.

Table 2.1 EMS dose-response curve experiment.

	Experiment 1				Experiment 2			
EMS concentration (%)	0.00	0.70	0.75	0.80	0.00	0.70	0.75	0.80
No. of treated seed	100	100	100	100	100	100	100	100
No. of plants after 4 weeks[a]	96	23	11	2	94	83	80	76
No. of plants which matured	-[b]	19	9	2	24[b]	80	78	76

[a] After EMS treatment, the seeds were sown in P40 trays and placed in a growth chamber at 20 °C and an 8 hour photoperiod. Four weeks later, the established plants were transplanted into 1L pots and grown in a summer greenhouse.
[b] Only 24 plants were kept for self-seed.

Table 2.2 Generation of *Sr44* EMS-mutated population and screening for susceptible mutants.

	Pilot experiment			Scaled up experiment		Total
EMS concentration (%)	0.70	0.75	0.80	0.70	0.80	
No. of treated seed[a]	200	200	200	900	900	2400
M_1 plants	99	87	78	680	399	1343
M_2 families tested with stem rust[b]						1171[c]
M_3 families tested with stem rust						24
Susceptible mutants						12

[a] Germinated seeds from pilot experiment and actual scaled up experiment were grown in 1L pots and deep-root trays respectively.
[b] M_2 families from the pilot experiment were harvested as individual spikes. Three packets of individual spikes and one packet of bulked remaining spikes per plant were harvested. M_2 families from scaled-up experiment were harvested per plant. One packet of one spike per plant and one packet of bulked remaining spikes per plant.
[c] Only M_2 families with more than five seeds were tested with wheat stem rust.

Table 2.3 Stem rust infection assays with *Pgt* race TTKSK on *Sr44* M_2 and M_3 families.

Mutant	Plant	ID	M_2 families reaction[a]	M_3 families reaction[a]		
				Plant A	Plant B	Plant C
M1	33d	BW_07362	3	3^+	3^+	4
M2	56d	BW_07386	4	3	3^+	3
M3	155a	BW_07494	3	3+	4	4
M4	176d	BW_07535	3^+	4	4	3^+
M5	252e	BW_07628	4	3	3^+	3
M6	353a	BW_07748	3	3	3^+	3
M7	398h	BW_07797	3^+	3	3^+	3
M8	498a	BW_07925	3	3	3^+	3
M9	754c	BW_08228	3 (4 tip)	3	3^+	3
M10	768a	BW_08245	3	3	3^+	3
M11	484a	BW_07910	3^+	2^+,3^-	3	3^{--}
M12	40f	BW_07370	4	3	3^+	3
	Sr44		1	1	1	1

[a] Lines with infection scores ; and 0 to 2+ (high to low levels of resistance, respectively) were classified as resistant while scores above 3 were classified as susceptible lines.

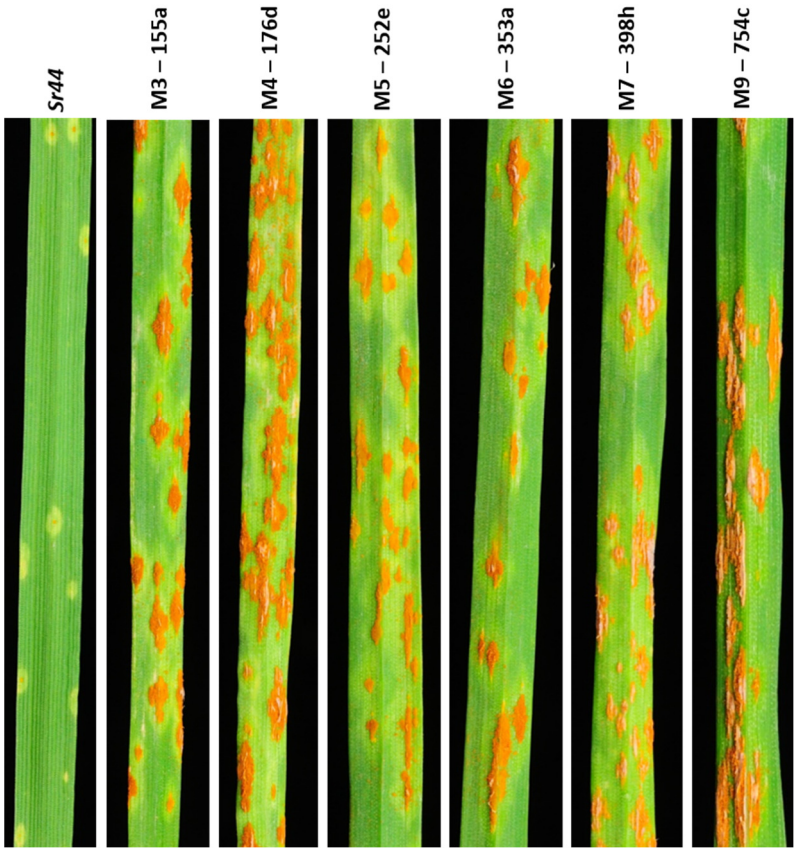


Figure 2.2 Susceptible *Sr44* loss-of-function mutants.

Stem rust infection phenotype with *Pgt* race TTKSK on selected M_3 mutants and comparison to *Sr44* wildtype.

2.2.1.2 Genotyping of Sr44 EMS mutants

To confirm that the generated EMS mutants carried the same introgression segment as in resistant *Sr44* introgression line (WT), we performed genotype-by-sequencing analysis on the resistant wildtype, susceptible recurrent parent Angas, and ten EMS-susceptible mutants The analysis revealed that all mutants carried the same *Sr44* introgression segment on chromosome 7A as in resistant *Sr44* introgression line except mutant M1 **(Figure 2.3 and Supplementary Figure 3)**.

Figure 2.3 GBS analysis on *Sr44* resistant wildtype and two susceptible mutants.

(a) Resistant introgression line (WT) (b) mutant M1 and (c) mutant M2. Red rectangular lines show single nucleotide polymorphisms (SNPs) associated with *Sr44* introgressed segment on chromosome 7A. SNPs as in wildtype can be observed in M2 but not in M1.

2.2.1.3 Complementation analysis

To investigate whether the mutations occur within the *Sr44* gene itself, or at a second, independent gene required for the resistance phenotype, I crossed six of the susceptible mutants, i.e. M2, M3, M4, M5, M6, and M7 to each other and screened the F_1 mutants with *Pgt* race TTKSK **(Table 2.4)**. Based on the stem rust reaction of each F_1 mutant cross, I generated a network showing the relationship and possible complementation groups of the mutants **(Figure 2.4)**. Mutants M8, M9, and M10 were excluded from the crossing scheme, either because of unsynchronous flowering or because the crosses were unsuccessful. Four of the F_1 mutant crosses, i.e. M3 x M4, M3 x M5, M4 x M5, and M6 x M7 showed susceptible reaction, i.e. did not complement, suggesting that the parents contain a mutation in the same gene. Moreover, the observation that all crossing combinations between M3, M4 and M5 gave rise to susceptible progeny **(Figure 2.4)**, suggest that these three parents are mutated in the same gene. However, M6 and M7, when crossed to either M4 or M5 gave rise to resistant progeny (i.e., they complemented each other, albeit that the M7 x M5 F1 was moderately resistant) suggesting that M4 and M5 are mutated in a different gene compared to that in the M3-M4-M5 trio **(Figure 2.4 and Table 2.4)**. In addition, M2 when crossed to either M3 or M4 in the M3-M4-M5 trio, also gave rise to resistant progeny, suggesting a third complementation group consisting solely of M2. In summary, the network suggests the existence of three non-complementation groups, namely M2, M6-M7, and M3-M4-M5. Other possible scenarios are considered in the discussion.

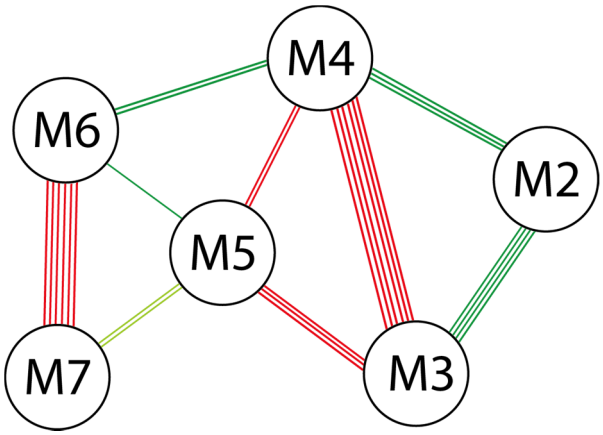


Figure 2.4 Complementation group analysis of *Sr44*.

Connecting lines represent complementation crosses performed between two mutants with the number of lines indicating the number of F_1 plants tested. Red lines represents susceptible reaction (i.e. non-complementation), while dark green lines represent high resistance reaction (i.e. complementation), and light green represents intermediate resistance reaction.

Table 2.4 Stem rust infection assays with *Pgt* race TTKSK on *Sr44* mutant F_1 complementation cross progeny.

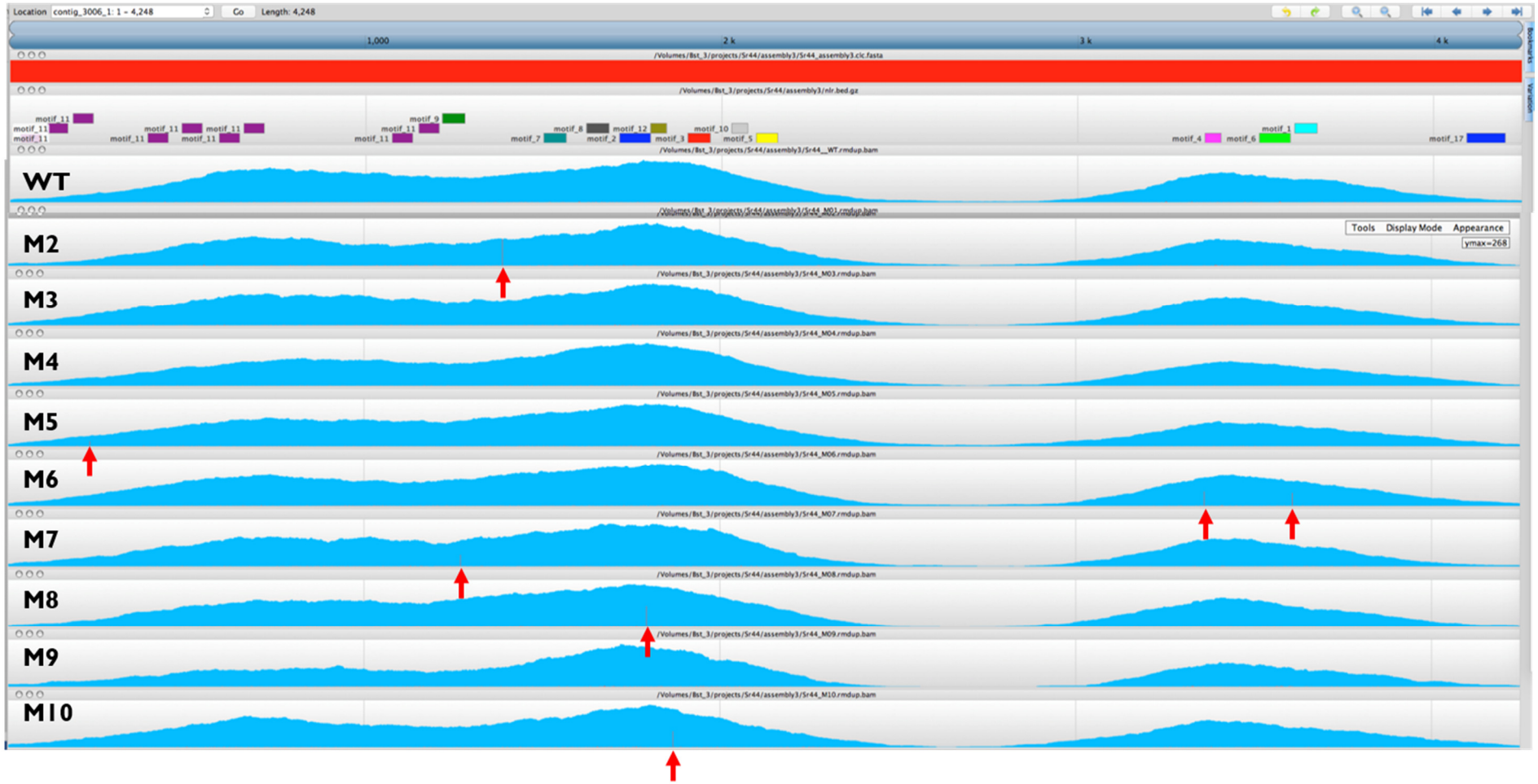
Mutant ID	Cross	No. of seeds available	No. of seeds tested	Plant 1	Plant 2	Plant 3
BW_22901	M5 x M3	1	1	4	NA	NA
BW_22902	M5 x M7	3	3	2+	2+	2
BW_22903	M5 x M3	1	1	4	NA	NA
BW_22904	M5 x M4	1	1	3	NA	NA
BW_22905	M5 x M7	1	0	NA	NA	NA
BW_22909	M7 x M6	6	3	4	4	4
BW_22910	M7 x M6	6	3	4	4	4
BW_22911	M3 x M2	2	2	11+	1	NA
BW_22917	M6 x M5	1	1	11-	NA	NA
BW_22918	M6 x M4	3	2	1	11+	NA
BW_22919	M4 x M3	4	3	33+	33+	33+
BW_22920	M4 x M2	3	2	11+	11+	NA
BW_22921	M4 x M3	4	3	33+	33+	33+
BW_22923	M4 x M2	1	1	11-	NA	NA
BW_22924	M3 x M2	3	2	11+	11+	NA
BW_22929	M5 x M3	2	1	4	NA	NA
BW_22930	M5 x M4	1	1	4	NA	NA
BW_21546	M2	10	3	33-	33-	3
BW_22936	M3	10	3	4	4	4
BW_07748	M4	5	0	NA	NA	NA
BW_22938	M5	10	3	33-	33-	3
BW_21553	M6	10	3	4	4	33+
BW_22940	M7	8	2	33+	33+	NA
BW_21536	Sr44-WT	10	3	0;1	11-	;11-

2.2.1.4 Reduced representation sequencing of Sr44 mutants

We sent high quality DNA of ten *bona fide* susceptible mutants and wildtype to MYcroarray for DNA library construction and enrichment using a Version 2 (V2) NLR RNA baits library similar to the V1 NLR library published previously (Steuernagel et al., 2016), which was used to clone *Sr22* and *Sr45* (Burkhard Steuernagel, unpublished data; **Supplementary Figure 2 and Supplementary Table 2)**. The V2 NLR bait library was improved relative to library NLR V1 by the amount of sequence information used to generate the baits. Instead of using only annotated genes available in (IWGSC, 2014), we now used additional transcriptome assemblies for *T. durum*, *T. urartu*, and *Ae. tauschii*. The V2 NLR library has successfully been used in collaborative projects to clone wheat rust *R* genes by MutRenSeq (unpublished data). The enriched DNA libraries were paired-end sequenced on the Illumina HiSeq platform and the sequences were analysed by Burkhard Steueurnagel using the MutantHunter pipeline (Steuernagel et al., 2016; https://github.com/steuernb/MutantHunter). This sequence

analysis did not identify a clear candidate. The best candidate contig (contig_3006_1) carries a point mutation with 100% allele frequency in M2. Five of the mutants (M5, M6, M7, M8, M10) also have point mutations but with a 50% allele frequency **(Figure 2.5a)**, indicative of a collapsed assembly. The candidate contig_3543_1 has nine mutants (M2, M3, M4, M5, M6, M7, M9, and M10) with many SNPs and the local alignment of these mutants to the bait source sequences are different to wildtype suggesting possible residual heterogeneity within the samples. Only one mutant (M8) carries a point mutation in this contig with 100% allele frequency **(Figure 2.5b)**.

a)


Location contig_3006_1: 1 - 4,248
Go
Length: 4,248
1,000
2 k
3 k
4 k
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_assembly3.clc.fasta
/Volumes/Bst_3/projects/Sr44/assembly3/nlr.bed.gz
motif_11
motif_9
motif_8
motif_12
motif_10
motif_7
motif_2
motif_3
motif_5
motif_4
motif_6
motif_1
motif_17
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44__WT.rmdup.bam
WT
M2
Tools
Display Mode
Appearance
ymax=268
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M03.rmdup.bam
M3
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M04.rmdup.bam
M4
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M05.rmdup.bam
M5
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M06.rmdup.bam
M6
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M07.rmdup.bam
M7
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M08.rmdup.bam
M8
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M09.rmdup.bam
M9
/Volumes/Bst_3/projects/Sr44/assembly3/Sr44_M10.rmdup.bam
M10

b)

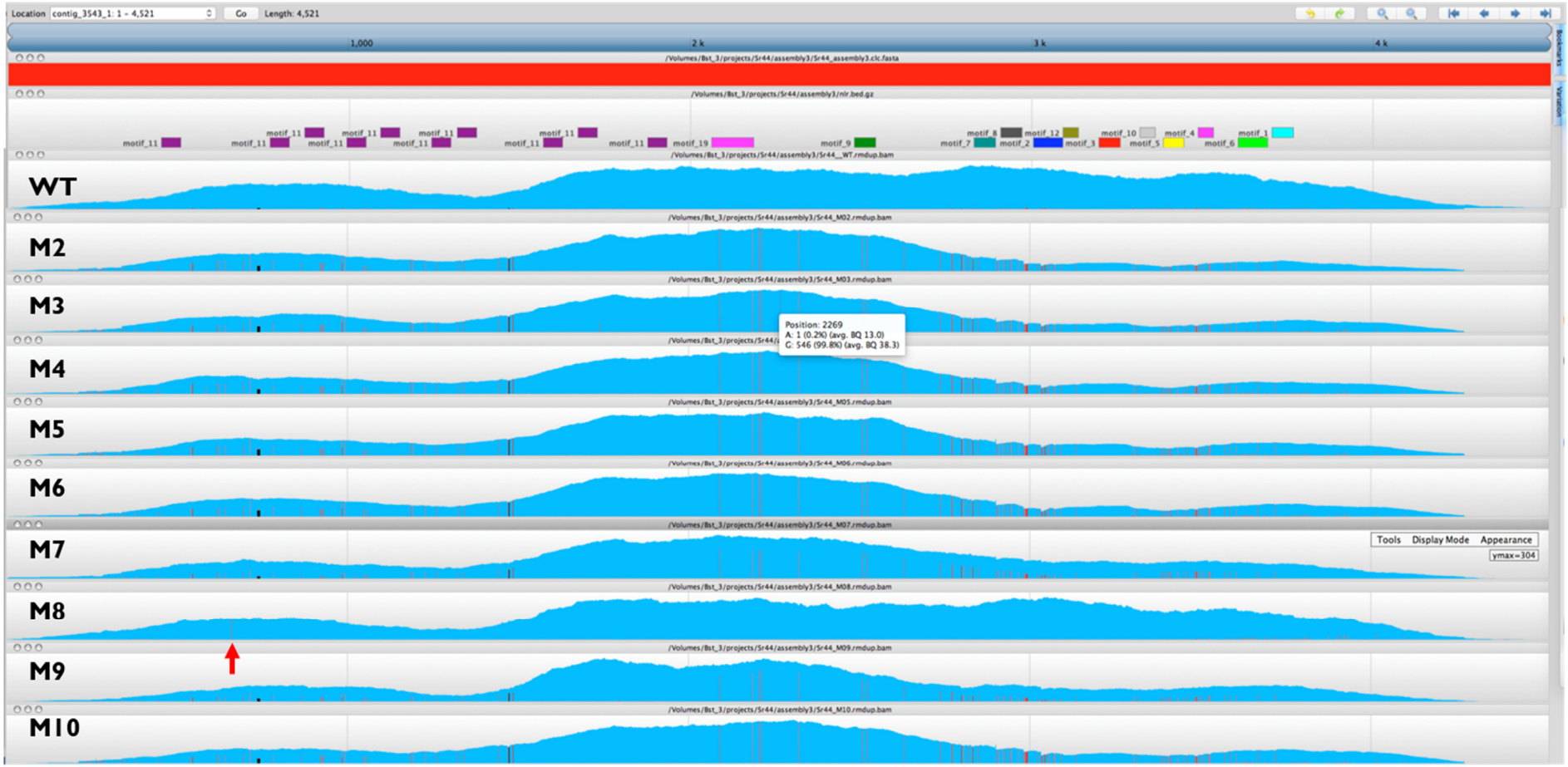


Figure 2.5 RenSeq analysis of nine *Sr44* mutants and wild-type.

The Savant genome browser screenshot shows candidate contigs of *Sr44*, (a) contig_3006_1 and (b) contig_3543_1. The upper panel corresponds to NLR motifs based on (Jupe et al., 2012a). The first row, labelled WT, corresponds to mapping of wild-type data against wild-type assembly aligned to the bait source sequence. The subsequent nine rows correspond to mapping of mutant data against *Sr44* resistant wild-type assembly aligned to the bait source sequence. SNV mutations within sub sequences with a local alignment to the bait source sequences are marked with red arrows. (a) M2 has a single SNP with 100% allele frequency. M5, M7, M8, and M10 have a single SNP with 50% allele frequency. M6 has two SNPs with 50% allele frequency. (b) Example of a contig with mapped reads containing many mismatches. M2, M3, M4, M5, M6, M7, M9 and M10 have many SNPs due to unspecific mapping. The local alignment of M2, M3, M4, M5, M6, M7, M9, and M10 to the bait source sequences are different compared to wildtype suggesting possible residual heterogeneity within the samples. M8 has a single SNP with 100% allele frequency.

2.2.2 Study of meristem cell fate in wheat

2.2.2.1 Phenotyping sister spikes from susceptible mutants

To investigate whether tillers from the same plant are derived from the same meristem cell in the seed (i.e. the tillers would be genetically related in a mutant population) or derived from distinct meristem cells (i.e. the tillers would be genetically independent of each other), we performed stem rust infection assays on sister spikes of 9 of the 12 M_2 families identified as segregating for susceptible mutant families in the initial screening and which were confirmed as *bona fide* through progeny testing. The majority of the sister spikes were found to be resistant against *Pgt* race TTKSK although for some of them only a small number of seed were available that could be tested **(Figure 2.5 and Supplementary Table 3)**. Nevertheless, all sister spikes from eight mutants, i.e. M2, M4, M5, M6, M7, M8, M9, and M10 (average of 1.4 ± 0.9 sister spikes per mutant, and 8.7 ± 9.2 seed per spike), showed a resistance phenotype suggesting that they are indeed genetically distinct from their susceptible primary spikes **(Table 2.5)**. Two of these mutants, M6 and M10, had single spikes which produced relatively large numbers of seed (35 and 17, respectively), which all gave rise to resistant seedlings. A dominant or incompletely dominant gene segregating for mutant and wildtype in the M_2 would be predicted to give rise to 25% susceptible seedlings. In this case, the probability of not getting a single susceptible seedling from 17 or 35 seedlings by chance alone would be 5.8×10^{-11} and 8.4×10^{-22}, respectively. One sister spike from mutant M3 (BW_07494_2) showed a segregation for resistance and susceptibility **(Table 2.5)**. The segregation phenotype suggests that it might be clonally related to its primary spike. Given the frequency of *Sr44* loss-of-function in this mutagenesis experiment (11 mutants in 1171 M_1 spikes) it would appear less likely that the loss-of-function *Sr44* mutant had arisen independently (*p* = (1-(11/1171)^13 = 0.88).

Table 2.5 Stem rust infection assays with *Pgt* race TTKSK on sister spikes of *Sr44* M_2 families.

Mutant	Spike ID	Sister spike ID	Line	No. of seeds available	No. of seeds tested	Susceptible (%)
M2	BW_07386	BW_07386_2	Sr44_M2	8	7	0
		BW_07386_3	Sr44_M2	7	6	0
		BW_07386_bulked	Sr44_M2	5	4	0
M3	BW_07494	BW_07494_2	Sr44_M2	19	17	11.7
M4	BW_07535	BW_07535_2	Sr44_M2	5	2	0
		BW_07535_3	Sr44_M2	6	1	0
		BW_07535_bulked	Sr44_M2	6	5	0
M5	BW_07628	BW_07628_1	Sr44_M2	2	2	0
M6	BW_07748	BW_07748_2	Sr44_M2	41	35	0
M7	BW_07797	BW_07797_1	Sr44_M2	4	4	0
M8	BW_07925	BW_07925_bulked	Sr44_M2	8	7	0
M9	BW_08228	BW_08228_bulked	Sr44_M2	9	8	0
M10	BW_08245	BW_08245_bulked	Sr44_M2	44	17	0
	McNair		Susceptible check		3	100
	Sr44		Resistant check		4	0
	Angas		Susceptible check		7	100

2.3 Discussion and conclusion

2.3.1 Targeting the wheat stem rust resistance gene *Sr44* using MutRenSeq

In EMS mutagenesis dose response curve test, we observed a large difference in number of recovered plants between two independent experiments. The number of recovered plants in the second experiment was a lot high as compared to the first experiment. This may be due to the degradation of EMS over time as the same bottle was used for both experiments.

From *Sr44*-mediated disease resistance suppressor screen, we identified twelve independent susceptible mutants. However, based on genotype-by-sequencing (GBS) analysis, mutant M1 is likely to be contaminated due to undetectable introgression segment on chromosome 7A, thus was excluded for further analysis.

Based on eight studies in diploids, polyploids, monocots and dicots, suppressor mutagenesis of major dominant *R* genes tends to give rise to a single complementation group defining the *R* gene itself (see Supplementary Table 1 in (Steuernagel et al., 2016)). However, some exceptions showed that this is not always the case. For example, an EMS suppressor screen for *Yr5*-mediated resistance to stripe rust, yielded 12 susceptible mutants. These were testcrossed to the yellow rust susceptible variety Avocet S. For seven mutants, the testcross progeny produced susceptible F_2 progeny, indicative of genetic lesions in *Yr5*. However, the other five mutants produced resistant testcross progeny, suggesting that these mutants contained lesions in one or more second site positive regulators required for *Yr5*-mediated resistance (McGrann et al., 2014). Another example is the leaf rust resistance gene *Lr10*, a CC-NB-LRR which was found to require a second CC-NB-LRR, *RGA2* (Feuillet et al., 2003; Loutre et al., 2009). In this case, screening for loss of *Lr10* mediated resistance to leaf rust yielded five mutants, three in *L10* and two in *RGA2* (Feuillet et al., 2003).

To determine the complementation groups of the *Sr44* susceptible mutants, we inter-crossed six of the mutants and screened the F_1 with *Pgt* race TTKSK. The complementation crosses suggested the presence of three complementation groups consisting of one (M2), two (M6 and M7) and three (M3, M4 and M5) mutants **(Figure 2.4)**. However, the interpretation was complicated by the observation that there was some degree of non-complementation between M5 and M7, suggesting that M5 and M7 could have a lesion in the same gene. This was not supported by the clear complementation between M6 and M5. Another possibility is that all the mutants carry lesions in *Sr44*, and that the resistant phenotype that we observed in some of the F_1 progenies results from the heterozygous combination of alleles with mutations in different parts of the gene. Indeed, inter-domain

interaction within the same NB-LRR protein molecule or homo-dimerisation between two NB-LRR protein molecules has been shown to play an important role for activating the downstream defence signalling pathway. For example, in the potato NB-LRR Rx protein a resistance response to *Potato virus X* (PVX) is mediated upon recognition of the PVX coat protein (CP) through an interaction between the CC-NB domain with the NB-LRR domain (Moffett et al., 2002), while homo-dimerisation of the TIR domain of the tobacco N protein occurs in response to the *Tobacco mosaic virus* (TMV) helicase (Mestre and Baulcombe, 2006). Such intra- and inter-molecular NB-LRR interactions may have obscured the complementation cross analysis in the case of *Sr44* by giving rise to allelic complementation in cases where the mutations occurred in different domains of *Sr44*, and assuming of course that *Sr44* encodes an NB-LRR.

It is also worth noting that some NLR proteins function in pairs to mediate immune signalling. Examples include the tobacco TIR-NB-LRR (TNL) N protein which depends on the CNL *NRG1* to mediate resistance against tobacco mosaic virus (Peart et al., 2005) and the tomato *Prf* (an NB-LRR) requires *NRC2a/b* and *NRC3* for Pto-mediated resistance to *Pseudomonas syringae* in *N. benthamiana* (Wu et al., 2016). For some NLR pairs, hetero-dimerisation between the NB-LRRs has been demonstrated. Examples include the rice CNL proteins RGA4 and RGA5 in which hetero-dimerisation of the CC domains leads to resistance signalling (Cesari et al., 2014b), and the Arabidopsis TIR-NBS-LRR RPS4/RRS1 proteins in which the TIR domains form a hetero-dimer to suppress effector-independent RPS4 activation that is induced by self-oligomerisation of the RPS4 TIR domain (Williams et al., 2014). Based on this, we cannot exclude the hypothetical possibility of hetero-dimerisation between a mutated *Sr44* gene product with another mutated NB-LRR in an F1 testcross, forming a non-activate resistance signalling complex. If this would reduce the level of functionally competent signalling complexes below a threshold required for downstream signalling, this would in turn give rise to non-complementation falsely suggesting that the mutants are in the same gene, whereas in fact they are in different genes.

To reveal more information on the nature of the *Sr44* mutants, vis-à-vis intragenic or extragenic mutations, we are planning to generate testcrosses between the mutants and the susceptible recurrent parent Angas and phenotype the F_1 with stem rust. Susceptible testcross progeny would indicate that the mutation responsible for the altered *Sr44* resistance phenotype is intragenic. In contrast, resistant testcross progeny would suggest that the mutation was outside of the *Sr44* locus (i.e. a second site positive regulator) as most likely downstream signalling components would be conserved between the *Sr44*

introgression line and the recurrent parent Angas. An exception would be if the second site positive regulator is a component (gene) which is under diversifying selection (i.e. a guardee) and has sufficiently diverged between Angas and introgression line, and this second gene is on the introgressed segment. However, examples of guardees which are physically close to their corresponding *R* genes are rare. A case in hand concerns the tomato *Prf* (an NB-LRR) and its guardee *Pto* (a protein kinase) genes. *Pto* was originally genetically defined as the *R* gene (Martin et al., 1993) and found to reside within a cluster of other kinases. The NB-LRR *Prf* also resides within this cluster (Salmeron et al., 1996) and was later found to guard *Pto* (Wu et al., 2016).

In an attempt to clone *Sr44* we performed NLR exome capture and sequencing and compared the sequence profile of nine mutants with wildtype (MutRenSeq, Steuernagel et al., 2016). We did not identify a clear *Sr44* candidate, i.e. a gene with independent mutations in the majority of the mutants. The two best candidate contigs were carefully scrutinized in the Savant genome browser **(Figure 2.5a and b)**. In the candidate contig_3006_1, four single nucleotide variants (SNVs) in mutants M6, M7, M8 and M10 with an approximate 50% allele frequency might at first glance suggest that the linker and NB-ARC domains from different NLRs have collapsed during the assembly **(Figure 2.5a)**. However, the observation that the SNV in mutant M2 has an allele frequency of 100% and lies in between mutants M7 and M8 makes this hypothesis unlikely. It would seem more likely that these SNVs correspond to residual heterogeneity from the original M_0 seed batch which was mutated, and which is still segregating in some of the mutants compared to wildtype. This in turn makes it less likely that these are causative mutations in the *Sr44* gene.

In the candidate contig_3543_1, a high number of single nucleotide polymorphisms (SNPs) in eight of the mutants is due to unspecific mapping from other loci, suggesting that these could be deletion mutants **(Figure 2.5b)**. Once again, a more likely scenario, is that this is indicative of residual presence/absence heterogeneity in the M_0 seed batch used for mutagenesis. By sequencing the samples to a higher depth and setting higher stringency during assembly, the extra SNPs should be removed during the mapping. In future cloning-by-sequencing experiments, it is therefore important to ensure that the initial starting material used for mutagenesis is as homogenous as possible. This could be achieved by developing a double haploid line from the introgression line. This does however impose a requirement for additional resources and longer timelines. Alternatively, the gene could be subjected to MutChromSeq. This gene cloning technology is based on chromosome flow sorting and sequencing of the chromosome to which the gene has been assigned (Sánchez-

Martín et al., 2016). It is less likely that there will be residual heterogeneity in the chromosome carrying *Sr44* since this chromosome was selected to be homozygous for *Sr44* after each backcross to the recurrent parent (Ian Dundas, personal communication).

The NLR assembly of the wildtype *Sr44* introgression line gave rise to 697 full length NLRs. By contrast, from unpublished analysis of the International Wheat Genome Sequencing Consortium (IWGSC) Chinese Spring Ref1 genome sequence, we know that there are about 2600 full length NLRs in bread wheat (Burkhard Steuernagel, personal communication). Therefore, it is possible that *Sr44* was not efficiently captured and assembled. If the target gene was split into several small contigs, and the mutations were spread evenly on these contigs, then *Sr44* contigs may not have been detected above the background mutation level. A new NLR library (V3) which carries introns and many more NLR source sequences (including *Sr44* transcriptome data) is currently being tested on *Sr44* and the mutants in a second attempt to clone *Sr44* by MutRenSeq.

2.3.1.1 Study of meristem cell fate in wheat

Knowledge of meristem cell fate of wheat tillers would be useful in the generation of mutant populations for suppressor screens. A more effective size of an M_2 mutant population could more readily be generated if independent tillers of the same plant are derived from independently mutated cells in the seed meristem. To investigate meristem cell fate in wheat, we carried out an *Sr44*-mediated diseases resistance suppressor screen in which we identified twelve independent mutants with complete loss of resistance from 1171 M_2 families. This suggests that there is only a single major *Sr* gene at the *Sr44* locus effective against the *Pgt* race TTKSK. If there were more effective *Sr* genes at the locus, then obtaining susceptible mutants would be considered less likely as it would require simultaneous loss of function of both genes, which would be expected to be a relatively infrequent event, given that EMS produces primarily point mutations (Krasileva et al., 2017). An exception includes a suppressor mutagenesis of the stem rust resistance gene *Sr33* in which two out of six mutants were due to deletions, rather than point mutations (Periyannan et al., 2013). Another alternative explanation, which cannot formally be excluded is that the mutations are in a shared downstream signalling component required for resistance mediated by the *Sr44* locus.

From 9 M_2 families, obtained from the primary spike of M_1 plants, and identified as segregating for resistance and susceptibility and confirmed *bona fide* susceptible in the M_3, we screened the available seed from M_2 sister spikes. The seed from the majority of the sister

spikes (n = 13) were all resistant, suggesting that they were genetically distinct from the primary spike.

One sister spike from mutant M3 segregated for resistance and susceptibility indicating that the spike is genetically identical or that the *Sr44* loss-of-function mutant is derived from an independent event. The occurrence of genetically identical spikes in an M_1 plant could be explained by the 'diplontic selection' theory in which after mutagenesis, a competition occurs between mutated cells in the mutagenised seed (Gaul, 1959). Those mutated cells that divide less efficiently due to heavy physiological or chromosomal damage will be selected against (Broertjes and Harten, 1978) which in turn could lead to a greater degree of clonal propagation in the meristems of some seed prior to tiller formation. Therefore, it is possible that in mutant M3, very few meristem cells survived the EMS treatment, and that this gave rise to genetically identical tillers.

Given that suppressor mutagenesis of major dominant *R* genes tends to give rise to a single complementation group defining the *R* gene itself (Steuernagel et al., 2016) it is likely that most of the *Sr44* mutants contain genetic lesions in *Sr44* itself. Therefore, the cloning of *Sr44* will likely allow us to reveal if the causative mutation in mutant M3 is identical in the sister spike.

To our knowledge, this is the first study in wheat documenting the propensity for genetic independence of tillers in wheat. Our results agree with previous study in which the tassel and ear shoot in X-ray-treated maize plants were shown to develop from independent lineage (Johri and Coe, 1983). In this study, we have observed that at least in the majority of cases (8 out of 9 mutants, 89%), each tiller is likely derived from an independent lineage. Therefore, in a mutagenesis programme, fewer M_1 plants need to be generated to obtain an effective population size for mutant screening. This would potentially reduce the cost, space and effort needed to identify enough susceptible mutants for cloning-by-sequencing such as MutChromSeq (Sánchez-Martín et al., 2016) and MutRenSeq (Steuernagel et al., 2016).

2.4 Materials and methods

2.4.1 Multiplication of Angas-*Sr44* alien introgression line

Seeds of Angas-*Sr44* alien-introgression line, IK1019 x Angas were obtained from Ian Dundas, The University of Adelaide, Australia. Ten seeds of the introgression line and three seeds of the recurrent parent Angas were sown on wet Whatman paper in three Petri dishes and imbibed at 4 °C for three days in the dark. They were then placed on the laboratory bench for another three days. The germinated seeds were sown into soil in 1 L pots in a summer greenhouse. To control self-pollination, the floral part of each plant was covered with a bag before anthesis. At physiological maturity, all heads from each individual plant were harvested and seeds were collected and stored at 4 °C.

2.4.2 DNA extraction and PCR marker amplification of bulked *Sr44* introgression lines

Approximately six inches of leaf samples from 3-week-old plants were collected in a 2 ml Eppendorf centrifuge tube with two tungsten beads. The samples were freeze-dried for 48 hours and then ground up using a QIAGEN Microtube Homogenizer at 29 cycle/s for 2 min in one orientation and 2 min in another orientation. The samples were ground until the tissue turned into a green-whitish powder. The powder was spun down at 2000 rpm. 800 µL pre-warmed extraction buffer was added [0.02% CTAB, 1.4 M NaCl, 0.1 M Tris pH 8.0, 0.02 M EDTA, 0.02% β-mercaptoethanol] to the sample and mixed well by shaking at a speed of 25 1/s for 1 min. The suspension was spun down at 2000 rpm. The tubes were placed in a water bath at 65 °C for 30 min, and carefully shaken for 10 s by hand every 10 min. The tubes were spun at 2000 rpm for 2 min. Then, in the hood, 800 µL of chloroform: isoamyl alcohol (24:1) was added and the caps were put on tightly. The tubes were put between tightening-boards and thoroughly mixed by hand shaking. The samples were centrifuged for 15 min at 12,000 rpm to separate the phases. 45 µL of NaOAc was added to each tube. 450 µL of the top phase was slowly transferred with a disposable 1000 µL pipet (without stirring the chloroform layer). 900 µL of absolute ethanol was added down the sides of each tube. The caps were put on and the plates were put in the refrigerator for 1 hour. Then, while holding the tube flat (sideways), the samples were carefully mixed for 1 min. The DNA globs were kept and the solution was removed using a pipet. 500 µL of 1000 µg/ml of RNase was added to 50 ml of TE. 500 µL of this TE + RNase was added to the tube. The tube was inverted to dislodge the DNA and gently re-suspended until no DNA was visible. The tubes were put into an incubator at 37 °C to allow RNA digestion for at least 1 hour. Then, 1000 µL of absolute ethanol was

added to wash the DNA for 1 hour or overnight. After that, DNA globs were kept and the solution was carefully removed. 1000 µL of 70% ethanol was added to wash the DNA and the solution was removed. The DNA was left to dry for 20 min. 100 µL of 1X TE buffer was added to the tube and the DNA was quantified using a Nanodrop. The DNA samples were stored at -20 °C or -80 °C.

Polymerase chain reaction (PCR) was used to screen for presence or absence of the *Sr44* gene in all the lines used for bulking up seed for mutagenesis. Amplification was performed using primers GWM 295-7D LEFT (GTGAAGCAGACCCACAACAC) and GWM 295-7D RIGHT (GACGGCTGCGACGTAGAG) as markers with a product of 800 bp. PCR was performed using REDTaq® ReadyMix™ PCR Reaction Mix (Sigma Aldrich) with the following program: 94 °C for 30 s, 60 °C for 30 s, (then drop the annealing temperature by 0.5 °C per cycle for the next 9 cycles) and 72 °C for 30 s during 9 cycles. 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s during 40 cycles, followed by 72 °C for 5 min.

2.4.3 EMS dosage response curve

We used seed from a bulk of 10 genotyped plants of the *Sr44* introgression line IK1019 x Angas. To identify the EMS concentration required for an LD50, the EMS dose response-curve analysis was carried out by treating four batches of 100 seeds with 0.00% (control), 0.70%, 0.75% and 0.80% EMS. All samples were treated for 16 hours at room temperature. Continous mixing of seed and EMS solution was achieved by placing the seed bottles on a gentle roller. After the EMS treatment, the seeds were washed with 10 ml tap water three times. Subsequently, the seeds were planted in a P40 tray in a growth chamber at 20 °C and 8 hour photoperiod.

2.4.4 EMS mutagenesis and screening for susceptible mutants

Based on the germination rate from the pilot experiment, treatments with a 0.70%, 0.75% and 0.80% concentration of EMS concentration were chosen to mutagenise 900 seeds in the respective concentrations in two independent experiments. After washing, the seeds were sown in P60 deep trays in a summer greenhouse. To allow self-pollination, the spike of each M_1 plant was covered with a bag before anthesis. At physiological maturity, the main head from each individual plant was harvested and M_2 seeds were collected. M_2 families with more than five seeds were screened for susceptible mutants using *Pgt* race TTKSK in the Biosafety Level 3 Containment Facility at the University of Minnesota campus in the lab of Brian Steffenson. The plants were inoculated with wheat stem rust 10 days after planting and rated for disease response 12-14 days after inoculation. M_2 families segregating for

resistance and susceptibility were grown on to produce self seed, which were harvested, disinfected, and removed from containment following approved procedures.

2.4.5 DNA extraction for RenSeq

Leaf material of ten *Sr44* susceptible mutants as well as wildtype was harvested by placing ~2.5 cm cut sections into a bag, sealing the bag, and placing into liquid nitrogen. 2 to 4 g of leaf material were placed in the chilled mortar and 2.0 to 4.0 g of grinding sand were added with some liquid nitrogen. The mixture was carefully ground to a very fine powder and transferred into a 50 ml polypropylene conical tube containing 20 ml CTAB extraction buffer. The powder was immediately mixed with buffer using a spatula. The tube was placed in a 65 °C water bath for 60 min and vortexed every ~10 min. In the fume hood, 20 ml of chloroform:isoamyl alcohol (24:1) were added and mixed thoroughly by inversion until a homogenous emulsion was formed. The tube was centrifuged for 15 min at 2500 *g* in a swing-out centrifuge to separate the aequeous and chloroform phases. The top phase (14 ml) was slowly removed with a disposable 25 ml pipet and transferred to a fresh 50 ml polypropylene conical tube. 28 ml of absolute ethanol (99.9% v/v) were added down the side of each tube (one tube at a time). The tube was mixed carefully by holding the tube flat (sideways) and rocking slowly so that the solution rocked back and forth from one end of the tube to the other. Using a pipette, DNA was carefully removed from the tube and placed into a fresh 15 ml tube. 5.0 ml 1x TE buffer were added and the tube was inverted to dislodge the DNA. The tube was gently re-suspended until no more DNA was visible. 50 µl of RNase A (1000 µg/ml) were added to a final concentration of 10 µg/ml, mixed gently and digested at 37 °C for at least 1 hr. ~5 ml of chloroform:isoamyl alcohol (24:1) were added to the tube and rocked until homogeneous. The tube was centrifuged for 15 min at 2500 *g* to separate the phases. 3 ml of the top phase were slowly removed with a disposable 1 ml pipette and transferred to a fresh 15 ml polypropylene conical tube. 6 ml of absolute ethanol (99.9% v/v) was added and carefully mixed by holding tube flat and rocking slowly so that the solution rolled back and forth from one end of the tube to the other. After mixing, a DNA blob became visible. The DNA blob was removed with a pipette and put into a fresh 15 ml tube. ~3 ml 70% EtOH were added and the DNA was washed overnight on an orbital shaker at 50 to 100 rpm at room temperature. The DNA was carefully transferred from the 15 ml tubes into 2 ml tubes with a pipette. Using a pipette tip, the DNA blob was squeezed until it had reduced in size and the alcohol had mostly gone. The tube with the DNA was placed in a fume hood until it had completely dried or left on counter overnight. 500 µl of TE buffer was added to dissolve the DNA. The tube was left in the fridge at 4 °C overnight to rehydrate it. The DNA

concentration was measured using a Nanodrop. An aliquot of DNA was also run on the gel to estimate the concentration and check for RNA contamination.

2.4.6 Genotype-by-sequencing (GBS) of recurrent parent Angas and *Sr44* introgression line

The DNA of recurrent parent Angas, introgression line (*Sr44* wild-type), *T. intermedium*, and ten EMS-susceptible mutants were sent to Jesse Poland at Kansas State University, USA for GBS. The analysis was done by Burkhard Steuernagel by mapping the GBS data from recurrent parent and introgression line to the IWGSC Chinese Spring Ref1 genome sequence. The SNPs between the recurrent parent and introgression line that had at least 3x coverage were extracted. Subsequently, the frequency of SNPs versus the number of positions with at least 3x coverage per 10 Mb interval on the chromomsomes was plotted.

2.4.7 Library construction, enrichment and sequencing of susceptible mutants

Library construction and target enrichment of ten *Sr44* susceptible mutants as well as wildtype were outsourced to MYcroarray, USA. DNA libraries were enriched using NLR RNA bait library version 2 designed by Burkhard Steuernagel (unpublished data) and synthesised by MYcroarray. The enriched libraries were sequenced with 150 bp PE reads on the Illumina HiSeq platform at Novogene, China.

2.4.8 Identification of candidate gene by RenSeq

The sequencing data of mutants and wildtype were run through the MutantHunter pipeline by Burkhard Steuernagel as described in (Steuernagel et al., 2016) to identify a candidate gene. Primary data from wild type was *de novo* assembled using CLC assembly cell (www.clcbio.com/products/clc-assembly-cell/) and standard parameters. Raw data of each mutant and wild type was aligned to the wild type assembly using Burrows-Wheeler analysis. The resulting SAM file was filtered for reads mapping as a proper pair using SAMtools and parameter –f 2. The result was converted to mpileup format using SAMtools mpileup and parameters –BQ0. *De novo* assembled wild type contigs were aligned to source sequences of the bait library using BLASTn. Only sequences of the contigs having local alignments to the source sequences were considered for further analysis. Using the mpileup format, potentially mutated nucleotide positions were identified. A position was considered for further analysis if the local coverage derived from the mapping of wild type raw data against wild type assembly was at least 10–fold and the alternative allele frequency of a mutant was at least 10%. The latter step is thought to be extremely sensitive and capable of identifying a large number of false positive positions that are filtered out in a subsequent step. Since it is highly

unlikely that two independently mutated plants have mutations at the same position, every position that was found for more than one mutant was filtered out. Regions with an average coverage less than 10% of the median overall coverage were considered as a deletion mutation for a line. Resulting candidate contigs were ranked by the number of SNV or deletion mutations within sub-sequences with a local alignment to the bait source sequences.

2.4.9 Complementation analysis

To determine whether the mutations occur in the same gene or different genes, the susceptible mutants were crossed to each other and the F_1 mutants were screened with wheat stem rust in the Biosafety Level 3 Containment Facility on the University of Minnesota campus in the lab of Brian Steffenson. To achieve uniform maturity between the mutants so as facilitate mutant intercrossing, the seed was sown every 7 days in 3 batches. In the first batch, only seeds from two late flowering mutant lines (M2 and M10) were sown (based on observation of previous multiplication). In the next two batches, seeds of all mutants were sown. Five seeds were sown in 1 L pots and grown in a growth chamber under speed breeding conditions (Watson et al., 2017). Only three germinated seed from each mutant line from all batches were kept for crossing. The female parent was emasculated to remove all immature anthers when the spike is in pre-anthesis but clear of the flag leaf. The spikelets at the base and very top were removed from the spike. The central florets were removed from the spike and the outer florets were left in place. The tops of the florets were cut off at an angle with scissors to allow easy access to the anthers. The three anthers from each floret were removed to produce a male sterile spike and the spikes were covered with bags to avoid accidental pollination. A tag was attached on each spike to identify the male and female involved in the cross. The stigmas in male sterile spikes were allowed to mature for 3 to 4 days before pollinating. During pollination, bags were removed from the emasculated spikes and seed development was checked, if all anthers were not removed, the spike was discarded. Half of the florets from a male plant that just started to shed pollen were cut off. After a few minutes, when the anthers had emerged from the floret, the pollen was dispersed into a maximum of four stigmas at a time using sterilised forceps dipped in 70% ethanol. The pollinated spikes were covered with bags and tagged with pollen recipient x pollen donor. At physiological maturity, all spikes from crossed plants were harvested and the seeds were collected and stored at 4 °C.

2.4.10 Phenotyping of sister spikes of susceptible mutants

Sister spikes of M_2 families of susceptible mutants were screened with *Pgt* race TTKSK in the Biosafety Level 3 Containment Facility on the University of Minnesota campus in the lab of

Brian Steffenson. For each M_2 family that was identified as segregating for resistance and susceptibility, and confirmed as susceptible in the M_3, the second and third spikes were selected (where available). In addition, the bulks of any remaining spikes were also included in the screening.

3 The wheat *Sr22*, *Sr33*, and *Sr45* genes confer disease resistance against stem rust in barley

3.1 Introduction

Stem rust caused by the fungus *Puccinia graminis* f. sp. *tritici* (*Pgt*) is a major threat to barley production in some parts of the world, although wheat has mostly been affected for many decades. This fungal disease is destructive to wheat and barley because it can cause severe devastation on stems and leaves resulting in a significant reduction in plant growth and yield (De Wolf et al., 2011). In 1999, a new virulent isolate of *Pgt* called Ug99 was detected in Uganda which had overcome the widely deployed wheat stem rust resistance gene, *Sr31* (Pretorius et al., 2000). Ug99 and its derivatives also have virulence against more than 80% of the world's wheat (Singh et al., 2008b). Effective ways of controlling this disease include fungicide application and breeding for resistant cultivars (McIntosh et al., 1995).

In contrast to wheat, where 82 stem rust resistance genes have been designated (McIntosh et al., 2017), only seven stem rust resistance genes have been reported in barley, including *Rpg1* (Brueggeman et al., 2002; Powers and Hines, 1933; Steffenson, 1992), *Rpg2* (Patterson et al., 1957), *Rpg3* (Jedel, 1990; Jedel et al., 1989), *rpg4* (Jin et al., 1994), *Rpg5* (Brueggeman et al., 2008; Sun and Steffenson, 2005; Sun et al., 1996), *rpg6* (Fetch et al., 2009) and *rpgBH* (Steffenson et al., 1984; Sun and Steffenson, 2005). *Rpg1* is the most widely deployed among these genes due to its broad-spectrum resistance which has remained effective for over 70 years (Brueggeman et al., 2002; Steffenson, 1992). However, a recent study revealed that this gene is not effective to the Ug99 (Steffenson et al., 2017), leaving *rpg4/Rpg5* as the only gene complex known to confer resistance to this race in barley (Steffenson et al., 2009). In addition, more than 95% and 97% of the tested lines of cultivated and wild barley, respectively, are susceptible, providing the incentive to identify novel sources of resistance (Steffenson et al., 2017).

Interspecies resistance (*R*) gene transfer can be achieved by introgression or transgenesis (Wulff and Moscou, 2014). For many decades, breeders have been introgressing resistance into wheat by performing wide crosses between wheat and its wild or domesticated relatives. Notable examples of such introgressions include the transfer of the stem rust resistance genes *Sr2* from emmer wheat (*Triticum turgidum* subsp. *dicoccum*) (McFadden, 1930a), *Sr31*, *Sr50* and *Sr1RS^{Amigo}* from rye (Mago et al., 2005b) (Mago et al., 2004) (The et al., 1991), *Sr24* and *Sr26* from *Thinopyrum ponticum* (Mago et al., 2005a), and *Sr36* from *T. timopheevi*

(McIntosh and Gyarfas, 1971). However, sexual incompatibility and long generation times impose significant barriers to successful gene introgression (Erickson, 1945) while deleterious linkage drag to undesirable alleles has hindered or at best delayed the deployment of many *Sr* genes in wheat, i.e. *Sr22* and *Sr43* due to yellow flour pigmentation and/or reduced yield and delayed heading date (Knott, 1984; Marais, 1992; Niu et al., 2014). Furthermore, when introgression lines carrying a single effective *R* gene against a certain disease are introduced into the field, this imposes a strong selection pressure on the pathogen population, which often leads to resistance breaking down and the outbreak of an epidemic (Stakman, 1957). Notwithstanding, there are a few cases where *R* genes have shown remarkable durability despite being deployed as a single gene for many years over a wide area where the disease is prevalent. Examples of such durability include *Sr31* which protected wheat from major losses for over 30 years until the Ug99 outbreak in 1999 (Ayliffe et al., 2008; Pretorius et al., 2000; Singh et al., 2006) and barley *Rpg1*, which has been widely deployed since the 1940s (Brueggeman et al., 2002). As an alternative strategy, the simultaneous deployment of several *R* genes within a cultivar is likely to prolong their efficacy in the field as there is no selective advantage for strains of the pathogen that have overcome of these *R* genes in the cultivar, thus imposing a barrier to the stepwise evolution of virulence (Dangl et al., 2013; Ellis et al., 2014). It is difficult, however, to ensure that multiple *R* genes, which may be scattered throughout the genome, remain together in a breeding programme and beyond, and thus prevent single genes from being exposed to the pathogen.

Another way of transferring *R* genes between species is by plant transformation. Advantages to this approach include that the transfer is not limited to sexually compatible species, there is zero linkage drag, and it becomes possible to stack multiple *R* genes at the same locus thus ensuring that the genes are inherited as a single unit. *R* genes tend to be functional when the transfer is between species within the same family (Wulff et al., 2011). For instance, the *Bs2* gene from *Capsicum annuum* was successfully transferred into *Solanum lycopersicum*, another Solanaceae, where it confers resistance to bacterial leaf spot (Tai et al., 1999). A few genes have also been successfully transferred to more than one species. A case in hand within the Poaceae includes the transfer of the wheat *Lr34* gene into barley (Risk et al., 2013), rice (Krattinger et al., 2016) and maize (Sucher et al., 2016). Transgenic expression of the rice receptor-like kinase (RLK) gene *Xa21* in *Citrus sinensis* (Mendes et al., 2010) and banana (Musa sp.) (Tripathi et al., 2014) is one of the few examples of gene transfer from the Poaceae family to multiple species from different families including a dicot. Interspecies gene transfer has also been achieved from dicot to monocot in the case of the transgenic

expression of the Arabidopsis EF-Tu receptor (EFR) pathogen-associated molecular pattern (PAMP) receptor-like kinase in wheat, which conferred more resistance against the cereal bacterial pathogen *Pseudomonas syringae* pv. o*ryzae* (Schoonbeek et al., 2015b). However, societal issue may limit the application of GM method due to strict regulations and public unawareness especially in Europe.

The majority of *R* genes cloned so far belong to the structural class of genes encoding nucleotide-binding and leucine-rich repeat (NLR) proteins. Plant genomes typically contain several hundred NLRs (Meyers et al., 2003). NLRs detect the presence of a pathogen by recognising pathogen effector molecules. This recognition can be direct, although more often it would appear to be indirect whereby the NLR (also known as the 'guard') recognises the interaction between an effector and its host pathogenicity target, (also known as the 'guardee') (Dodds and Rathjen, 2010). In the case of the *L6* gene from flax (*Linum usitatissimum*, a member of the Linacea) it would appear that a direct interaction with *AvrL567* enabled the functional transfer of this gene into *Nicotiana benthamiana* (a member of the Solanaceae) (Dodds et al., 2004). On the other hand, gene transfer between distantly related species is also possible in cases of indirect interaction if both the *R* gene (guard) and the plant effector target protein (guardee) are transferred. Examples include the transfer of the *Arabidopsis thaliana* (a Brassicaceae*)* guard and guardee pairs *RPS2* (Day et al., 2005) or *RPM1* (Chung et al., 2011) with *RIN4* and *RPS5* with *PBS1* (Ade et al., 2007) into *N. benthamiana.*

Barley (*Hordeum vulgare*) and wheat (*T. aestivum*) diverged from a common ancestor approximately 10 to 14 million years ago within the Poaceae family (Schlegel, 2013). We hypothesised that wheat NLR genes are likely to be functional in barley, and that wheat *Sr* genes could be used to improve the immunity of barley against wheat stem rust. Five major dominant *Sr* genes have been cloned so far in wheat namely *Sr22* which originated from *T. boeticum* (Steuernagel et al., 2016), *Sr33* and *Sr45* from *Aegilops tauschii* (Periyannan et al., 2013), *Sr35* from *Triticum monococcum* (Saintenac et al., 2013)*,* and *Sr50* from *Secale cereale* (Mago et al., 2015). All these genes encode coiled-coil (CC)-NLR proteins and confer resistance to the Ug99 race. In this study, we transformed barley with *Sr22*, *Sr33,* and *Sr45.* The resultant transgenic lines expressed high-level resistance to *Pgt* indicating that wheat *Sr* genes can be used to engineer immunity towards wheat stem rust in barley.

3.2 Results

3.2.1 Wheat transgenic lines expressing *Sr45* confer resistance against wheat stem rust

A candidate gene for *Sr45* was previously identified based on six independently-derived ethyl methane sulphonate mutants, which all had a mutation in the same gene. This candidate gene also genetically co-localised with *Sr45*-mediated resistance (Steuernagel et al., 2016). However, no *Sr45* wheat transgenics were generated to confirm this candidate as a functional resistance gene. We therefore transformed the *Sr45* candidate gene with native and non-native regulatory elements into the susceptible wheat cultivar Fielder **(Supplementary Table 8)**. A total of 12 primary transgenic plants carrying the *Sr45* candidate gene encoding a 6481 bp including 885 bp of 5' and 1508 bp of 3' native regulatory regions were recovered and inoculated with the Australian *Pgt* race 98-1, 2, 3, 5, and 6. Seven lines (PC110-1, -2, -4, -5, -7, -10, -12) had an infection type ;1-, four lines (PC110-3, -6, -9, -11) had an infection type 1 and one plant (PC110-8) had an infection type 3+, confirming the identified candidate as a functional *R* gene **(Figure 3.1)**. In addition, a total of 10 primary transgenic plants carrying the *Sr45* candidate gene driven by non-native regulatory elements *Sr33* promoter and terminator were scored with Australian race 98-1, 2, 3, 5, and 6 and all lines (except line 9) produced infection type ;1 **(Figure 3.1)**.

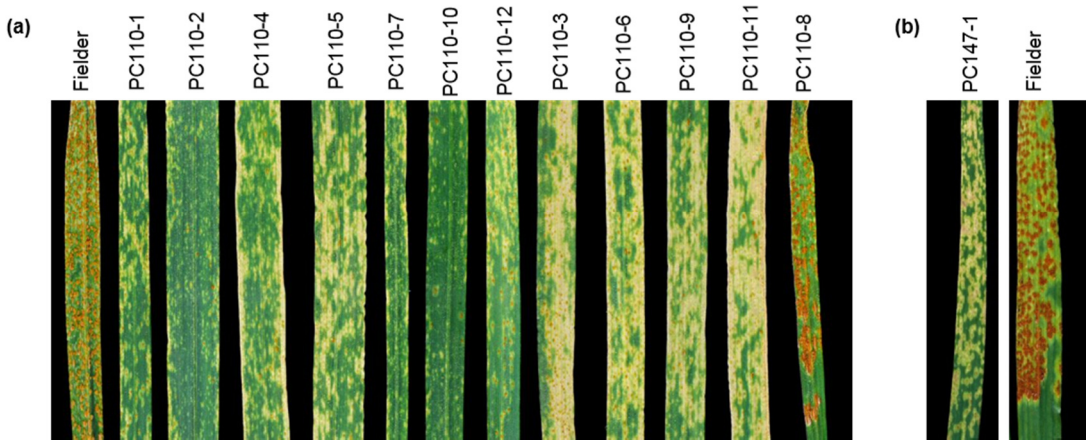


Figure 3.1 Stem rust infection assays using Australian *Pgt* race 98-1,2,3,5, and 6 on representative *Sr45* wheat T_1 transgenic lines at the seedling stage.

(a) PC110 (1-12) represent independent transgenic lines carrying the native *Sr45* gene construct and comparison to susceptible cultivar Fielder. (b) PC147-1 is a representative resistant transgenic line carrying the *Sr45* gene with the non-native regulatory elements *Sr33* promoter and terminator. The susceptible cultivar Fielder is shown on the right.

3.2.2 *Sr22, Sr33,* and *Sr45* transgenic barley lines confer resistance against wheat stem rust

To test the functionality of the cloned wheat *Sr* genes *Sr22*, *Sr33*, and *Sr45* in barley against stem rust, I transformed barley cv. Golden Promise with constructs carrying these genes using *Agrobacterium*-mediated transformation **(Supplementary Table 8)**. Single copy segregating T_2 families were selected for functional analysis. Eight plants from each line were infected with the North American *Pgt* race MCCFC which is virulent on Golden Promise (Arora et al., 2013; Kleinhofs et al., 2009). Race MCCFC is avirulent on *Sr22* (Rouse and Jin, 2011a), *Sr33*, and *Sr45* (unpublished data). Therefore, MCCFC can be used to test the functionality of *Sr22*, *Sr33*, and *Sr45*.

We selected 10 single copy *Sr22* primary transgenic lines and selfed these to produce segregating T_2 families. Eight random plants from each T_2 family were inoculated with *Pgt* race MCCFC. Four families derived from lines 1370-11-01 **(Figure 3.2a)**, 1370-17-01, 1370-19-01, 1372-08-01 segregated for resistance, while one line (1370-01-01) showed resistance in all eight tested plants **(Supplementary Table 4)**. For *Sr33*, nine out of twelve T_2 families showed segregation for resistance although the resistance infection type (IT) was not as clear as *Sr22* except for line 1024-13-01 which showed a very clear segregation between resistance and susceptibility **(Figure 3.2b and Supplementary Table 5)**. For *Sr45,* four out of eleven T_2 families showed clear segregation between resistance and susceptibility **(Supplementary Table 6)**. These results suggest that *Sr22, Sr33* and *Sr45* confer resistance against wheat stem rust in barley. In all experiments, cv. Golden Promise was included as a susceptible control. The wheat *Sr33*-containing cultivar Chinese Spring and two EMS-derived mutants carrying non-functional alleles of *Sr33* (Periyannan et al., 2013) were included in a second *Sr33* experiment as resistant and susceptible controls, respectively **(Supplementary Table 5)**.

To demonstrate that resistance to MCCFC correlates with the presence of the *Sr22,* and *Sr33* transgenes in segregating T_2 families, we performed PCR with primers specific for the *Sr* gene or the selectable marker gene **(Supplementary Table 7)** on representative resistant and non-resistant plants within a single segregating family. As expected, we obtained PCR bands corresponding to the *HPTII* transgene on the *Sr22*-resistant plant 1370-11-01 A, B, and C, while no bands was observed on the susceptible sibling 1370-11-01-D and Golden Promise **(Supplementary Figure 4c)**. On the contrary, the barley endogenous CONSTANS gene could be amplified from all samples indicating that the lack of amplification of *HPTII* from 1370-11-01-D and Golden Promise was not due to poor quality DNA **(Supplementary Figure 5)**. Similar results were obtained in the *Sr33* experiment where the *NPTII* and *HPTII* genes could be

amplified from the resistant plants 1024-13-01-A and 1024-13-01-B but not from the susceptible sibling 1024-13-01-C, or the Sr33 wheat line, the *Sr33* susceptible mutant E2 or non-transgenic Golden Promise **(Supplementary Figure 4a and 4b)**. T_2 families showed segregation for resistance although the resistance infection type (IT) was not as clear as *Sr22* except for line 1024-13-01 which showed a very clear segregation between resistance and susceptibility. These results indicate that the resistance phenotype observed in the *Sr22* and *Sr33* transgenic lines was correlated with the presence of the transgene.

To rule out the possibility that the resistant lines express an ectopic, non-specific defence reaction, we tested the *Sr33* transgenic plants with barley leaf rust race 4. All *Sr33* transgenic lines, as well as Golden Promise showed susceptibility against barley leaf rust, while the resistant control cultivar PI531901-4 was resistant, indicating that the resistance observed in the stem rust infection assays was due to a specific defence reaction to wheat stem rust **(Figure 3.3 and Supplementary Table 9)**.

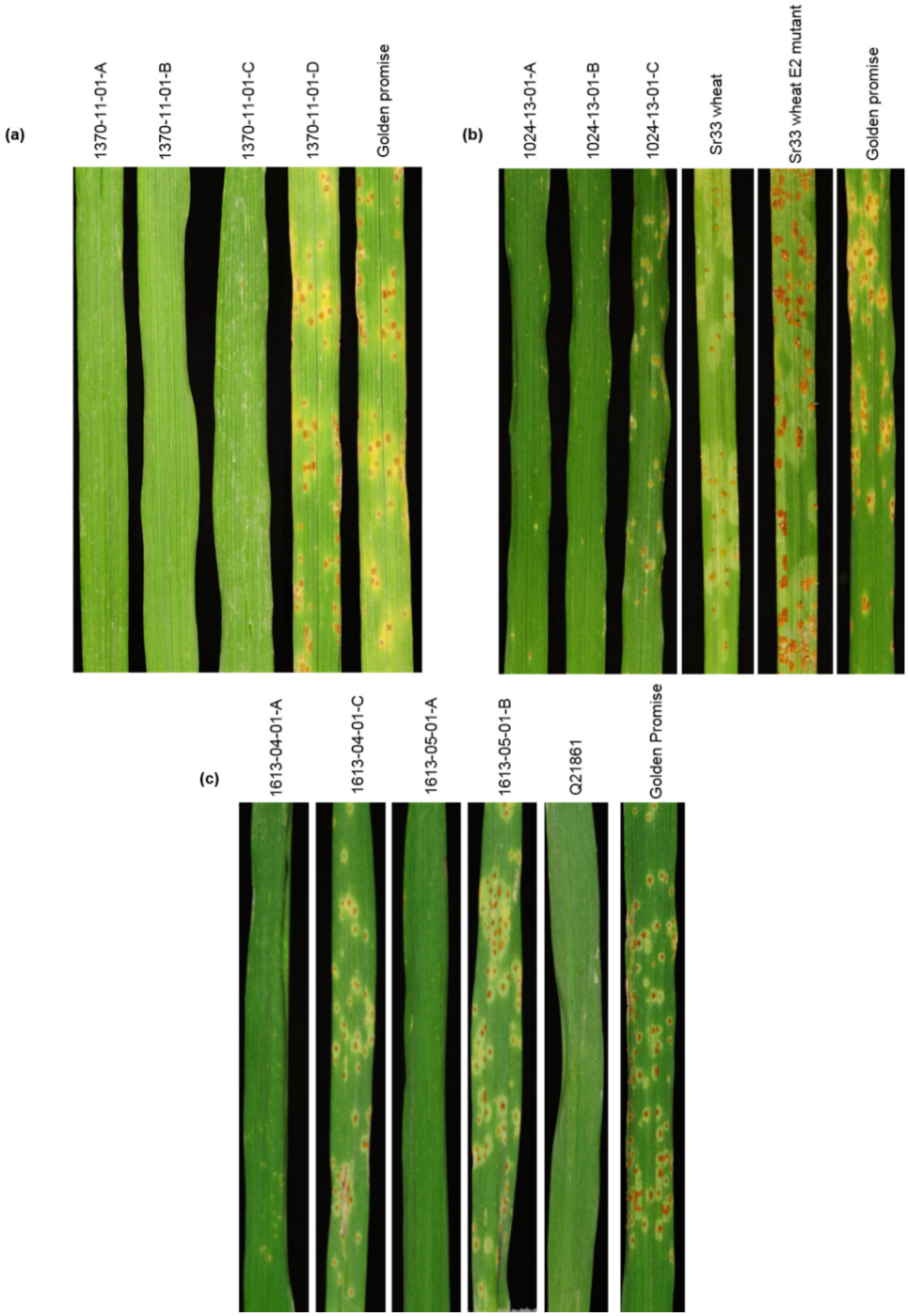


Figure 3.2 Stem rust infection assays using *Pgt* race MCCFC on *Sr22*, *Sr33,* and *Sr45* representative barley T_2 families at the seedling stage resulted in segregation of resistance and susceptibility.

(a) *Sr22* transgenics 1370-11-01 (A-D) and comparison to susceptible cultivar Golden Promise. (b) *Sr33* transgenics 1024-13-01 (A-C) and comparison to resistant *Sr33* wheat, susceptible *Sr33* EMS-induced mutant wheat line (E2) and susceptible cultivar Golden Promise. (c) *Sr45* transgenics 1613-04-01 (A and C), 1613-05-01 (A and B) and comparison to resistant check Q21861 and susceptible cultivar Golden Promise.

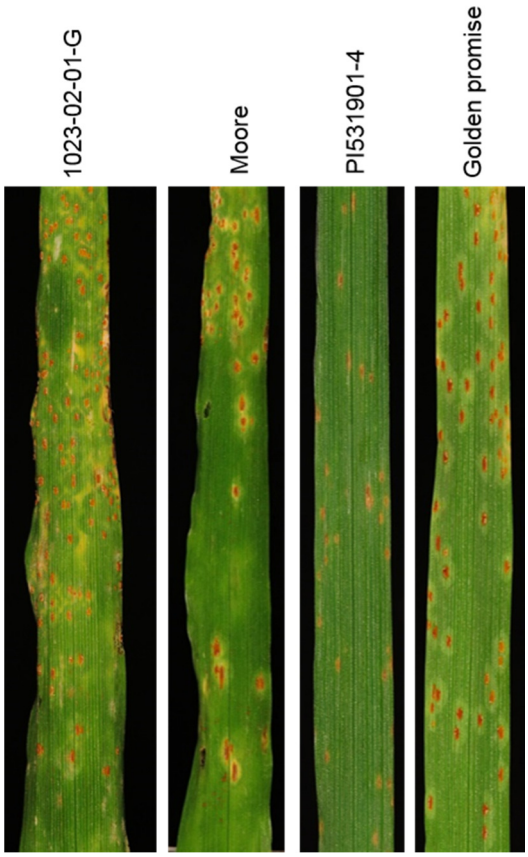


Figure 3.3 Leaf rust infection assays with *Puccinia triticina* race 4 on *Sr33* representative T_2 families at the seedling stage.

Sr33 transgenic 1023-02-01-G and comparison to the susceptible control Moore, resistant control PI531901-4, and susceptible cv. Golden Promise.

3.3 Discussion

R genes typically function when transferred from one species to another within the same family (Wulff et al., 2011). In this study, we showed that the *Sr22, Sr33,* and *Sr45* genes function when transferred into barley by conferring disease resistance against the wheat stem rust pathogen. This is consistent with previous reports on *R* gene transfer in monocots such as transfer of the maize *Rxo1* gene into rice to confer resistance to bacterial streak disease (Zhao et al., 2005) and single-cell transient expression assays of the barley *Mla6* gene in wheat to confer *AvrMla6*-dependent resistance specificity to *B. graminis* f. sp. *hordei* (*Bgh)* (Halterman et al., 2001). The observation that the wheat *Sr22*, *Sr33,* and *Sr45* genes function in barley also indicates that the downstream signalling between wheat and barley has remained conserved since the divergence of wheat and barley approximately 10 to 14 million years ago (Schlegel, 2013).

Barley is a major food staple in some parts of the world including Central Asia, North Africa, and the Baltic region (Grando and Macpherson, 2005). The emergence of the wheat stem rust Ug99 race poses a threat to barley production. A recent study revealed a very limited source of resistance to this stem rust isolate in barley (Steffenson et al., 2017). The successful transfer from wheat into barley of the *Sr22, Sr33,* and *Sr45* genes, which confer resistance to Ug99, provides additional sources of resistance. It was difficult in this study to confirm that the race specificity was maintained due to the difficulty in identifying races of stem rust virulent on Golden Promise and avirulent on either *Sr22*, *Sr33* or *Sr45*. Therefore, to show that the resistance that we observed with *Sr33* was not due to a non-specific defence reaction, we tested *Sr33* barley transgenics with another pathogen species, namely barley leaf rust. The observation that the *Sr33* transgenic lines were susceptible to barley leaf rust, indicates that the stem rust resistance is a pathogen-specific resistance reaction and unlikely due to a non-specific reaction.

In the last couple of years, many significant improvements have been made in the field of *R* gene cloning. For example, sequence comparison of multiple independently-derived mutants, facilitated by various genome complexity reduction technologies, i.e. NLR exome capture (Steuernagel et al., 2016) or chromosome flow sorting (Sánchez-Martín et al., 2016) was used to rapidly clone *Sr22*, *Sr45* and *Pm2* from hexaploid wheat. These great advances coupled with the recent availability of a wheat reference genome will greatly accelerate the *R* gene cloning process in the coming years. As more *R* genes are cloned in wheat, these could easily be tested in barley using the strategy discussed in this paper and, in the case of stem rust, this opens up a much-needed avenue for engineering genetic disease control.

While the transfer of *R* genes by transgenesis from wheat to barley has great potential as a source of novel pathogen resistance, there are some limitations that need to be considered. We need to firstly identify a cultivar that is transformable to a high efficiency and susceptible to the disease with a clear phenotype. Golden Promise is the cultivar of choice for barley transformation. However, although little resistance to stem rust has been reported in barley in general, Golden Promise does show a high level of resistance to many *Pgt* races. This made it challenging in this study to identify races which were virulent to Golden Promise and avirulent on the *R* genes that we were testing.

In summary, functional transfer of the *Sr22*, *Sr33,* and *Sr45* genes into barley has opened up a new source of resistance to barley stem rust. As more novel rust *R* genes are cloned and shown to be functional in barley, these could subsequently be deployed in a stack to provide

broad-spectrum resistance and reduce the risk of resistance breakdown. Future GM field experiments with barley plants expressing single or multiple *Sr* transgenes will be useful to assess the agronomic value of wheat *Sr* genes for barley breeding.

3.4 Materials and methods

3.4.1 Generation of binary constructs carrying *Sr* genes

For the *Sr22* construct, a 9855 bp *Sr22*-containing fragment including 2377 bp of 5' regulatory sequence (i.e. 5' of the predicted start codon) and 1560 bp of 3' regulatory sequence (i.e. 3' of the STOP codon) was synthesised by a commercial DNA synthesis provider (Life Technologies Ltd) with flanking *Not*I sites and cloned into the pVec8 binary vector (Wang et al., 1998) at the *Not*I site.

The *Sr33* and *Sr45* genes were synthesised as either a whole gene or smaller modular promoter, gene and terminator components respectively and assembled by Golden Gate cloning (Weber et al., 2011) into a toolkit vector, and then transferred to pVec8 as *Not*I fragments. For wheat transformation of *Sr45*, the assembled components were transferred to pVecBARII (a derivative of pWBvec8 in which the 35S hygromycin gene has been replaced with a 35S BAR selectable marker gene) (Wang et al., 1998). Prior to synthesis, all native *Bsa*I and *Bpi*I sites were domesticated (i.e. removed by editing) without changing the predicted amino acid coding sequence within the exons of the coding regions, and without changing the intron splice donor/acceptor sites. The domestication was performed using the program Genious.

For the *Sr33* construct, a domesticated synthetic gene of a 7854 bp *Sr33*-containing fragment including 2381 bp of 5' and 1405 bp of 3' native regulatory regions was synthesised as Level 1 modules by introducing *Bpi*I sites upstream and downstream of the respective 5' and 3'. A Level 2 reaction was performed to assemble the synthetic gene with the *NPTII* selection cassette.

The *Sr45* construct for wheat transformation was made by amplifying a genomic fragment of 6481 bp that encodes the *Sr45* gene including 885 bp of 5' regulatory sequence and 1508 bp of 3' regulatory sequence using primers S45F1 and S45R5 **(Supplementary Table 7)**. The amplified fragment was cloned into the *Not*I site of the binary vector pVecBARII.

For the *Sr45* construct for barley transformation, the *Sr33* promoter and *Sr33* terminator (Periyannan et al., 2013) were used to regulate expression of the gene. *Sr33* promoter and terminator and domesticated *Sr45* open reading were synthesised as Level 0 modules by

introducing *Bsa*I sites at the start and at the end of each module. In addition to introduction of flanking *Not*I sites, the sequence AATG was added to link the *Sr33* promoter and *Sr45* coding region and the sequence GCTT was added to link the *Sr45* coding region with the *Sr33* terminator. The assembled *Sr45* Level 1 module was transferred into the pVec8 binary vector for barley transformation. This *Sr45* Level 1 module was also transferred into the pVecBARII binary vector for wheat transformation **(Supplementary Table 8)**.

All binary plasmids containing the desired insert were transformed into *Agrobacterium tumefaciens* strain AGL1 for barley transformation.

3.4.2 Wheat transformation

The *Sr45* gene constructs with native and non-native regulatory regions were introduced into wheat cultivar Fielder by Michael Ayliffe and colleagues at the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia using the *Agrobacterium*-transformation protocol described by (Ishida et al., 2015) and phosphinothricin as a selective agent. Twelve and ten independent primary transgenic plants carrying the *Sr45* gene with native and non-native regulatory elements respectively as well as sibling lines without the *Sr45* transgene were recovered and grown in an automated growth cabinet set with day and night temperature of 23˚C, 16 hours light and 8 hours dark conditions. The plants were inoculated with the Australian *Pgt* race 98-1,2,3,5, and 6, which is virulent on Fielder at the fully developed third leaf stage. After 24 hours of incubation in a closed transparent plastic box under high humidity, the plants were restored to the original growth conditions and observed for rust development. Scoring for rust infection were done 14 days post inoculation.

3.4.3 Barley transformation

Agrobacterium-mediated transformation of *Sr* gene constructs into barley cv. Golden Promise was performed as described in (Harwood, 2014b). Ten to 12 independent primary transgenic (T_1) plants carrying the *Sr* gene construct were recovered. Confirmation that the transformants carried the *Sr* gene was done by PCR on genomic DNA using gene specific markers **(Supplementary Table 7)**. The copy number analysis of *Sr22*, *Sr33*, and *Sr45* by qPCR was outsourced to iDNA Genetics, Norwich Research Park, UK. Plants with a single copy transgene were selected and propagated for phenotyping.

3.4.4 Stem rust inoculations and phenotypic evaluations

Sr barley T_2 plants alongside with the susceptible control cv. Golden Promise were infected with *Pgt* race MCCFC at the University of Minnesota in the lab of Brian Steffenson 10 days

after planting. The inoculated plants were rated for disease response 12-14 days after inoculation.

3.4.5 PCR primers and amplification

PCR assay was performed to confirm the presence of the transgene in transgenic plants or the ability to PCR-amplify the control, the endogenous *CONSTANS* gene. Specific PCR primers for each (trans)gene were designed using the web-based application Primer3 (http:// http://primer3.ut.ee/) **(Supplementary Table 7)**. PCRs with a final volume of 20 µl contained 10 ng of genomic DNA, 10 µl of REDTaq ReadyMix PCR Reaction Mix (Sigma-Aldrich, St Louis, MO, USA) and 10 µM of each primer. The reaction schedule for each transgene was; Neomycin phosphotransferase II (*NPTII*) gene 94 ˚C for 5 min, 35 cycles of 94 ˚C for 30 seconds, 60 ˚C for 40 seconds and 72 ˚C for 70 seconds, 72 ˚C for 10 min and 16 ˚C; Hygromycin phosphotransferase II (*HPTII*) gene 95 ˚C for 5 min, 29 cycles of 94 ˚C for 30 seconds, 54 ˚C for 30 seconds and 72 ˚C for 30 seconds, and 16 ˚C; *Sr22* and *Sr45* genes 95 ˚C for 5 min, 29 cycles of 94 ˚C for 30 seconds, 59 ˚C for 30 seconds and 72 ˚C for 30 seconds, and 16 ˚C. The reaction schedule for *CONSTANS* gene was; 95 ˚C for 5 min, 33 cycles of 94 ˚C for 30 seconds, 57 ˚C for 30 seconds and 72 ˚C for 30 seconds, and 16 ˚C.

4 Allelic variation at the *Sr22* wheat stem rust resistance gene locus

4.1 Introduction

The LRR domain of R proteins has been shown to mediate ligand-receptor interactions through repeated LRR units that consists of xxLxLxx motifs. These individual units form β-strand/β-turn structures in which the variable amino acids are solvent exposed and predicted to interact with the complementary pathogen effector protein or its host target (Luck et al., 2000). Residues at these solvent-exposed positions are often found to be highly variable, suggestive of diversifying selection imposed by the pathogen on the *R* gene (Ellis et al., 2000; Parniske et al., 1997; Ravensdale et al., 2012; Seeholzer et al., 2010). The role of the LRR domain in mediating recognition specificity has been demonstrated through domain-swapping in several studies. *In vitro* intragenic exchanges of the *L2*, *L6*, and *L10* alelles from the flax *L* locus has revealed that the recognition specificity between alleles is controlled by the LRR region (Ellis et al., 1999). Other examples include the flax *P2* gene (Dodds et al., 2001), the tomato *Cf-4/Cf-9* genes (Wulff et al., 2009; Wulff et al., 2001), the potato Rx/Gpa2 genes (Rairdan and Moffett, 2006), and the barley *Mla* genes (Shen et al., 2003).

The interaction between plant resistance proteins and effector proteins can be either direct as receptor and ligand or indirect through resistance-protein sensing of effector-mediated modification of a host pathogenicity target (or guardee) (Cui et al., 2015). The direct interaction between an R protein and its corresponding *Avr* effector has been demonstrated in the rice-rice blast pathosystem, where the *Magnaporthe grisea* effector *AvrPita* was found to physically bind the LRR region encoded by the rice resistance (*R*) gene *Pi-ta* in a yeast two-hybrid assay and *in vitro* (Jia et al., 2000). Similar biochemical approaches were used to describe the direct association of the LRR domain of the flax *L* alleles with the flax rust *AVR-L567* alleles (Dodds et al., 2006). The direct interaction between LRR and effector has also been observed between the Arabidopsis RPPI resistance protein and the *Hyaloperonospora arabidopsis* ATR1 effector using a co-immunoprecipitation assay (Krasileva et al., 2010). Although the LRR domain appears to be the major determinant of recognition specificity in most studied systems, the CC domain has also been found to be involved in some *R-Avr* interactions. Examples include the rice *Pik* and *M. grisea Avr-Pik* (Kanzaki et al., 2012), and the potato R protein RB and the *Phytophthora infestans Avr* protein IPI-O (Chen et al., 2012). In a rather rare case, in addition to the LRR region, the N-terminal TIR domain of *L* alleles in flax has also been demonstrated to influence the rust resistance specificity (Luck et al., 2000).

Recent genetic and genomic studies have shed new insights on the molecular evolution of *R* genes and the mechanism that generate sequence diversity in these genes. Genome sequencing revealed that the majority of NB-LRR–encoding genes reside in clusters. For example, 73% of the mapped NB-LRR genes grouped into 63 clusters in the potato genome (Jupe et al., 2012b). This is also the case for other plant species such as Arabidopsis and rice (Bai et al., 2002; Meyers et al., 1999). Some NB-LRR genes appear to be in a simple loci with a single gene family such as *RPM1* (Stahl et al., 1999) and *RPP13* (Bittner-Eddy et al., 2000) in Arabidopsis while most of them reside in a complex loci with related paralogs such as *Cf-4/Cf-9* in tomato (Parniske et al., 1997; Parniske and Jones, 1999), *Dm3/13* in lettuce (Meyers et al., 1998b), and *I2* in tomato (Ori et al., 1997). The clustered arrangement of these genes provides a reservoir of genetic variation to generate new resistance specificities through recombination or gene conversion (Ellis et al., 2000; Hulbert et al., 2001; Michelmore and Meyers, 1998). Extensive historical recombination between paralogs within *R* genes clusters have been observed in tomato (Parniske et al., 1997), Arabidopsis (Meyers et al., 1998a), and flax (Dodds et al., 2001; Ellis et al., 1999) based on a patchwork or mosaic pattern of sequence blocks. Genetic analyses revealed that recombination played a central role in the evolution of new specificities at the *L, M, N,* and *P* loci in flax (Dodds et al., 2001; Ellis et al., 1999), *Rp1* rust resistance complex in maize (Hulbert, 1997), and *Cf4/Cf-9* homologs in tomato (Parniske et al., 1997; Parniske and Jones, 1999).

Unequal recombination can either occur frequently in some *R* gene clusters (such as *Rp1* gene clusters of maize) or rarely in others (such as *Dm3* of lettuce) (Michelmore and Meyers, 1998). Meyers and his colleagues demonstrated that tandem duplication events between paralogous genes have been an important mechanism in generating sequence variation in Arabidopsis NB-LRR genes (Meyers et al., 2003). Gene duplication can also explain the observation of gene copy number variation among haplotypes within a same species such as in lettuce *Dm3* (Kuang et al., 2004) and potato *MLB* clusters (Jupe et al., 2012b). In *R* gene clusters with a simple locus structure, strong diversifying selection can also result in the evolution of many alleles through equal recombination between allelic variants. This was observed in the flax *L* locus in which frequent interallelic recombination events gave rise to at least a dozen novel alleles (Dodds et al., 2006). Very rarely, ectopic recombination between different clusters, even between clusters located on different chromosomes, may also occur (Parniske and Jones, 1999). As a consequence, although the *R* gene clusters usually contain closely related sequences, individual genes may exhibit closer phylogenetic

relationships to sequences from other clusters within the same genome than the genes within the same cluster.

The *Sr22* gene was recently cloned using a mutagenesis and sequence capture approach from a hexaploid wheat introgression line, Schomburgk, in which the gene had been previously introgressed into chromosome 7A from the diploid A genome relative, *T. boeoticum* (Steuernagel et al., 2016). This gene confers broad spectrum resistance to multiple races of *Pgt* including the Ug99 race group (Steuernagel et al., 2016). The wild relatives of wheat can display a wide allelic diversity at the loci contributing to a particular trait. Indeed, we previously (Steuernagel et al., 2016) identified fourteen *Sr22* sequence variants by PCR screening and sequencing of accessions of *T. boeoticum* (and its domesticated form *T. monococcum*) which had been postulated to carry functional or non-functional alleles of *Sr22* (Rouse and Jin, 2011a).

In this study, we describe an additional eight *Sr22* alleles (obtained from eight accessions). However, six of the alleles were found to share identical sequences, and were thus removed from further analysis. For the same reason, two of the variants from the fourteen that were reported in the previous study were also removed, bringing the total to fourteen alleles from twenty-two accessions. We show that the *Sr22* locus is a simple, single gene locus based on the phylogenetic organisation in the Chinese Spring wheat genome. We also demonstrate through nucleotide sequence analysis that some alleles have undergone historical sequence exchange in the LRR region. To confirm the gene postulation of two of the previously identified *Sr22* alleles, wheat transgenics were generated and phenotyped. We also compared their postulated function with their predicted amino acid sequences but did not identify a region associated with functional resistance against wheat stem rust.

4.2 Results

4.2.1 Analysis of the Chinese Spring NB-LRRs identifies a homolog of *Sr22* for each homoeologous chromosome in the Chinese Spring wheat genome

R genes are often present in clusters of related paralogs. Recently, a whole genome shotgun and sequence assembly of the Chinese Spring wheat genome was released by the International Wheat Genome Sequencing Consortium (IWGSC, unpublished data). To study the phylogenetic organisation of the *Sr22* locus in Chinese Spring, we consulted the Chinese Spring NB-LRR annotation and phylogenetic tree generated by Burkhard Steuernagel and colleagues (unpublished data). We identified one homologue of *Sr22* for each homoeologous chromosome A, B, and D. Only one NB-LRR gene was found at 1 Mb upstream of the *Sr22*.

However, based on dot plot, this gene is completely unrelated to *Sr22.* These data indicate that only one *Sr22* orthologue is found in the Chinese Spring wheat A, B and D genomes, and that these orthologues occur as a singleton rather than an NB-LRR cluster **(Figure 4.1)**.

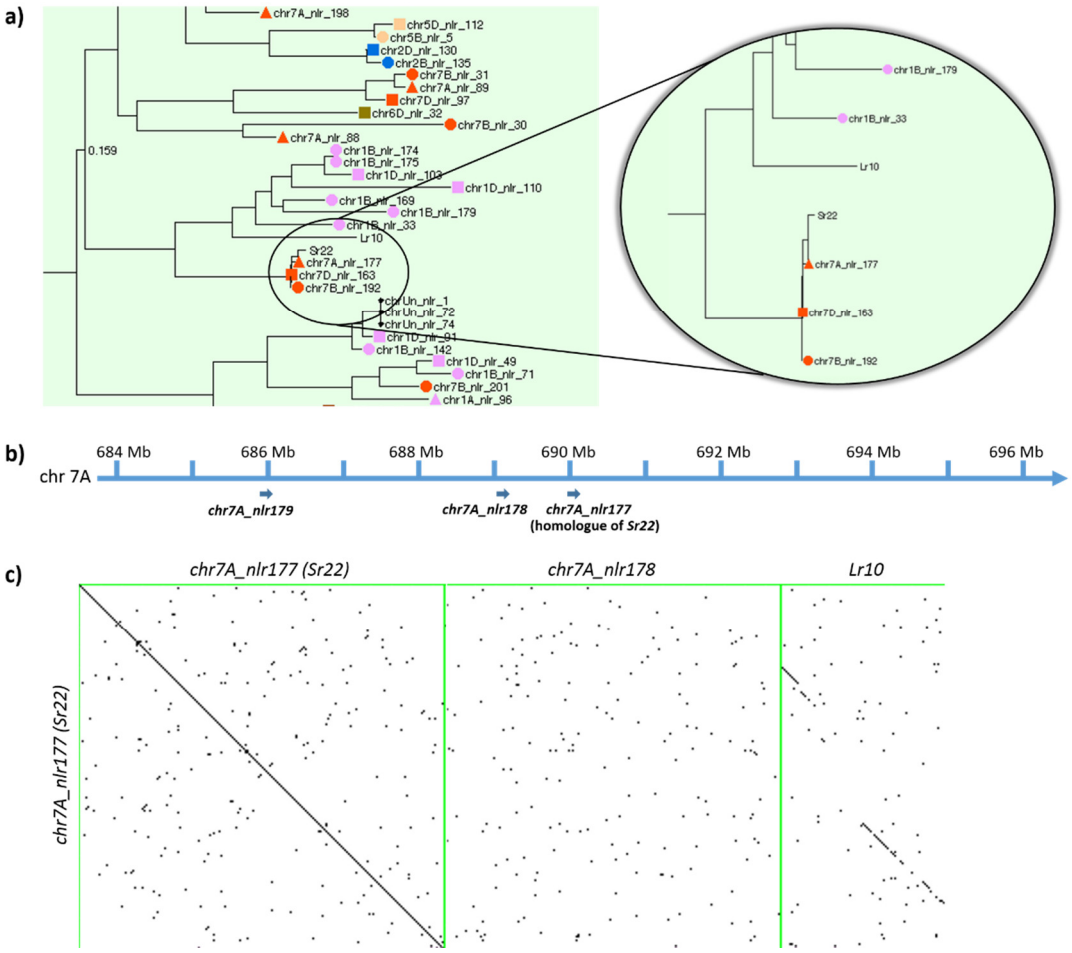


Figure 4.1 Phylogenetic organisation of the *Sr22* gene locus in the Chinese Spring wheat genome.

All NB-LRRs were annotated from the Chinese Spring wheat genome by Burkhard Steuernagel (unpublished data). (a) One homologue of *Sr22* for each homoeologous chromosome was identified based on a phylogenetic tree of all the NB-LRRs. (b) Closer observation of the locus in the sequenced Chinese Spring wheat genome revealed only one NB-LRR, chr7A_nlr178 at 1 Mb upstream of the locus. (c) Dot plots showed that chr7A_nlr178 is completely unrelated to *Sr22* homologue (chr7A_nlr177). As a positive control, *Lr10* showed some degrees of homology with *Sr22* homologue (chr7A_nlr177).

4.2.2 Historical recombination at the *Sr22* locus

Fourteen *Sr22* variants were previously identified based on PCR screening and sequencing of accessions of *T. monococcum* and *T. boeoticum* that had been postulated to carry *Sr22* (Steuernagel et al., 2016). Two of the variants were removed from further analysis due to possible germplasm redundancy. Further screening and sequencing done by Sambasivam

Periyannan at CSIRO revealed an additional eight sequences from both species (Sambasivam Periyannan and Matt Rouse, unpublished data). Similar to previous screening, six of the eight sequences are identical, bringing the total tally to fourteen alleles. To compare the nucleotide sequences of these fourteen *Sr22* variants, we aligned them using the multiple sequence alignment tool Clustal Omega and analysed the alignment using Jalview. To obtain a better overview of the alignment, we aligned only informative polymorphic sites, IPS (a polymorphic nucleotide which is shared between at least two sequences within the alignment) of the variants. This reduced the total length of the sequence alignment from 2,841 bp to 135 IPSs **(Figure 4.2)**. Based on stretches of near-continuous sequence affiliation the *Sr22* alleles could be broadly grouped into three groups; (i) IG44855, Schomburgk, and IG44857 (ii) PI289605, PI330550, and PI90945 and (iii) PI573523, IG44878, and DV92 while the remaining five alleles did not share a strong sequence affiliation with any of these groups. The first group shared almost 100% sequence identity except for IG44857 at nucleotide positions 281 to 520 **(Figure 4.2)**. Within the other two groups, more pronounced sequence variation was observed. We also highlighted sequence patches of at least three consecutive IPS that deviate from the consensus sequence and are shared between at least two *Sr22* variants **(Figure 4.2)**. PI289605, PI330550, PI90945, PI573523, and IG44878 share almost continuous sequence affiliations from position 2400 to 2572, but PI289605 and PI330550 are most closely related to Schomburgk from position 276 to 2272. This patchwork pattern is indicative of sequence exchange between *Sr22* alleles resulting from crossing over or gene conversion as observed in the tomato *Hcr9* genes (Parniske et al., 1997; Parniske and Jones, 1999). A similar pattern can also be seen in PI90945 which is most closely related to IG44857 from position 284 to 2272 although it shares an IPS sequence signature with PI289605 and PI330550 from position 768 to 2809. Several possible recombination breakpoints within the codons were found **(Figure 4.2)**.

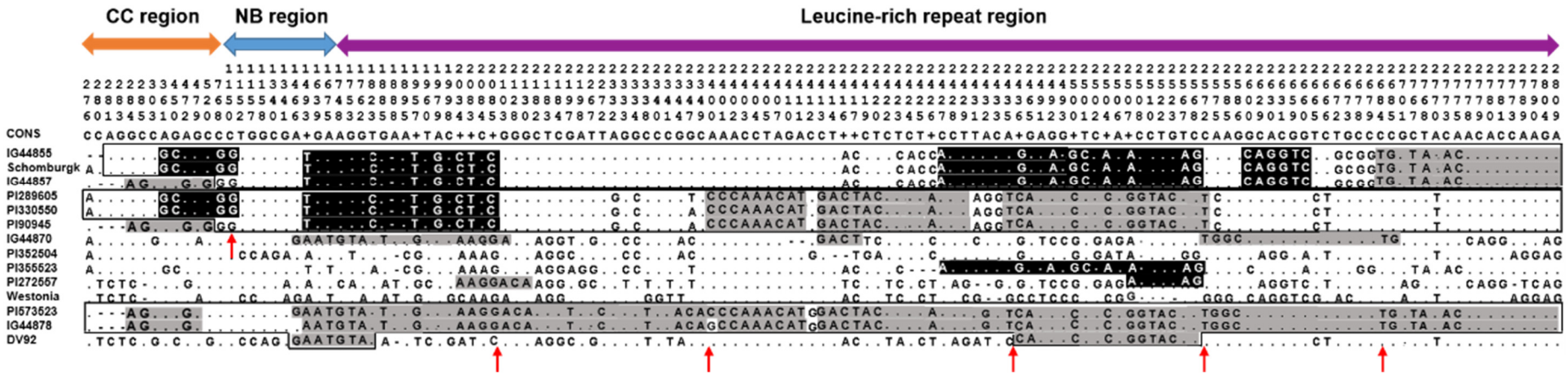


Figure 4.2 Extensive historical intragenic recombination between *Sr22* variants.

Only informative polymorphic sites, IPS (a polymorphic nucleotide which is shared between at least two sequences within the alignment) of the nucleotide sequences of the fourteen *Sr22* variants are shown. This reduces the total length of the sequence alignment from 2,841 nucleotides to 135. Vertical numbers refer to nucleotide position within the global alignment of 2,841 nucleotides. The domain structure of the *Sr22* protein according to Steuernagel et al. 2016 is indicated in the upper panel. The red arrows show possible recombination breakpoints within the codons. The almost continuous sequence affiliations within the variants of IG44855, Schomburgk, and IG44857 (upper), PI289605 and PI330550 (middle), and PI573523 and IG44878 (bottom) are boxed by lines. Sequence patches of at least three consecutive IPS that deviate from the consensus sequence and are shared between at least two *Sr22* variants are highlighted by black and grey boxes. Patches shared with IG44855, Schomburgk, IG44857, PI289605, PI330550, PI90945, PI355523, and PI272557 are shown by white letters in black boxes. Patches shared with all variants except PI352504, PI355523, and Westonia are highlighted by grey boxes. A possible recombination breakpoint within these codons is found between nucleotide position 2435 and 2436 of DV92 and IG44878, PI573523, PI289605, PI330550, or PI90945. Cons, consensus sequence (modal value shared by more than 1 residue is indicated by “+” symbol in the display for the simple reason that it is not possible to display multiple characters in a single character space); (dots) nucleotides identical to the consensus; (dashes) gaps in the alignment.

4.2.3 Comparison of amino acid sequences of functional and non-functional *Sr22* variants does not identify a clear region associated with resistance against wheat stem rust

All fourteen variants encode N-terminal CC domains followed by NB and LRR domains. To compare the amino acid sequences of these variants, we aligned them using the multiple sequences alignment tool Clustal Omega and analysed the alignment using Jalview. The total length of the sequence alignment was reduced from 948 consensus residues to 139 polymorphic sites **(Figure 4.3a)**.

We identified thirteen polymorphic sites that correspond to putative solvent-exposed residues of the LRR β-strand/β-turn motif (denoted by x's above the alignment in **(Figure 4.3a)**. Two resistant *T. boeoticum* accessions, IG44855 and IG44857 were shown to possess a near-identical sequence to the *Sr22* reference (from Schomburgk) except for a predicted two amino acid deletion at the N terminal region and five predicted amino acid positions, respectively **(Figure 4.3a)**. A resistant *T. monococcum* accession, PI190945 showed similar resistance responses to *Pgt* races as the other two resistant *T. monococcum* accessions, PI289605 and PI330550 but differed by five predicted amino acid positions **(Figure 4.3a)**.

The correlation importance score between residue and phenotype was calculated based on the score of major allele (score 10), minor allele (score 1), resistance allele (sign plus, +) and susceptible allele (sign minus, -). Based on a correlation importance score, the highest scores were found at amino acid position 94 and 780 with scores of 65 followed by amino acid position 559, 630 and 631 with scores of 56 **(Figure 4.3a)**. The lowest score was found at amino acid positions 384 and 660 with score of -34. A resistance-associated region within the alignment could be identified between amino acid position 498 and 756 for all resistant accessions except for IG44870 and PI352504 **(Figure 4.3a)**. Unexpectedly, these two resistant accessions grouped with the susceptible accessions in the phylogenetic tree **(Figure 4.3b)**.

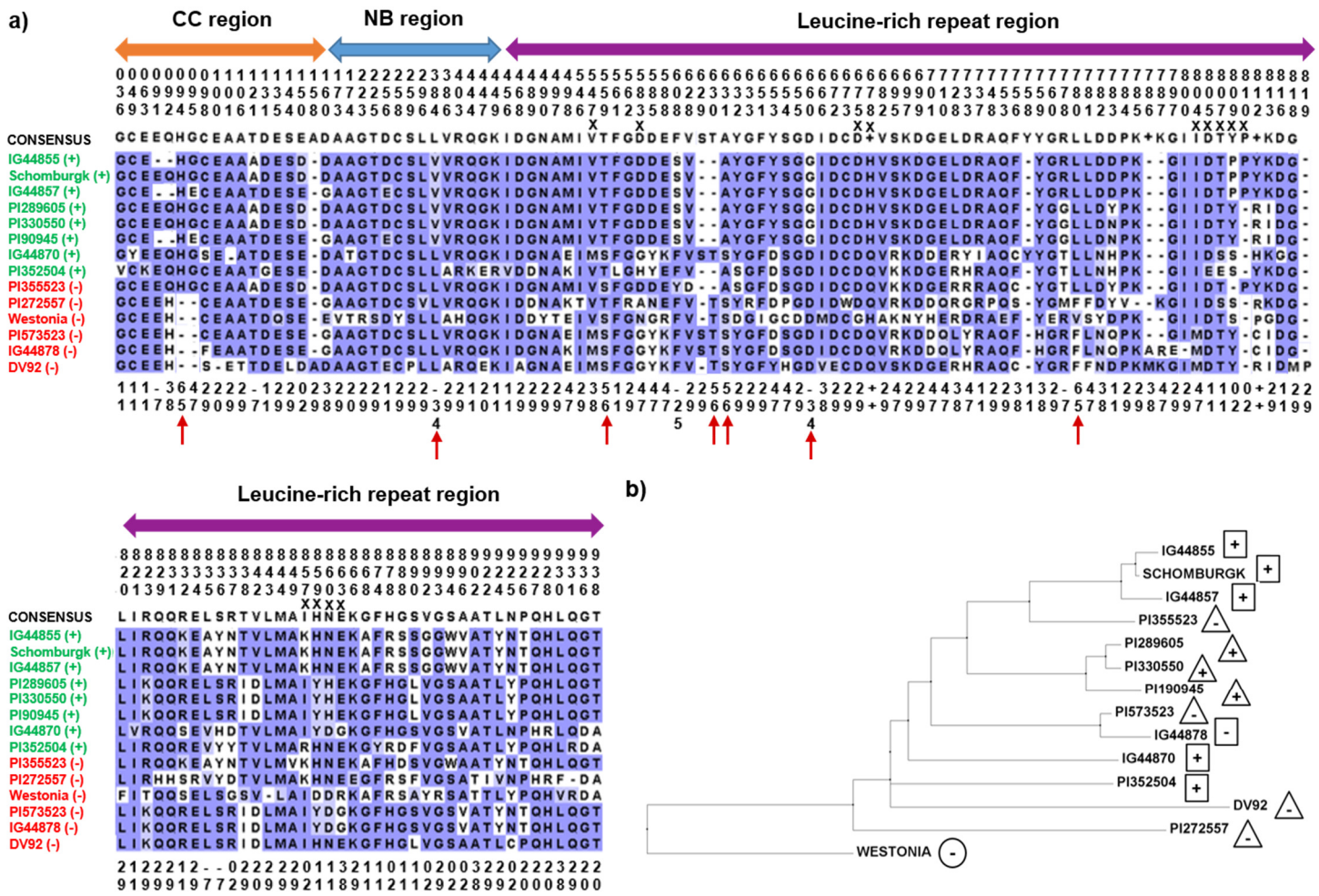


Figure 4.3 Amino acid sequence alignment of *Sr22* variants.

Amino acid sequence alignment of *Sr22* variants. (a) Only polymorphic sites in the amino acid sequences of the fourteen *Sr22* variants are shown in the alignment. This reduces the total length of the sequence alignment from the 948 consensus residues to 139. Their domain structure according to Steuernagel et al. 2016 is indicated in the upper panel. Vertical numbers below the domain structure refer to amino acid positions in the alignment. Amino acids that correspond to putative solvent-exposed residues of the LRR β-strand/β-turn motif are indicated below a boldface 'x'. Consensus sequences are indicated above the alignment (modal value shared by more than 1 residue is indicated by a "+" symbol in the display for the simple reason that it is not possible to display multiple characters in a single character space). Residues that match the consensus sequence residue at that position are highlighted with a dark blue background. Residues that do not match the consensus residue but where the two residues have a positive BLOSUM62 score are highlighted with a light blue background. Gaps are shown with a white background. The *Sr22* variants are indicated on the left with "+" or "-" signs in parenthesis as well as green and red colour to denote the postulated presence or absence of functional *Sr22*, respectively, based on resistance to wheat stem rust (Rouse and Jin, 2011a) (Sambasivam Periyannan and Matt Rouse, unpublished data). Vertical numbers below the alignment refer to an importance score of correlation between the residue and phenotype, in which the higher the value the better the correlation. The dark red arrows show amino acid positions with the scores of 65, 56, and -34. (b) Neighbour-joining tree analysis of amino acid sequence of *Sr22* gene variants identified from diploid and hexaploid wheat **(Table 4.3)**. "+" or "-" signs denote the postulated presence or absence of *Sr22*, respectively, based on resistance to wheat stem rust. Box, triangle or circle symbols indicate a *T. boeoticum*, *T. monococcum* or *T. aestivum* origin, respectively.

4.2.4 Confirmation of postulated *Sr22* variant function in stable wheat transgenics

Previously, a screen of 1,061 *T. monococcum* accessions against race TTKSK (Ug99) and four additional *Pgt* races with widely different virulence profiles, TRTTF, TTTTF, QFCSC, and MCCFC, identified 55 accessions that were resistant to all five races. Fifteen of these accessions were postulated to carry functional alleles of *Sr22* based on comparison of the observed infection type patterns to the expected patterns of known monogenic and digenic lines (Rouse and Jin, 2011b). To confirm the function of three of these postulated *Sr22* alleles, we transformed them into the stem rust susceptible wheat cv. Fielder. To exclude potential variation in phenotype being contributed by differences in regulatory sequences upstream and downstream of the ATG and STOP codons, respectively, we decided to explore the use of a synthetic promoter and terminator. We first synthesised a 9.7 kb sequence in which the original *Sr22* sequence cloned from Schomburgk (Steuernagel et al., 2016) was fused to the *Sr33* promoter and terminator (Periyannan et al., 2013). Ten primary (T_1) transgenic Fielder plants carrying this sequence, PC132, were all found to be resistant to the Australian wheat stem rust race *Pgt* race 98-1,2,3,5, and 6 **(Table 4.2 and Figure 4.4)**. Moreover, the infection type was comparable to the original *Sr22* transgenics in which the transgene was driven by the native promoter (construct PC103; (Steuernagel et al., 2016)) **(Table 4.2 and Figure 4.4)**. We next investigated whether multi-segment Golden Gate assembly (Weber et al., 2011) could be used as an alternative to full-length synthesis to facilitate the rapid, cost-effective generation of constructs with additional *Sr22* alleles regulated by the *Sr33* promoter and terminators since these regulatory sequences contain no occurrence of the Type IIS restriction enzyme *Bsa*I, which must be removed by domestication for Golden Gate assembly. In addition, we domesticated the *Sr22* open reading and terminator sequences by removal of one occurrence of a *Bsa*I site at each sequence. Even though the *Sr22* open reading could be faithfully maintained through domestication, there was a small risk that the removal of these internal *Bsa*I sites, or the addition of non-native Golden Gate linker sequences at the fusion junctions would disrupt function. To test this, we assembled an *Sr33*_promoter::*Sr22*_gene::*Sr33*_terminator synthetic gene (construct PC127) and sent the construct to Mick Ayliffe and his colleagues at CSIRO, Australia for wheat transformation into cv. Fielder. Ten primary transgenics were tested as above and found to be resistant with an infection type similar to the equivalent undomesticated synthetic construct (PC132) **(Table 4.2 and Figure 4.4)**. This result motivated us to generate similar Golden Gate constructs for

the *Sr22* alleles PI90945 and PI573523, which had been postulated to carry functional and non-functional *Sr22* alleles, respectively (**Table 4.3**;(Rouse and Jin, 2011a)). As above, we evaluated ten primary transgenic Fielder plants carrying these constructs. All *Sr33*_promoter::*Sr22*-PI90945_gene::*Sr33*_terminator (construct PC130) plants were found to be resistant, while all *Sr33*_promoter::*Sr22*-PI573523::*Sr33*_terminator (construct PC146) plants were found to be susceptible **(Table 4.2 and Figure 4.4)**, in line with the postulations **(Table 4.3)**.

4.2.5 Fusion of *Sr22* to the maize ubiquitin constitutive promoter has a negligible effect on resistance to wheat stem rust in seedling assays

The expression levels of resistance genes in unchallenged hosts are generally low (MacQueen and Bergelson, 2016). We hypothesised that the intermediate infection type (IT) of transgenics expressing *Sr22*-Schomburgk (IT of 2=) and the susceptible infection type of *Sr22*-PI573523 (IT of 4) **(Table 4.2 and Figure 4.4)** could be reduced (i.e. resistance increased) if these sequences were driven by the maize ubiquitin constitutive promoter. To test this, we generated the following constructs maize_ubiquitin_promoter::*Sr22*-Schomburgk:: Schomburgk_terminator (construct PC126) and maize_ubiquitin::*Sr22*-PI573523::Schomburgk_terminator (construct PC131) **(Table 4.1)**. As above, we generated 10 transgenic lines for each construct and tested these against the *Pgt* race 98-1,2,3,5, and 6. There was no appreciable increase in resistance when *Sr22*-Schomburgk was driven by the maize ubiquitin promoter, whereas *Sr22*-PI57352 under control of the maize ubiquitin promoter reduced the infection type marginally from a 4 to a 3+ **(Table 4.2 and Figure 4.4)**.

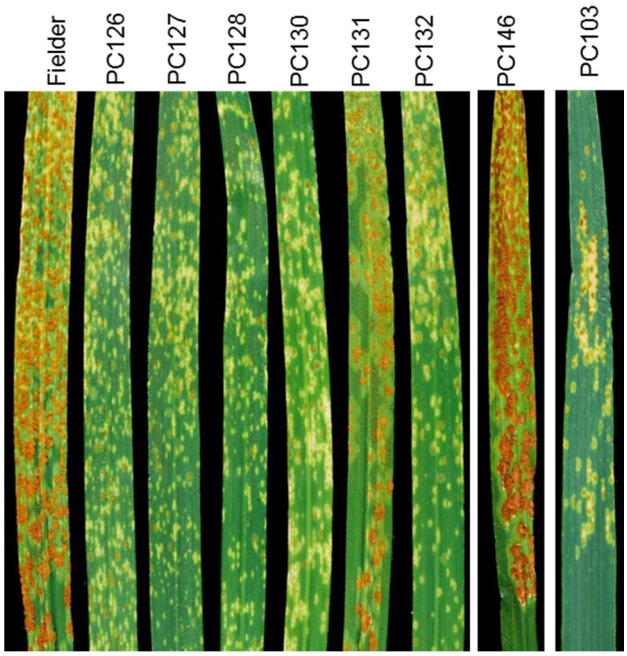


Figure 4.4 Phenotypic reaction of transgenic *Sr22* variants against Australian *Pgt* race 98-1,2,3,5, and 6.

Stem rust infection assays using the Australian *Pgt* race 98-1,2,3,5, and 6 on representative seedlings of primary transgenic lines (T_1) of *Sr22* variants and comparison to the susceptible control cultivar Fielder. PC126 to PC128 and PC132 represent independent transgenic lines carrying the *Sr22* allele Schomburgk driven by different promoter and terminator combinations. PC130 represents a transgenic line carrying the *Sr22* PI190945 gene driven by the *Sr33* promoter and terminator. PC131 and PC146 represent transgenic lines carrying the *Sr22* PI573523 gene driven by different promoter and terminator combinations. PC103 represents transgenic lines carrying the *Sr22* Schomburgk under native regulatory elements (**Table 4.1**).

Table 4.1 List of binary constructs of *Sr22* variants.

Name	ID	Promoter	*Sr22* variant	Terminator
pBW_0001	PC126	Maize ubiquitin	Schomburgk	Sr*22*_Schomburgk
pBW_0002	PC127	*Sr33*	Schomburgk	*Sr33*
pBW_0003	PC128	*Sr33*	Schomburgk	*Sr22*_Schomburgk
pBW_0004	PC130	*Sr33*	PI190945	*Sr33*
pBW_0005	PC131	Maize ubiquitin	PI573523	*Sr22*_Schomburgk
pBW_0006*	PC132	*Sr33*	Schomburgk	*Sr33*
pBW_0007	PC146	*Sr33*	PI573523	*Sr33*
	PC103**	*Sr22*_Schomburgk	Schomburgk	*Sr22*_Schomburgk

*Synthesised as native, undomesticated gene.

**Published by Steuernagel et al., 2016.

Table 4.2 Stem rust resistance (infection type) scores of transgenic Fielder lines with *Sr22* variants assayed with the Australian *Pgt* race 98-1,2,3,5, and 6.

Plant ID	Line	Plant A	Plant B	Plant C	Plant D	Plant E	Plant F	Plant G	Plant H	Plant I	Plant J
PC126	PC126_T1	2=	2=	2=	2=	2=	2=	2=	2=	2=	2=
PC127	PC127_T1	2=	2=	2=	2=	2=	2=	2=	2=	2=	2=
PC128	PC128_T1	2=	2=	2=	2=	2=	2=	2=	2=	2=	2=
PC130	PC130_T1	2=	2=	2=	2=	2=	2=	2=	2=	2=	2=
PC131	PC131_T1	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+
PC132	PC132_T1	2=	2=	2=	2=	2=	2=	2=	2=	2=	2=
PC146	PC146_T1	4	4	4	4	4	4	4	4	4	4
PC103**	PC103_T1	2-	2-	2-	2-	2-	2-	NA*	NA	NA	NA
Fielder	Susceptible check	4	4	4	4	4	4	4	4	4	4

*Not available.

**Published by Steuernagel et al., 2016.

Table 4.3 Stem rust resistance score for selected diploid accessions (*T. boeoticum* and *T. monoccum*) and hexaploid wheats.

Species	Accession	*Sr22* postulation	Infection type							
			Australian *Pgt*			Rouse and Jin 2011				
			17-1,2,3,7=142	98	194	TRTTF	TTKSK	TTTTF	QFCSC	MCCFC
T. monococcum	PI289605	+	NA*	NA	NA	2-	2-	2-	;1	;1
	PI355523	-	NA	NA	NA	3+	3	4	3+	4
	PI330550	+	NA	NA	NA	2	2	2/2+	;1-	;1
	PI190945	+	NA	NA	NA	1;	1	2-;	;1-	;1
	W3534	+	22-	NA	NA	NA	NA	NA	NA	NA
	DV92	-	NA	NA	NA	NA	NA	NA	3+	NA
	P1573523	-	NA	NA	NA	2-	;N	2	;2-	;1-
	PI272557	-	NA	NA	NA	3+	4	4	4	4
	PI362553	+	NA	NA	NA	2-	1	2-	;2=	;1
	PI362554	+	NA	NA	NA	2-	;1	2-	;2-	;1
	PI355522	+	NA	NA	NA	1;	;1	1;/2-	;12-	;1
T. boeoticum	Schomburgk	+	2-	NA	NA	NA	NA	NA	NA	NA
	IG44855	+	2=	NA	NA	NA	NA	NA	NA	NA
	IG44857	+	2=	NA	NA	NA	NA	NA	NA	NA
	IG44921	+	2=	NA	NA	NA	NA	NA	NA	NA
	IG44878	-	3+	NA	NA	NA	NA	NA	NA	NA
	PI352504	+	NA	NA		2	2-	2	;2-	;1
	IG44870	+	2=	1	1	NA	NA	NA	NA	NA
	IG44866	+	2=	1	1-	NA	NA	NA	NA	NA
	IG44868	+	2=	NA	NA	NA	NA	NA	NA	NA
	IG44919	+	2=	2=	NA	NA	NA	NA	NA	NA
T. aestivum	Westonia	-	3+	NA	NA	NA	NA	NA	NA	NA

*Not available.

4.3 Discussion

Disease resistance genes are often clustered in complex loci with related paralogs. The clustered distribution provides a dynamic structure to generate new alleles through intergenic or intragenic recombination. However, analysis of the phylogenetic organisation of the *Sr22* locus in the Chinese Spring wheat genome revealed a single gene at each of the homoeologous loci in chromosomes 7A, 7B and 7D **(Figure 4.1)**. This suggests that the locus has remained as a simple, single gene locus since the separation of the A, B and D lineages 7 million years ago (Marcussen et al., 2014). This physical organisation is similar to the *Rpm1* (Stahl et al., 1999) and *RPP13* (Bittner-Eddy et al., 2000) genes in Arabidopsis but in contrast to most resistance genes which are members of large gene families and organised in complex clusters or paralogs, such as the tomato *Cf-4* and *Cf-9* genes at the Milky Way locus (Parniske et al., 1997), the flax *M* locus (Anderson et al., 1997) and the maize *Rpi* locus (Sun et al., 2001).

The nucleotide sequence alignment of the fourteen alleles indicated that five accessions, which shared >95% sequence identity, contained the greatest variation in the LRR region. Comparison of the nucleotide sequences revealed extensive historical recombination between different alleles based on a patchwork or mosaic pattern of sequence blocks **(Figure 4.2)**. This 'gene shuffling' resulting from multiple, iterative intragenic sequence exchange events, has been proposed to be a major mechanism in generating sequence variation from which novel specificities can be selected during the co-evolution of the host and the pathogen (Parniske et al., 1997). As the phylogenetic organisation analysis suggests that the locus complexity is physically simple (i.e. there is a singleton at the locus), the observed sequence exchange between *Sr22* alleles is more likely due to equal crossing over between allelic variants, rather than unequal crossing over between paralogues (Parniske et al., 1997) or homologues from another chromosome (Parniske and Jones, 1999) as observed in the tomato *Hcr9* genes.

From twenty-two variants that were initially identified, eight are 100% identical. This could reflect that they were derived from genetically identical parents (which were erroneously given different accession numbers) or that there is conservation of sequence at this locus among genetically distinct accessions. The generation of additional genotype data outside of the *Sr22* locus would be required to elucidate this question.

In this study, we compared the sequence of fourteen alleles of *Sr22* with postulated resistance (n = 8) and susceptible (n = 6) phenotypes based on infection type reactions to

diagnostic races of wheat stem rust on the *T. boeoticum* and *T. monococcum* accessions from which the alleles were isolated. We confirmed the function of two of these *Sr22* alleles by generating and phenotyping transgenic wheat plants. We generated and studied an alignment of the fourteen *Sr22* alleles and found three amino acid residues that are enriched in the resistant *Sr22* alleles at the N-terminal region of the LRRs **(Figure 4.3a)**. Amino acid residues in the LRRs have been found to act as major determinants of specificity in other CC-NB-LRR proteins. For example, in the rice blast resistance genes *Pita* and *Pid2*, a single amino acid change at LRR domain and transmembrane (TM) domain, respectively, differentiates resistant and susceptible alleles (Bryan et al., 2000; Chen et al., 2006). It would be interesting therefore to target these residues for mutagenesis to see if this would result in gain-of-resistance in a susceptible allele and *vice versa*.

Based on the sequence alignment, we also observed a general association with resistance and susceptibility between amino acid position 498 and 756, except that *Sr22*-IG44870 and *Sr22*-PI352504 buck the trend **(Figure 4.3a)**. The sequences of these two alleles postulated to be resistant are more closely related to those alleles postulated to not confer resistance **(Figure 4.3a and b) (Table 4.3)**. The generation and phenotyping of transgenics carrying these alleles will provide essential information on their functionality. Alternatively, or in addition, an infection assay of the accessions with a wheat stem rust race lacking *Avr-Sr22* could further confirm the gene postulation.

The majority of the *Sr22* sequence encompassing the CC and NB domains is conserved in contrast to the LRR region, in which most of the variation can be found between alleles. Enhanced sequence variation in the LRR domain is consistent with its proposed role in pathogen recognition specificity, which is believed to explain the higher degree of diversifying selection typically found in this part of the protein (Dodds et al., 2006; Krasileva et al., 2010), in particular within the putative solvent-exposed residues in the LRRs (Dodds et al., 2001; Ellis et al., 1999; Rairdan and Moffett, 2006; Shen et al., 2003; Wulff et al., 2009; Wulff et al., 2001).

We showed that multi-segment Golden Gate assembly (Weber et al., 2011) can be used to rapidly generate constructs in a cost effective manner. We used synthetic versions of *Sr22* promoter and terminator*, Sr33* promoter and terminator*,* and maize ubiquitin promoter to test three allelic variants of *Sr22.* All these synthetic versions of regulatory elements were undomesticated except the *Sr22* terminator. In the future, these modular components can be used for allele mining or functional testing of new candidate stem rust resistance genes

(i.e. identified through map-based approximation or mutational genomics screens such as MutRenSeq (Steuernagel et al., 2016) and MutChromSeq (Sánchez-Martín et al., 2016)).

4.4 Materials and methods

4.4.1 Acquisition and verification of *Sr22* sequence variants

Fourteen *Sr22* sequence variants were identified in Steuernagel et al., 2016 by PCR screening and sequencing of accessions of *T. boeoticum* (and its domesticated form *T. monococcum*) which had been postulated to carry functional or non-functional alleles of *Sr22.* Further screening and sequencing revealed an additional eight sequences from both species (Sambasivam Periyannan and Matt Rouse, unpublished data).

4.4.2 Nucleotide sequence analysis of *Sr22* variants

Nucleotide sequence alignment of *Sr22* variants was performed using Clustal Omega (Sievers et al., 2011) with the default parameters except the order in which the sequences appear in the final alignment was set to "input sequence order". The alignment output was visualised and analysed using Jalview version 2 (Waterhouse et al., 2009). Only polymorphic sites are shown in the alignment in which two or more alternative nucleotides, with each occurring in at least two sequences within the alignment. Nucleotides that correspond to putative solvent-exposed residues of the LRR β-strand/β-turn motif were identified manually by the position of "xxLxLxx".

4.4.3 Amino acid sequence analysis of *Sr22* variants

Amino acid sequence alignment of *Sr22* variants was performed using Clustal Omega (Sievers et al., 2011) using default parameters except the order in which the sequences appear in the final alignment was set to "input sequence order". The alignment output was visualised and analysed using Jalview version 2 (Waterhouse et al., 2009). In addition, some manual adjustments of the alignment were also done. A BLOSUM62 (Blocks Substitution Matrix) score was applied to colour residues within each column in relation to the consensus sequence. The conserved amino acid sequences of the twenty-two *Sr22* variants were discarded, leaving only polymorphic sites shown in the alignment. Amino acids that correspond to putative solvent-exposed residues of the LRR β-strand/β-turn motif were identified manually by the position of "xxLxLxx". The correlation importance score between residue and phenotype was calculated based on the score of major allele (score 10), minor allele (score 1), resistance allele (sign plus, +) and susceptible allele (sign minus, -). A neighbour-joining tree analysis of amino acid sequence of *Sr22* gene variants was created using Jalview version 2 according to the BLOSUM62 calculation (Waterhouse et al., 2009).

The postulated presence or absence of *Sr22* based on resistance to wheat stem rust was according to Rouse and Jin, 2011 and Sambasivam Periyannan and Matt Rouse, unpublished data.

4.4.4 Acquisition and verification of *Sr33* regulatory sequence

Since the 5' region of *Sr33* was not available in GenBank and could not be obtained from the authors of the *Sr33* cloning paper due to IP restrictions, a PCR amplification was performed to obtain the sequence by amplifying an 8 kb length of genomic fragment comprising all of the *Sr33* exons and introns, as well as 2.4 kb upstream of the ATG and 1.5 kb downstream of the STOP codon. *Sr33* wild type (CSID5405) genomic DNA obtained from Evans Lagudah, CSIRO was used as a template. A PCR was performed with a final volume of 50 µL containing 100 ng of genomic DNA, 25 µL of GoTaq Long Range PCR master mix (recombinant *Taq* DNA polymerase and a recombinant proofreading DNA polymerase) (Promega Corp.) and 5 µM of each primer (Sr33P1 and Sr33P2 (Periyannan et al., 2013)). The reaction schedule was; 95 °C for 2 min, 30 cycles of 94 °C for 30 seconds, 65 °C for 8 min and 72 °C for 10 min. The 8 kb *Sr33* gene fragment was cloned into the pCR-XL-TOPO vector (Life Technologies Ltd) and transformed into *E. coli* one-shot Top10 chemically competent cells (Life Technologies Ltd). Eleven colonies were picked for colony PCR to screen for positive clones. This PCR with a final volume of a 20 µL contained 10 µM of each primer (Sr33F6 and Sr33R8), and 10 µL of 2X REDTaq ReadyMix PCR Reaction Mix (Sigma-Aldrich, St Louis, MO, USA). The reaction schedule was; 95 °C for 5 min, 33 cycles of 94 °C for 30 s, 59 °C for 40 s and 72 °C for 1 min. The positive clones from the colony PCR were cultured for DNA extraction using QIAPREP Miniprep Kit (Qiagen Ltd.) following the company's protocol. From restriction analysis using *Eco*RI, four positive clones were sequenced to confirm the clones. From sequencing results, new primers were designed towards the 5' region to obtain additional unknown promoter sequence from each positive clone. This step was repeated few times until a consensus sequence of the *Sr33* 5' upstream region was obtained from these four positive clones. The 3' region of *Sr33* was obtained from GenBank. Both the 3' region and the consensus sequence of the 5' region of *Sr33* were synthesised by a commercial DNA synthesis provider (Life Technologies Ltd). The domesticated version of the *Sr33* promoter was tested for its ability to drive *Sr33* in transgenic barley (Chapter 3).

4.4.5 Generation of binary constructs carrying *Sr22*

The maize ubiquitin and the *Sr33* promoter regions were used for designing *Sr22* chimeric constructs. *Sr22* coding regions (sequence from START to STOP codon, including both exons and introns), from three different sources, including *Sr22*-Schomburgk, PI190945, and

PI573523, were used for the designs. For the simplicity of the construct synthesis, the coding regions were divided into two sections, namely coding region 1 and 2. The *Sr22*-Schomburgk and the *Sr33* terminator regions were used for the designs. All these modular components were designed with an assistance from Guotai Yu and synthesised by a commercial DNA synthesis provider (Life Technologies Ltd). The synthesised components were assembled by Golden Gate cloning (Weber et al., 2011) into the toolkit vector pICH47732, and then transferred to pVecBARII (a derivative of pWBvec8 in which the 35S hygromycin gene has been replaced with a 35S BAR selectable marker gene (Wang et al., 1998) as *Not*I fragments. Prior to synthesis, all native *Bsa*I sites were domesticated (i.e. removed by editing) without changing the predicted amino acid coding sequence within the exons of the coding regions, and without changing the intron splice donor/acceptor sites. The domestication was performed using the program Genious version 7.1.7. All modular components were domesticated and synthesised as Level 0 modules by introducing *Bsa*I sites at the start and at the end of each module. In addition to introduction of flanking *Not*I sites, the sequence AATG was added to link promoter and coding regions and the sequence GCTT was added to link coding regions and terminators. The sequence ACGT was added to link coding region 1 and 2. The sequences GGAG and CGCT were added to link left and right borders of the vector, respectively. The assembled Level 1 module was transferred into the pVecBARII binary vector for wheat transformation. In addition, one whole synthetic gene, which was not domesticated for *Bsa*I sites, and composed of the *Sr33* promoter fused to the *Sr22* coding region fused to the *Sr33* terminator was designed and synthesised. All binary plasmids containing the desired insert were transformed by electroporation into *Agrobacterium tumefaciens* for wheat transformation.

4.4.6 Wheat transformation, stem rust inoculations and phenotypic evaluations

The *Sr22* binary constructs were introduced into wheat cultivar Fielder by Michael Ayliffe and his colleagues at CSIRO, Australia using the *Agrobacterium*-transformation protocol described by (Ishida et al., 2015) and phosphinothricin as a selective agent. Ten independent primary transgenic plants were recovered and grown in an automated growth cabinet with 16 hours light and 8 hours dark conditions and a constant temperature of 23 °C. The plants were inoculated with the Australian *Pgt* race 98-1,2,3,5, and 6, which is virulent on Fielder at the fully developed third leaf stage. After 24 hours of incubation in a closed transparent plastic box under high humidity, the plants were restored to the original growth conditions and observed for rust development. Rust infection types were scored 14 days post inoculation.

5 A method for CRISPR/Cas9-mediated reiterative gene addition

5.1 Introduction

The importance of gene stacking for the production of cultivars expressing multiple traits or genes has increased enormously. One reason is, many important agronomic traits are controlled by complex protein interactions of different genes, requiring integration of multiple loci (Munns and Tester, 2008). Also, plant-microbe interaction studies have revealed that the host and pathogen are under a competitive evolutionary arms race for their continuous survival. Multiple genes are required to manage evolving pathogens in the field including the causative agent of wheat stem rust. This fungal disease caused by *Puccinia graminis* f. sp. *tritici* is a major threat to wheat production in many regions of the world. A single resistance (*R*) gene tends to break down when deployed over a large area where the pathogen is prevalent. In comparison, stacking multiple wheat stem rust resistance (*Sr*) genes against this disease could result in a more durable resistance.

Gene pyramiding can be achieved by conventional crossing of independently generated transgenic plants. However, the dispersal of the transgenes throughout the genome complicates maintenance of the stack in subsequent generations. T-DNA integration at a single transgene locus can be attained by molecular stacking using a binary vector carrying multiple genes (Jo et al., 2014). The logistical challenges of this approach increases with the number of genes to be stacked in addition to vector construction costs, size and time constraints. Therefore, it would be desirable to be able to add transgenes sequentially to the same locus, and indeed be able to subtract individual components of the stack as and when required, so as to provide the maximum level of flexibility.

The advent of genome editing tools such as meganuclease, Zinc Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9) could mediate the site-specific transgene integration required to assemble a transgene stack *in vivo*. Sequence specific nucleases can be designed to introduce a DNA double-strand break (DSB) at a chosen specific chromosomal location and utilise the double-strand break repair mechanism either via homologous recombination (HR) or non-homologous end joining (NHEJ) to promote precise integration of the transgene. In HR, the DSB is repaired using a DNA repair template containing a sequence of the introduced gene with homology to the cut site. In NHEJ, the ends of the

break sites are repaired, most often imperfectly, which introduces small insertions or deletions (indels). The indels occurring within coding sequence often destroy gene function as a result of frame-shift mutations or in-frame deletions (Luo et al., 2016).

An example of engineered nuclease-mediated transgene stacking includes the use of a designer meganuclease to integrate two herbicide tolerance genes, *epsp* and *hppd* into the pre-existing transgene locus of the *cry2Ae* insecticidal gene in cotton (D'Halluin et al., 2013). In another study, a ZFN was engineered to recognise an endogenous target next to a pre-existing herbicide resistance phosphinothricin acetyltransferase (*pat*) transgene and used to introduce a selectable marker gene (*aad1*) in corn (Ainley et al., 2013). These two studies employed an HR-mediated gene stacking, while to our knowledge, no study has reported *in vivo* transgene stacking via the NHEJ pathway.

Gene targeting or gene knock-in using designer nuclease have also been reported in numerous studies, providing a strong incentive to stack multiple genes using this approach. In maize, DSB generated by CRISPR/Cas9 (Svitashev et al., 2015) and ZFN (Shukla et al., 2009) was used to achieve insertion of the herbicide resistance gene phosphinothricin acetyltransferase (*pat*) by the HR repair mechanism. On the other hand, gene knock-ins have also been achieved via the NHEJ pathway including TALEN-mediated gene knock-in of ssDNA encoding His-tag and Myc-tag pepetides in wheat (Wang et al., 2014b) and an intron-mediated targeted gene insertion in the rice endogenous gene *EPSPS* using CRISPR/Cas9 (Li et al., 2016).

Nevertheless, targeted gene integration using designer nucleases is a rare outcome as shown in previous studies in which the frequency of recovering plants with successful integration event was relatively low, ranging from 1.4% to 5% (Ainley et al., 2013; D'Halluin et al., 2013; Li et al., 2016; Svitashev et al., 2015; Wang et al., 2014b). This has proved to be a major bottleneck owing to the requirement of generating and screening a high number of transgenics.

Alternatively, targeted gene integration can also be achieved by site-specific recombination (SSR). In contrast to designer nuclease-mediated transgene integration, this approach is simple and requires less transformation efforts to obtain transgenics that contain precise site-specific integration (Nandy et al., 2015). Site-specific recombinases such as Cre/lox, R/RS and FLP/FRT have been used to mediate integration of transgenes into pre-integrated recombination loci. This has been demonstrated in various plants including Arabidopsis

(Louwerse et al., 2007), maize (Kerbach et al., 2005), rice (Srivastava et al., 2004), soybean (Li et al., 2009), and tobacco (Nanto et al., 2009). However, a major constraint of this approach is that the recombination is reversible as it introduces two recombination sites that can recombine, limiting its application in gene stacking. To improve this approach, a recent study demonstrated the use of an engineered nuclease to excise one of the recombination sites embedded in the marker gene for subsequent recombination as well as generating a marker free SSR-mediated gene integration plant. Cre-*lox* and FLP-*FRT*-mediated gene integration in combination with I-*Sce*I-mediated marker gene excision in rice has extended the application of recombinases for gene stacking (Nandy et al., 2015).

The assembly of DNA fragments into a large and complex plasmid vector can be a tedious and time-consuming step for plant transformation. Traditional restriction digestion and ligation-based cloning method generates only short single-stranded overhangs at specific sites. This limits the applicability for rapid multi DNA fragments assembly. Recently, Golden Gate cloning has been developed, which uses type IIS restriction endonucleases that cleave outside of their recognition sequences (Engler et al., 2008). The incorporation of the Type IIS restriction sites allows generation of user-defined overhangs in which a simultaneous assembly of multiple fragments can be achieved with no existing scar between adjacent assembly fragments. This assembly method has dramatically decreased the amount of time required to design and develop complex DNA molecules. However, one major concern is that the fragments to be assembled must be free from recognition site of the selected type IIS restriction endonuclease. This will require domestication (removal of the sites) of the sequences and may affect gene expression if it involves some essential part of the genome such as untranslated regions and promoters.

In this study, we have attempted to use CRISPR/Cas9 to repair the hygromycin phosphotransferase II (*HPTII*) selectable marker gene as a proof-of-concept to stack multiple wheat stem rust resistance genes. The proposed method involves targeted insertion of the additional genes to a pre-characterised locus, known as a landing pad containing the first gene of interest and two different selection cassettes, one complete and the other as a half version. In the next round of transformation, a repair template containing the second gene of interest and the two selection cassettes in half version is delivered into the landing pad line. The complete selection cassette is replaced with the half version of itself, while the other selection cassette is repaired in allowing selection of the regenerated plants. The half version of the selection cassette can then be used for the subsequent round of stacking **(Figure 5.1)**. This would allow all inserted transgenes to be inserted at the same physical

location in the genome and thus co-segregate and also permit removal of subunits, if necessary. Theoretically, an unlimited number of *Sr* genes can be integrated sequentially at different times at such a locus. In this study, we generated barley T_0 and T_1 transgenics containing an integrated landing pad. Super transformation of immature embryos isolated from landing pad lines did not yield positive transformants. However, we did identify one deletion event out of twenty-four calli derived from immature embryos of T_0 transgenics.

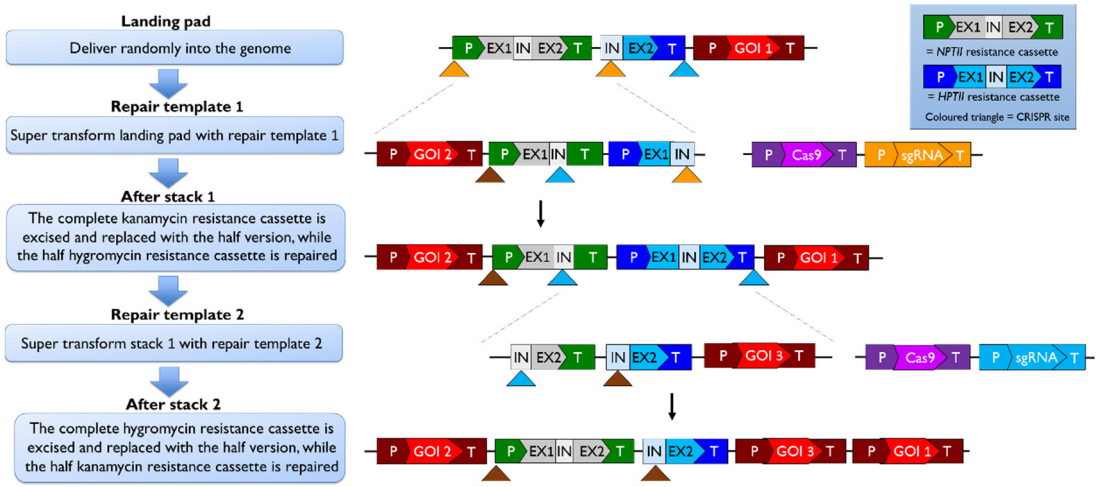


Figure 5.1 Scheme of transgene pyramiding using CRISPR/Cas9-mediated gene addition.

Landing pad containing gene of interest (GOI) 1, a complete kanamycin selectable marker gene, half a hygromycin selectable marker gene, two CRISPR 1 sites and one CRISPR 2 site (orange and blue triangles, respectively) is delivered randomly into the genome. Repair template 1 containing GOI2, half a hygromycin and half a kanamycin selectable marker gene, and one CRISPR 1, 2 and 3 site (orange, blue and brown triangles, respectively) as well as a plasmid containing Cas9 and sgRNA targeting CRISPR 1 sites are delivered during super transformation of transformant containing landing pad. The complete kanamycin selectable marker gene is excised and replaced with the half version of itself, while the half hygromycin selectable marker gene is repaired. The regenerated plants are selected on the repaired hygromycin gene. The half kanamycin selectable marker gene will be used for the subsequent round of stacking. Repair template 2 containing GOI3, half a kanamycin and half a hygromycin selectable marker gene, and one CRISPR 2 and 3 site (blue and brown triangles, respectively) as well as a plasmid containing Cas9 and sgRNA targeting CRISPR 2 sites are delivered during super transformation of transformant containing stack 1. The complete hygromycin selectable marker gene is excised and replaced with the half version of itself, while the half kanamycin selectable marker gene is repaired. The regenerated plants are selected on the repaired kanamycin gene. The half hygromycin selectable marker gene will be used for the subsequent round of stacking. Each CRISPR site (coloured triangle) contains two protospacers.

5.2 Results

5.2.1 Kanamycin as a selectable marker gene in barley transformation

To test functionality of the neomycin phosphotransferase II (*NPTII*) resistance gene under the control of a constitutive maize ubiquitin promoter as a selectable marker gene in barley transformation, we introduced two constructs generated by Mark Smedley, Crop Genetics Department, JIC, into *H. vulgare* cv. Golden Promise; pB214::ZmUbi::*NPTII* and pB214::ZmUbi::*NPTII*-intron **(Supplementary Table 10)**. In addition to the *NPTII* selection cassette, these constructs also contain an *HPTII* selection cassette conferring resistance to the antibiotic hygromycin. We used the antibiotics hygromycin and G418 (*NPTII* confers resistance to this antibiotic) for selection and recovered regenerated shoots from pB214::ZmUbi::*NPTII* but not from pB214::ZmUbi::*NPTII*-intron **(Figure 5.2b and c)**. We extracted DNA from the regenerated shoots and verified the presence of the *HPTII* and *NPTII* transgenes by PCR **(Figure 5.2d and Supplementary Table 11)**.

As excision within an intron is unlikely to disrupt gene function, we decided to insert CRISPR sites within the intron of *NPTII* and use these as target sites for cutting and repairing the selection cassette in the subsequent round of stacking. Therefore, I engineered a construct to test functionality of *NPTII* containing an intron under the control of another constitutive promoter, rice actin, pBW_0036 **(Supplementary Table 10)**. Using antibiotic G418 as a selection, I recovered shoots regenerated from immature embryos of *H. vulgare* cv. Golden Promise inoculated with this construct **(Figure 5.3b)**. The presence of the *NPTII* transgene was verified by diagnostic PCR using genomic DNA extracted from the regenerated shoots as a template. As the *NPTII* gene is derived from bacteria, I also used specific primers of the *vir*D2 gene in the diagnostic PCR to show that the amplification of the *NPTII* in the initial PCR were indeed from the transgene but not likely to come from the *Agrobacterium* **(Figure 5.3c and Supplementary Table 11)**.

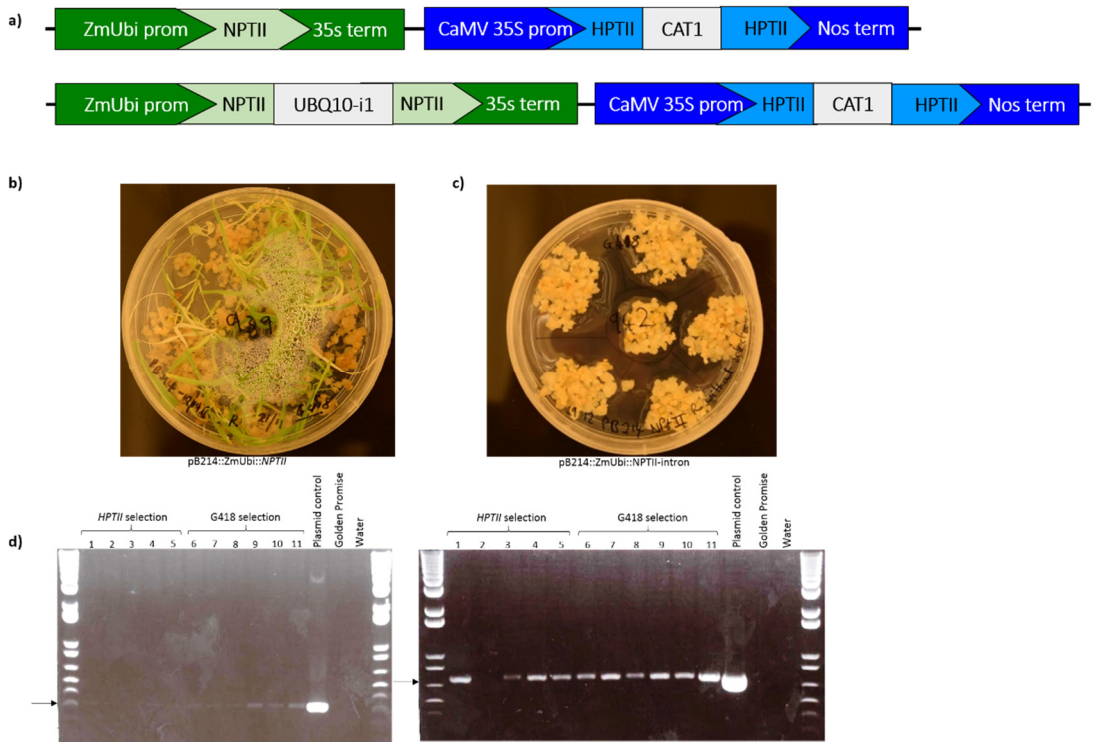


Figure 5.2 Functional testing of kanamycin resistance gene under the control of a maize ubiquitin promoter in barley transformation.

(a) Schematic diagram of construct pB214::ZmUbi::*NPTII* and pB214::ZmUbi::*NPTII*-intron. (b) Regeneration of shoots on regeneration media (after 8 weeks of callus induction) derived from immature embryos of *H. vulgare* cv. Golden Promise transformed with construct pB214::ZmUbi::*NPTII* and selected using the antibiotic G418. (c) Development of callus but no regeneration of shoot on regeneration media derived from immature embryos of *H. vulgare* cv. Golden Promise transformed with construct pB214::ZmUbi::*NPTII*-intron and selected using the antibiotic G418. (d) Left: Amplification products of *HPTII* transgene specific marker (Partial_Hyg_F and Partial_Hyg_R) using genomic DNA (five samples of pB214::ZmUbi::*NPTII* selected using antibiotic hygromycin, six samples of pB214::ZmUbi::*NPTII* selected using antibiotic G418 and Golden Promise) and plasmid DNA (positive control) as a template. The amplification product of 373 bp (arrowed) is present in all samples. Right: Amplification products of *NPTII* transgene specific marker (NPTII_F and NPTII_R) using genomic DNA (five samples of pB214::ZmUbi::*NPTII* selected using antibiotic hygromycin, six samples of pB214::ZmUbi::*NPTII* selected using antibiotic G418 and Golden Promise) and plasmid DNA (positive control) as a template. The amplification product of ~700 bp (arrowed) is present in all except sample 2.

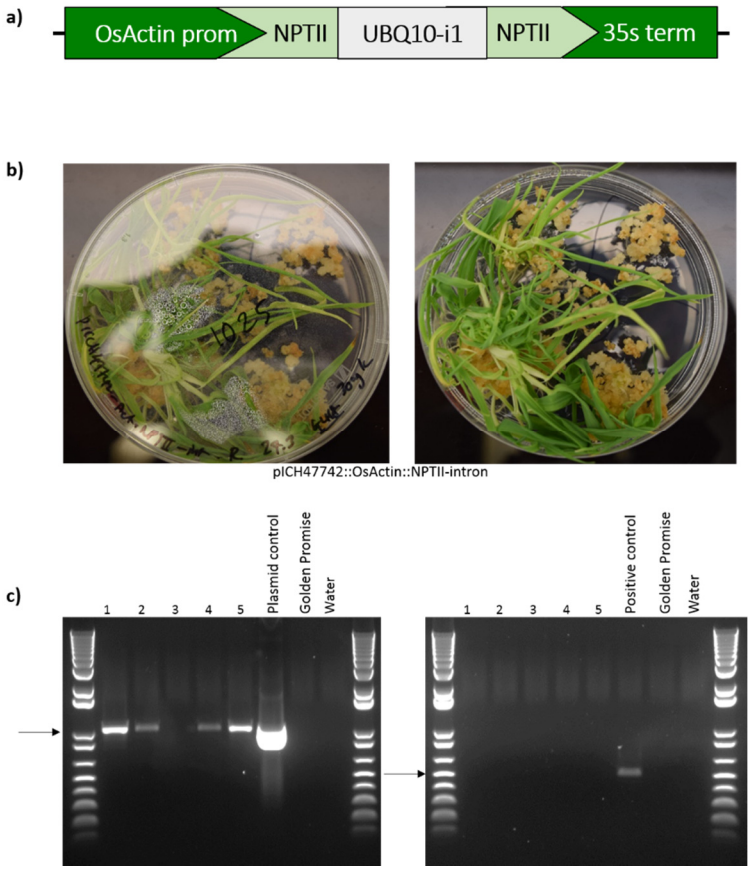


Figure 5.3 Functional testing of kanamycin resistance gene under the control of a rice actin promoter in barley transformation.

(a) Schematic diagram of construct pBW_0036 (pCH47742::OsActin::NPTII-intron). (b) Regeneration of shoots on regeneration media (after 8 weeks of callus induction) derived from immature embryos of *H. vulgare* cv. Golden Promise inoculated with construct pBW_0036 and selected using antibiotic G418. (c) Left: Amplification products of *NPTII* transgene specific marker (NPTII_F and NPTII_R) using genomic DNA (five samples and Golden Promise) and plasmid DNA (positive control) as a template. The amplification product of 1004 bp (arrowed) is present in all except sample 3. Right: Amplification products of *vir*D2 gene specific marker (VIRD2 F and VIRD2 R) using genomic DNA (five samples and Golden Promise) and small aliquot of agrobacterium culture (positive control) as a template. The amplification product of ~487 bp (arrowed) is only present in positive control.

5.2.2 Transformation of landing pad into barley and characterisation of primary transgenics

Using the Golden Gate cloning method (Engler et al., 2014), I engineered a landing pad construct, pBW_0041 **(Supplementary Table 10)** containing a complete *NPTII* resistance gene, a half *HPTII* resistance gene, two CRISPR sites (one at the 5' of the *NPTII* gene and one within the *HPTII* gene intron), and *Sr33*. I introduced this construct into 700 immature embryos of *H. vulgare* cv. Golden Promise by *Agrobacterium*-mediated transformation and recovered 189 transformed plants following a protocol described in (Harwood, 2014a)

except the positive transformants were selected on antibiotic G418. Based on qPCR results, 112 of the transformants contained multiple copy insertions (2 to 25 copies), 50 contained single copy insertions and 27 were nulls containing no PCR-detectable insertion **(Supplementary Table 12)**.

5.2.3 Screening for repair activity of the *HPTII* gene in barley protoplasts

To investigate whether I could transiently repair the *HPTII* gene, I isolated protoplast from the T_2 landing pad line and transfected this with the PCR fragment of repair template amplified from plasmid DNA (pBW_0057) **(Figure 5.4a and Supplementary Table 10)**. I also co-transfected the PCR fragment with Cas9 plasmid DNA (pBW_0140) **(Supplementary Table 10)**. As a positive control, I included YFP construct (pBW_0143) **(Supplementary Table 10)** and examined the YFP fluorescence under the microscope 18 hours post transfection **(Figure 5.4b)**. As negative controls, I transfected the protoplasts with repair template only and Cas9 only. To screen for editing events, I extracted DNA from the transfected protoplast 18 hours post transfection and ran a diagnostic PCR using the *HPTII* transgene specific marker **(Figure 5.4a and c and Supplementary Table 11)**. The amplification product was only present in the plasmid control indicating that the repair event was unsuccessful.

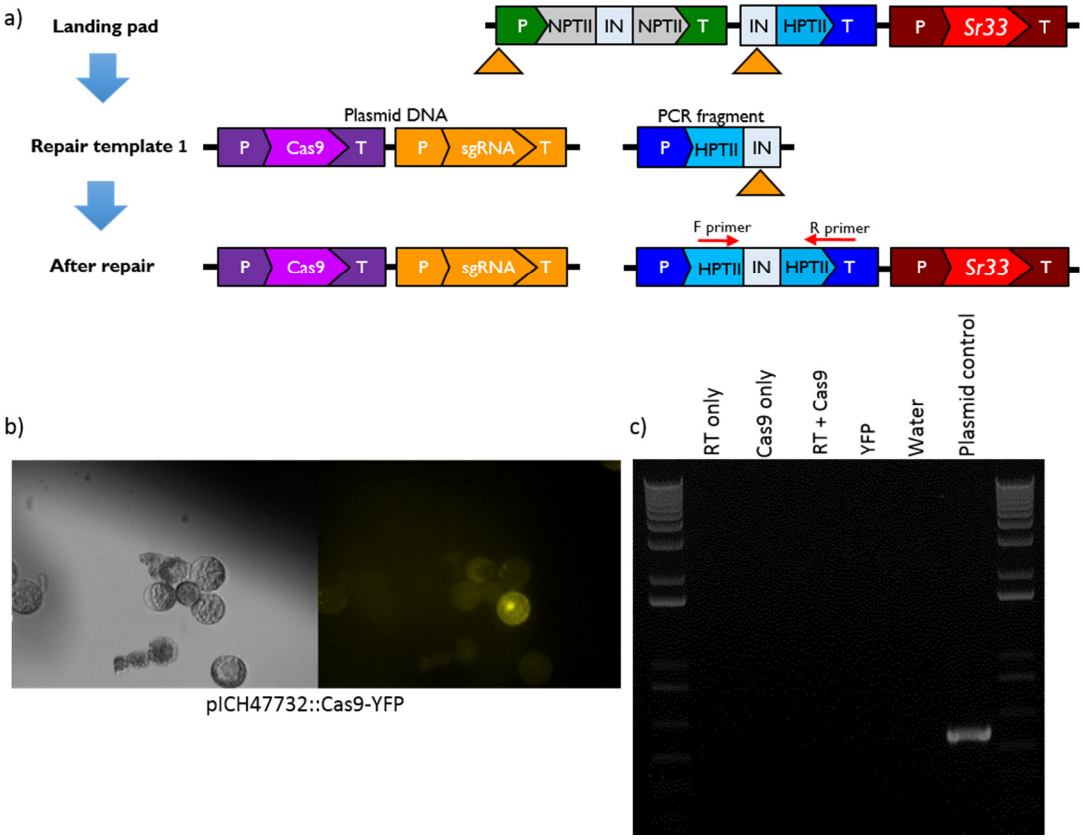


Figure 5.4 Transient assay to detect repaired *HPTII* gene events using barley protoplasts isolated from a T_2 landing pad line.

(a) Super transformation scheme to repair half version of the *HPTII* gene. The landing pad in T_2 line is excised using CRISPR/Cas9-mediated excision. PCR fragment of repair template amplified from plasmid DNA (pBW_0057) is co-transfected with Cas9 plasmid DNA (pBW_0140) into protoplasts isolated from a T_2 landing pad line. After the repair event, the half *HPTII* gene is re-constituted with another half in the landing pad. Arrows indicate primers used to amplify the repaired *HPTII* gene. (b) Microscopy image of protoplast transfected with control construct, pBW_0143 (right: bright field, left: YFP filter). (c) Amplification product of *HPTII* transgene specific marker (Partial_Hyg_F and Partial_Hyg_R) using DNA extracted from protoplast (protoplast transfected with repair template only, protoplast transfected with Cas9 only, protoplast co-transfected with repair template and Cas9, and protoplast transfected with YFP) and plasmid DNA (positive control) as a template. The amplification product of 373 bp (arrowed) is present only in the plasmid control.

5.2.4 *Agrobacterium*-mediated super transformation of landing pad with repair template

To achieve a stable repair event at an early stage of the project, I isolated immature embryos directly from T_0 landing pad lines and inoculated with the repair template construct pBW_0057. I inoculated 625 and 275 immature embryos from lines containing single and multiple copy insertion of the landing pad, respectively **(Table 5.1)**. In comparison with immature embryos derived from *H. vulgare* cv. Golden Promise and T_0 landing pad line inoculated with pB214::ZmUbi::*NPTII*, I obtained no regenerated shoots from super transformation of landing pad line **(Figure 5.5)**. In the next generation, I isolated 1375 immature embryos from single copy, hemizygous T_1 landing pad lines and inoculated with another version of the repair template construct, pBW_0139 **(Supplementary Table 10)**. This improved version of the repair template construct containing only sgRNA1 was used as we speculated that the undetected repair events might be due to the deletion of the repaired *HPTII* gene by both sgRNA1 and sgRNA2 **(Supplementary Figure 8)**. As in the previous experiment, I recovered no shoots in comparison with immature embryos derived from *H. vulgare* cv. Golden Promise inoculated with pB214::ZmUbi::*NPTII*.

Table 5.1 Number of immature embryos derived from T_0 and T_1 landing pad lines inoculated with repair template construct.

Construct	No. of immature embryos		
	T_0 plant[a]		T_1 plant[b]
	Single copy	Multiple copies	Single copy
pBW_0057[c]	625	275	0
pBW_0139[d]	0	0	1,375

[a] 75 and 125 immature embryos derived from *H. vulgare* cv. Golden Promise and T_0 landing pad line, respectively were inoculated with pB214::ZmUbi::*NPTII* as positive controls.
[b] 50 immature embryos derived from *H. vulgare* cv. Golden Promise were inoculated with pB214::ZmUbi::*NPTII* as a positive control.
[c] Repair template with sgRNA1 and sgRNA2.
[d] Repair template with sgRNA1.

5.2.5 PCR screening for deletion and repair event of *HPTII* gene

After super transforming the T_0 landing pad line with repair template construct (pBW_0057) (1309), I recovered no regenerated shoots in comparison to the T_0 landing pad line transformed with control construct (pB214::ZmUbi::*NPTII*-intron) (1308) **(Figure 5.5)**. However, some of the calli of 1309 varied in size, from small to large **(Figure 5.5a)**. I extracted DNA from 24 representative calli of 1307, 1309, and 1310 and ran a diagnostic PCR targeting the *HPTII* transgene specific marker and obtained no positive band in comparison with controls 1308 and 1311 **(Figure 5.5c and Supplementary Table 11)**.

To screen for deletion events, I used another pair of primers and ran a diagnostic PCR **(Figure 5.6a and b and Supplementary Table 11)**. One of the 24 calli appeared to carry a deletion event (1310-6) **(Figure 5.6c)**. Unexpectedly, I also obtained a similar band in the plasmid control. Since there are two CRISPR sites at two different positions, the observed band in the plasmid control could be due to amplification of the forward primer at 5' of *NPTII* and reverse primer at 3' of *HPTII.* PCR and Sanger sequencing of 1310-6 indeed revealed that the *NPTII* gene was deleted **(Figure 5.6d)**.

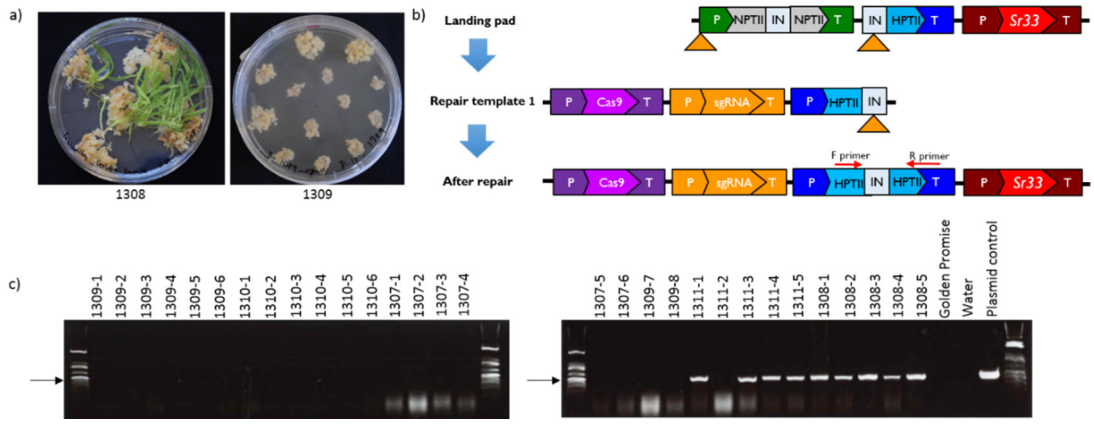


Figure 5.5 Screening of *HPTII* repair events on callus.

(a) Regeneration of shoots on regeneration media (after 8 weeks of callus induction) derived from immature embryos of T_0 landing pad line inoculated with construct pB214::*HPTII*::*NPTII*-intron (1308) and selected using antibiotic hygromycin. Development of callus but no regeneration of shoot on transition media (after 6 weeks of callus induction) derived from immature embryos of T_0 landing pad line inoculated with repair template construct, pBW_0057 (1309) and selected using antibiotic hygromycin. (b) Super transformation scheme to repair half version of *HPTII* gene. The landing pad in the T_0 line is excised using CRISPR/Cas9-mediated excision. Plasmid DNA containing repair template and Cas9 cassette is delivered by *Agrobacterium*-mediated transformation into immature embryo isolated from T_0 line containing landing pad. After repair event, half *HPTII* gene is re-constituted with another half in the landing pad. Arrows indicate primers used to amplify repaired *HPTII* gene. (c) Amplification product of *HPTII* transgene specific marker (Hyg_int_span_For and Hyg_int_span_Rev) using DNA extracted from callus (24 calli generated from immature embryo of T_0 landing pad line transformed with repair template (1307, 1309, 1310), 5 calli generated from immature embryo of T_0 landing pad line transformed with control construct pB214::*HPTII*::*NPTII*-intron (1308), five calli generated from immature embryo of Golden Promise transformed with control construct pB214::*HPTII*::*NPTII*-intron (1311)), from genomic DNA (Golden Promise) and from plasmid DNA (positive control) as a template. The amplification product of 588 bp (arrowed) is present only in 1308, 1311 (controls) and plasmid control.

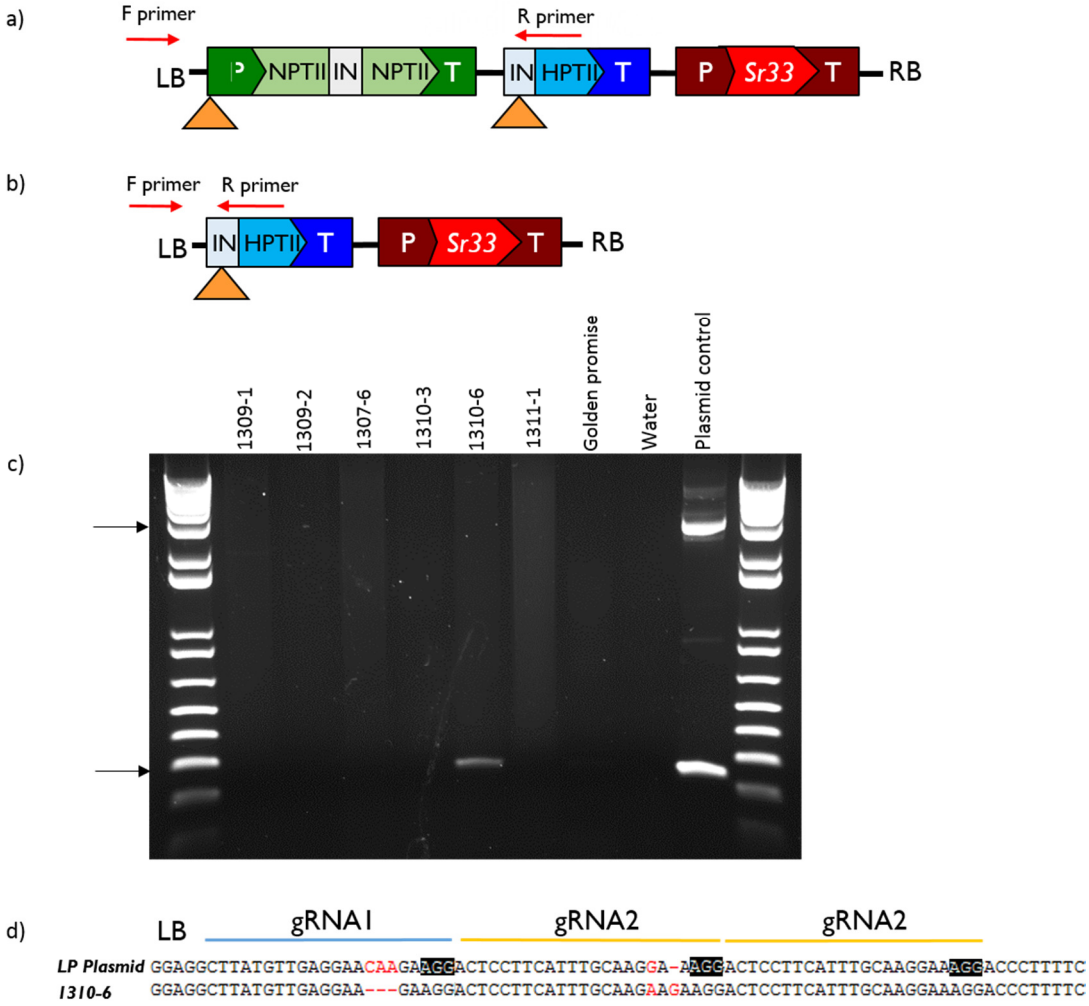


Figure 5.6 Screening for deletion events in callus.

Arrows indicate primer pairs used to screen for (a) non-deletion event and (b) deletion event in callus generated from barley immature embryo of T_0 landing pad line transformed with repair template construct, pBW_0057. Each CRISPR site (orange triangle) contains two protospacers should one of them not be functional. (c) Amplification product of deletion event diagnostic marker (F_del_LP and R_del_LP) using DNA extracted from callus (six representative samples and Golden Promise) and plasmid DNA (positive control) as a template. The amplification product of 3,051 bp for non-deletion event is only present in the plasmid control and 192 bp for deletion event is present in sample 1310-6 and the plasmid control. (d) Sanger sequencing results of PCR amplicon from callus 1310-6 carrying the deletion event. The second row in the DNA alignment shows sequences of 1310-6 in comparison with sequences of the landing pad (LP) plasmid DNA. Highlighted sequences are PAM motifs (NGG) and red sequences are deletion sites (three base pair upstream of the PAM motif). Deletion events are indicated by insertion/deletions (indels) (red sequence) and missing parts of the *NPTII* sequences.

5.3 Discussion and conclusion

We have proposed a strategy to stack multiple *Sr* genes in barley based on marker exchange of two selectable marker genes in each stacking step **(Figure 5.1)**. Hence, identifying efficient and less laborious selectable marker genes for barley transformation is crucial. In addition, selectable marker genes containing an intron within the coding sequence are also desired as (NHEJ)-mediated repair often results in frame-shift mutations or in-frame deletions (Luo et al., 2016). The bialaphos resistance (*Bar*) gene conferring resistance to the glufosinate group of herbicides (bialaphos or phosphinothricin, PPT) and hygromycin phosphotransferase II (*HPTII*) gene conferring resistance to the antibiotic hygromycin are among selection systems that have been used for barley transformation. The latter is the favoured selectable marker as 'escape' plants are extremely rare and the selection protocol is less laborious (Harwood, 2014a). The *HPTII* gene containing a castor bean catalase-1 (CAT-1) gene intron was selected as it is an efficient selectable marker gene in barley in addition to offering a 2.5-fold increase in expression (Wang et al., 1997). The neomycin phosphotransferase II (*NPTII*) gene conferring resistance to the antibiotic kanamycin and some related aminoglycosides (G418 and paromomycin) has been used in *Agrobacterium*-mediated wheat transformation (Binka et al., 2012; Cheng et al., 1997). To our knowledge, no report has been made of using the *NPTII* gene in barley transformation. *NPTII* with the UBIQ10 intron is available in the BRACT lab and we therefore first tested this gene under the control of a maize ubiquitin promoter. Immature embryos of barley cv. Golden Promise inoculated with a construct containing *NPTII* with an intron did not result in any regenerated shoots in contrast to *NPTII* without an intron **(Figure 5.2b and c)**. We hypothesised that the strong constitutive promoter coupled with the presence of the UBIQ10 intron in the selection cassette may have impaired the ability to regenerate shoots from transformed calli. To test this hypothesis, we generated another *NPTII* containing intron construct under the control of a slightly weaker constitutive promoter, rice actin, and inoculated immature embryos of barley cv. Golden Promise (unpublished data). After 8 weeks of callus induction, we recovered shoots which then confirmed positive by PCR for presence of the transgene **(Figure 5.3b and c)**. Based on this result, we used *NPTII* containing the UBIQ10 intron and *HPTII* containing the CAT-1 intron as a selectable marker gene for landing pad transformation and super transformation with repair template, respectively.

We managed to achieve 27% transformation efficiency using the *NPTII* gene containing the UBIQ10 intron as a selectable marker gene in the landing pad transformation of barley cv. Golden Promise with 26% of the primary transformants containing single-copy insertion

integrated into the genome. One of the limiting factors to achieve gene knock-in or gene targeting is delivering a sufficient amount of repair template (Svitashev et al., 2015). For this, particle bombardment is often used as a method in delivering a high amount of repair template as described in previous studies (Ainley et al., 2013; D'Halluin et al., 2013; Li et al., 2016; Svitashev et al., 2015). However, *Agrobacterium*-mediated transformation has largely replaced particle bombardment as the method of transformation in many labs, including the JIC, due to its higher transformation efficiency and because it is less prone to generating complex events, i.e. containing multiple tandemly repeated insertions. We therefore decided to use *Agrobacterium*-mediated transformation to deliver our repair template into the landing pad line. Another limiting factor that could influence the efficiency of editing activity is the size of the repair template. As a proof of concept, we generated a repair template without the second selectable marker gene, *NPTII* and second *Sr* gene. This substantially reduced the size of the T-DNA from ~20 kb to ~9 kb **(Supplementary Figure 6)**.

To test the possibility of achieving integration events at an early stage of this project, we isolated immature embryos directly from single-copy T_0 landing pad lines and inoculated them with the repair template construct, pBW_0057 **(Supplementary Table 10)**. It is important to note that the chance of repairing the *HPTII* gene of the progeny of primary transformants (T_1) is reduced by 25%, since 50% will be hemizygous for the T-DNA, 25% will be homozygous for the T-DNA, and 25% of the segregants will be null **(Supplementary Figure 7)**. In addition to super transforming single-copy lines, we also super transformed multi-copy lines so as to have more template and thereby increase the chance of repairing the *HPTII* gene. However, out of 900 immature embryos derived from both single and multi-copy lines, none regenerated into shoots despite the observation that some of the developed calli appeared to be large in size **(Table 5.1)**. PCR-based screening of some of these large calli for repair events did not result in a positive band **(Figure 5.5c)**. However, we did detect one deletion event in one of the calli based on the diagnostic PCR and sequencing **(Figure 5.6c and d)**, suggesting that both sgRNAs are functional.

We generated the landing pad construct with two CRISPR sites, each containing two protospacers (an Arabidopsis and endogenous protospacer with slight modifications), in case one of them would not be functional. These two protospacers were also included in the repair template design. However, in a later discussion, we identified a possible problem with deploying a pair of different protospacers at the two different positions. The presence of two different combinations of protospacers might result in excision of the *NPTII* gene and repair of the *HPTII* gene by the first combination. This may then be followed by deletion of the

repaired *HPTII* gene by the second combination **(Supplementary Figure 8)**. Since we knew from the first experiment that both protospacers are functional based on the detected deletion event **(Figure 5.6c and d)**, we generated another version of the repair template containing only the sgRNA1 expression cassette and delivered this into immature embryos of single-copy, hemizygous T_1 landing pad lines. Despite super transforming 1375 immature embryos, no regenerated shoots were obtained.

In the beginning of this project, a barley transient assay had not been established at JIC. With the help of Oleg Raitskin (a postdoc in Nicola Patron's lab, Earlham Institute), we set up a protoplast-based transient assay for barley. This transient system was used to transiently express the repair template in protoplasts isolated from T_2 landing pad lines followed by PCR-screening for a repair event **(Figure 5.4)**. In comparison with plasmid control, we obtained no amplification product from the repair template co-transfected with Cas9 plasmid, suggesting that repair events were unsuccessful or below the detectable limit. It is important to note that although we observed YFP fluorescence in the control (as a control for transfection efficiency), this might not reflect the overall transformation efficiency as a proper measurement was not conducted (the percentage of YFP-expressing cells was not counted). Based on these results, it is crucial to test the feasibility of obtaining editing events using the transient assay prior to stable transformation. In addition, transient assays are also an ideal way to test the functionality of protospacers and measure their cutting efficiency. By incorporating all this information, we could then estimate the number of immature embryos required for recovering editing events in the stable transformation.

As mentioned earlier, a low copy number of repair template delivered into the target site could be a limiting factor. This has indeed been shown in maize in which *Agrobacterium*-mediated delivery did not recover integration events in comparison to biolistic transformation (Svitashev et al., 2015). Therefore, the next logical step would be to bombard the immature embryos isolated from a landing pad line with massive amounts of repair template DNA to increase the chance of obtaining editing events.

Besides that, a viral replicon could also be utilised to increase the amount of repair template. Gene targeting using CRISPR/Cas9 and TALEN to insert a *35S* promoter upstream of a gene controlling anthocyanin biosynthesis was previously successfully used to obtain accumulation of purple pigments in tomato tissue. This study demonstrates that an increased amount of delivered DNA was achieved when nuclease constructs were encoded on a modified bean yellow dwarf virus genome with 10% efficiency (Čermák et al., 2015). A

similar approach was also applied in wheat in which CRISPR/Cas9-mediated gene targeting of an endogenous ubiquitin locus resulted in a 12-fold increase in editing events when using a modified wheat dwarf virus (WDV) in comparison to non-viral delivery methods (Gil-Humanes et al., 2017). However, this method might be efficient for replication of small genes only as the replication efficiency decreases with the size of the replicated gene (Gil-Humanes et al., 2017), whereby *Sr* genes, which are typically ~8 kb in size (Periyannan et al., 2013; Saintenac et al., 2013; Steuernagel et al., 2016), would be too large.

5.4 Materials and methods

5.4.1 Generation of kanamycin selection cassette

pB214::ZmUbi::*NPTII* and pB214::ZmUbi::*NPTII*-intron were generated by Mark Smedley from BRACT using pBRACT vector **(Supplementary Table 10)**. The *NPTII*–intron gene under the control of rice actin promoter (pBW_0036) was generated by amplifying domesticated sequence using the primer For Kan_Int_Frag1, Rev Kan_Int_Frag1, For Kan_Int_Frag2, and Rev Kan_Int_Frag2 **(Supplementary Table 11)** and native *NPTII* plasmid as a template. Prior to PCR, a *Bpi*I site within the UBIQ10 intron was domesticated (removed by editing) without changing the intron splice donor/acceptor sites using the program Vector NTI. A random bp was introduced at the start of each primer sequence followed by (i) a *Bpi*I site, (ii) 2 bp to bridge the recognition site and cut site, (iii) 4 bp overhangs defined by the standard part being made, and (iv) 23 bp that anneal with the sequence of the *NPTII* gene. The two amplified PCR products were purified with QIAGEN PCR Purification Kit (Qiagen Ltd.) following the manufacturer's protocol. The purified PCR fragments were cloned into the level 0 acceptor vector pICH41308 by Golden Gate cloning (Engler et al., 2014; Weber et al., 2011). The *NPTII*-intron level 0 module was assembled with rice actin (SynBio #363), 35S terminator (pICH41414), and level 1 position 2 acceptor vector (pICH47742). All toolkit vectors and Golden Gate modules were obtained from SynBio TSL.

5.4.2 Generation of landing pad construct

Two protospacers (an Arabidopsis and endogenous protospacer with slight modifications) were used in two CRISPR sites within the landing pad. The Arabidopsis protospacer has been successfully targeted in previous work at TSL and the presence of possible off-site targets have been checked in barley genomes prior to construct design. The first CRISPR site was synthesised as a level 1 position 1 module by a commercial DNA synthesis provider (Life Technologies Ltd).

The two protospacer sequences were added into the middle of the CAT-1 intron of the *HPTII* gene. In addition, *Bpi*I sites were added at the start and at the end of the sequence followed by 2 bp to bridge the recognition site and cut site, and 4 bp overhangs defined by the standard part being made. The modified *HPTII*-intron was synthesised by a commercial DNA synthesis provider (Life Technologies Ltd). An inverse PCR was performed using the primers For_hyg_int and Rev_hyg_int **(Supplementary Table 11)** and the level 0 *HPTII*-intron module as a template. This was done using Phusion DNA polymerase (NEB) following the manufacturer's instructions. Prior to PCR, a random bp was introduced at the start of each primer sequence followed by a *Bpi*I site, 2 bp to bridge the recognition site and cut site, 4 bp overhangs defined by the standard part being made, and 27 bp that anneal with the sequence of *HPTII*. The linear amplicon, including the *HPTII* exons and the backbone of the plasmid, was digested with the methylation sensitive enzyme *Dpn*I to cut out the circular plasmid template. The digested product was run on gel and gel-purified using the Qiagen Gel Purification Kit (Qiagen Ltd) following the manufacturer's protocol. The purified amplicon was used in a *BpiI*-mediated Golden Gate reaction with the synthetic version of the CAT-1 intron synthesised earlier to insert the modified intron into the *HPTII* gene. The *HPTII*-intron level 0 module was assembled with the 35S promoter (pICH51288), the Nos terminator (pICH41421), and the level 1 acceptor vector (pICH47732). The half version of the *HPTII* gene was amplified using the primers For_partial_hyg_int and Rev_partial_hyg_int, and the *HPTII* level 1 module as a template **(Supplementary Table 11)**. The primers were designed by adding random bp at the start of each primer sequence followed by *Bsa*I sites, one bp to bridge the recognition site and cut site, 4 bp overhangs defined by the standard part being made, and 24 bp that anneal with the *HPTII* sequence. The purified PCR amplicon of the half-version of the *HPTII* gene was cloned into the level 1 position 3 acceptor vector (pICH47751).

The domesticated synthetic gene of a 7,854 bp *Sr33*-containing fragment including 2,381 bp of 5' and 1,405 bp of 3' native regulatory regions was synthesised as a level 1 position 4 module by introducing *Bpi*I sites upstream and downstream of the respective 5' and 3' ATG and STOP codons. A level 2 reaction was performed to assemble the synthetic protospacers of CRISPR site 1, the *NPTII* selection cassette (pBW_0036), the half version of the *HPTII* selection cassette (pBW_0038), and the synthetic *Sr33* gene (pBW_0032) into the level 2 acceptor pAGM4723. All toolkit vectors and Golden Gate modules were obtained from SynBio TSL.

5.4.3 Generation of repair template construct

A level 1 Cas9 expression cassette was generated by assembling the maize ubiquitin promoter (pICSL120009), Cas9 (pICSL90004), Octopine synthase terminator (pICH41432), and level 1 position 1 vector (pICH47732). The two protospacers were integrated into a double stranded DNA molecule ready for assembly with a U6 promoter in a level 1 Golden Gate reaction using a 5' tailed oligonucleotide tailed primer to amplify the sgRNA from an existing sgRNA containing plasmid (pICSL70001) **(Supplementary Table 11)**. The primers of gRNA1_F, gRNA1_R, and gRNA2_F were designed as follows:

tgtggtctca CTTG NNNN NNNNN NNNNN NNNNN ***gttttagagctagaaatagcaag***

(The *Bsa*I recognition site is in blue; the four base pair overhang produced by digestion with *Bsa*I is in underlined capitals – this fuses to the last four base pairs of the U6 promoter in plasmid pICSL70001; the 20 bp target sequence is in red; the portion of the oligonucleotide that anneals to the sgRNA template is in bold italics)

This was done using Phusion DNA polymerase (NEB) following the manufacturer's instructions. Amplicons were verified by agarose gel electrophoresis from which they were subsequently cut and purified using the Qiagen Gel Purification Kit (Qiagen Ltd) following the company's protocol. The purified DNA of sgRNA1 was used in a Level 1 assembly reaction with U6 promoter (pICSL90003) and level 1 positon 2 acceptor vector pICH47742. The purified DNA of sgRNA2 was used in a Level 1 assembly reaction with the U6 promoter (pICSL90003) and level 1 positon 3 acceptor vector pICH47751.

The half version of the *HPTII* gene was amplified using primer For_partial_hyg_int_5' and 2Rev_partial_hyg_int_5'_NGG and *HPTII* level 1 module as a template **(Supplementary Table 11)**. The primers were designed by adding random bp at the start of each primer sequence followed by a *Bsa*I site, one bp to bridge the recognition site and cut site, 4 bp overhangs defined by the standard part being made, and 25 bp that anneal with the sequence of *HPTII*. The purified PCR amplicon of the half-version of the *HPTII* gene was cloned into the level 1 position 5 acceptor vector (pICH47772).

A level 2 reaction was performed to assemble the Cas9 expression cassette (pBW_0044), the sgRNA 1 and 2 expression cassette (pBW_0045 and pBW_0046, respectively), the position 4 dummy (pICH54044), and the half-version of the *HPTII* selection cassette (pBW_0048) into a level 2 acceptor pAGM4723. Another version of the repair template containing only the sgRNA1 expression cassette was generated by assembling the Cas9 expression cassette

(pBW_0044), the sgRNA 1 cassette (pBW_0045), the position 3 dummy (pICH54033), the position 4 dummy (pICH54044), and the half-version of the *HPTII* selection cassette (pBW_0048) into the level 2 acceptor pAGM4723. All toolkit vectors and Golden Gate modules were obtained from SynBio TSL.

5.4.4 Golden Gate cloning

Level 0 assembly reactions contained 100 ng of CDS1 acceptor plasmid-pICH41308 as well as PCR amplicon such that inserts to be included in the acceptor backbone were at a 2:1 molar ratio to the acceptor. 0.5 µl of 10 units/µl *Bpi*I (5 units, NEB), 1.5 µl Bovine Serum Albumin (10x), 0.5 µl of 400 units/µl T4 DNA ligase (200 units, NEB), 1.5 µl T4 DNA ligase buffer (10x, NEB) and reaction volumes were made up to 15 µL using sterile distilled water. The reactions were incubated in a thermocycler as follows: 37 °C for 20 seconds, 26 cycles of 37 °C for 3 min/16 °C for 4 min followed by 50 °C for 5 min and 80 °C for 5 min and then stored at 16 °C.

Level 1 assembly reactions contained 100 ng of level 1 acceptor plasmid as well as level 0 plasmids or sgRNA amplicon such that inserts to be included in the acceptor backbone were at a 2:1 molar ratio to the acceptor. 0.5 µl of 10 units/µl *Bsa*I (5 units, NEB), 1.5 µl Bovine Serum Albumin (10x), 0.5 µl of 400 units/µl T4 DNA ligase (200 units, NEB), 1.5 µl T4 DNA ligase buffer (10x, NEB) and reaction volumes were made up to 15 µL using sterile distilled water. Reactions were incubated in a thermocycler as follows: 37 °C for 20 seconds, 26 cycles of 37 °C for 3 min/16 °C for 4 min followed by 50 °C for 5 min and 80 °C for 5 min and then stored at 16 °C.

Level 2 assembly reactions contained 100 ng of level 2 acceptor plasmid as well as level 1 plasmids such that inserts to be included in the acceptor backbone were at a 2:1 molar ratio to the acceptor. 0.5 µl of 10 units/µl *Bpi*I (5 units, NEB), 1.5 µl Bovine Serum Albumin (10x), 0.5 µl of 400 units/µl T4 DNA ligase (200 units, NEB), 1.5 µl T4 DNA ligase buffer (10x, NEB) and reaction volumes were made up to 15 µL using sterile distilled water. Reactions were incubated in a thermocycler as follows: 37 °C for 20 seconds, 26 cycles of 37 °C for 3 min/16 °C for 4 min followed by 50 °C for 5 min and 80 °C for 5 min and stored at 16 °C.

5.4.5 *E. coli* transformation

5 µl of ligation reaction was used for transformation of LIBRARY EFF DH5 COMPETENT CELLS (Life Technology Ltd.) following the company's protocol. 20-80 µl of transformed DH5 was spread onto a plate containing the appropriate antibiotic and incubated at 37 ˚C overnight. White colonies were picked the next day and cultured in LB media containing the appropriate antibiotic. Seven ml of the culture was pelleted and the plasmid was extracted with the

QIAPREP SPIN MINIPREP KIT (Qiagen Ltd.) following the company's protocol. The fidelity of the clone was confirmed by restriction digest analysis and Sanger sequencing.

5.4.6 *Agrobacterium* transformation

0.5 µl (100 ng) of DNA was mixed with 500 µl of *Agrobacterium* strain AGLI electro competent cells. The mixture was placed on ice for a few minutes before transferring into a pre-chilled cuvette. The moisture was removed from the outsides of the cuvette and the cuvette was subjected to electroporation with the following settings: 400 Ω resistance, 25 µFD, and 2.5 mV. 100 µl of LB medium was added into the cuvette and slowly mixed by pipetting up and down. The mixed cells were then transferred into 200 of LB medium. The cells were incubated for 4 to 6 hours at room temperature at 100 rpm. 20-80 µl of transformed AGL1 was spread onto a plate containing the appropriate antibiotic and incubated at 28 ˚C for two days. White colonies were picked after two days and cultured in LB media containing the appropriate antibiotic. The culture was incubated at 28 ˚C for two days. Standard inoculum was prepared by mixing 2 ml culture with 2 ml sterile 30% glycerol and aliquot into 800 µl.

5.4.7 Barley transformation

Agrobacterium-mediated transformation of immature embryos of the barley cv. Golden Promise and landing pad transgenic line was performed and transformed plants were recovered following the protocol described in (Harwood, 2014a). Constructs **(Supplementary Table 10)** containing selection cassettes conferring resistance to the antibiotic hygromycin were introduced.

5.4.8 Isolation of protoplast and transfection

10 ml of enzyme solution was prepared consisting of 0.3% macerozyme, 1.5% cellulose, 0.4 M mannitol, 20 mM KCl, 20 mM MES pH 5.6, 10 mM $CaCl_2$, and 0.1% BSA. The enzyme solution was warmed at 50 ˚C for 5- 10 minutes until the enzymes were completely dissolved. Leaf material of 2 to 3 cm lengths were harvested and placed into a sterile petri dish. 10 mL of enzyme solution was added into a plate and the leaf material was slowly chopped vertically into 5 cm segments using a sharp scalpel blade in the presence of the leaf digest mix (to prevent cell damage resulting from the leaf material drying out). The plate was wrapped with foil and incubated for 2 hours in controlled growth chamber at 22 ˚C. After two hours, 2 ml of W5 buffer consisting of 154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, and 2 mM MES pH 5.6 was added and the digested protoplast solution was filtered through a 75 µm filter into a 50 ml plastic tube. The tube was centrifuged for 2 min at 100 rcf and the supernatant was quickly discarded. The pellet was resuspended in 2 ml of W5 buffer and left

on ice for 30 minutes until the protoplasts precipitated by gravity. The supernatant was carefully removed and the protoplasts were resuspended to a concentration of $2*10^5$ in MMg buffer consisting of 0.4 M mannitol, 15 mM $MgCl_2$, and 4 mM MES pH 5.6.

10-50 µg of DNA was gently mixed with 100 µl of protoplast solution in a 2 ml tube. 100 + x µl of PEG20 solution (x µl of volume of the DNA added) was slowly added to the wall of the tube and incubated on the bench at room temperature for 5 minutes. 1 ml of W5 solution was added and mixed followed by centrifugation at 0.8 rpm for 2 minutes. The supernatant was pipetted out very carefully and resuspended in 300 µl of W5 buffer in a multi-well plate pre-incubated with 1% BSA. The multi-well plate was incubated overnight in Growth Control Chamber with 22 ˚C, light period of 18 hours with intensity not exceeding 75 to 100 µmol m^{-2} s^{-1}, and a dark period of 6 hours.

5.4.9 Copy number analysis

Copy number analysis was outsourced to iDNA genetics, Norwich Research Park.

5.4.10 DNA extraction from leaf tissue and protoplast

Approximately six inches of leaf samples from 3-week-old plants (or callus) were collected in a 2 ml Eppendorf centrifuge tube with two tungsten beads. The samples were freeze-dried for 48 hours and then ground up using a QIAGEN Microtube Homogenizer at 29 cycle/s for 2 min in one orientation and 2 min in another orientation. The samples were ground until the tissue turned into a green-whitish powder. The powder was spin down at 2000 rpm. 800 µL pre-warmed extraction buffer was added [0.02% CTAB, 1.4 M NaCl, 0.1 M Tris pH 8.0, 0.02 M EDTA, 0.02% β-mercaptoethanol] to the sample and mixed well by shaking at a speed of 1600 rpm for 1 min. The suspension was spun down at 2000 rpm. The tubes were placed in a water bath of 65 °C for 30 min, and carefully shaken for 10 s by hand for every 10 min. The tubes were spun at 2000 rpm for 2 min. Then, in the hood, 800 µL of chloroform: isoamyl alcohol (24:1) was added and the caps were put on tightly. The tubes were put between tightening-boards and thoroughly mixed by hand shaking. The samples were centrifuged for 15 min at 12,000 rpm to separate the phases. 45 µL of NaOAc was added to each tube. 450 µL of the top phase was slowly transferred with a disposable 1000 µL pipet (without stirring the chloroform layer). 900 µL of absolute % ethanol was added down the sides of each tube. The caps were put on and the plates were put in the refrigerator for 1 hour. Then, while holding the tube flat (sideways), the samples were carefully mixed for 1 min. The DNA globs were kept and the solution was removed using a pipet. 500 µL of 1000 µg/ml of RNase was added to 50 ml of TE. 500 µL of this TE + RNase was added to the tube. The tube was inverted

to dislodge the DNA and gently re-suspended until no DNA was visible. The tubes were put into an incubator at 37 °C to allow RNA digestion for at least 1 hour. Then, 1000 µL of absolute ethanol was added to wash the DNA for 1 hour or overnight. After that, DNA globs were kept and the solution was carefully removed. 1000 µL of 70% ethanol was added to wash the DNA and the solution was removed. The DNA was left to dry for 20 min. 100 µL of 1X TE buffer was added to the tube and the DNA was quantified using a Nanodrop. The DNA samples were stored at -20 °C or -80 °C.

DNA extraction of protoplast was performed by pelleting the protoplasts by centrifugation and removing the supernatant. 550 µl of extraction buffer consisting of 0.2 M Tris-HCl pH 7.5, 0.05 M EDTA, 2 M NaCl, 2 % CTAB and pH adjusted to 7.44 was added to the protoplast. The tube was incubated for 60 minutes at 65 ˚C in a water bath. 250 µl of chloroform was added and each sample was vigorously vortexed for 5 minutes (in a fume hood). The phases were separated by centrifugation at 20,000 rcf (max speed) for 5 to 10 minutes. 250 µl supernatant was added to a new tube filled with 250 µL of isopropanol (1:1 ratio with amount of supernatant) and mixed by inversion for 1 minute. The tube was incubated at -80 ˚C overnight. The DNA was pelleted by centrifuging the samples for 30 minutes at 20,000 rcf, 4 ˚C. The isopropanol was carefully poured off and the pellet was washed with 500 µl of 70% ethanol. The tube was centrifuged for 10 minutes at 20,000 rcf. The ethanol was discarded and the tube was left in the fume hood for at least 60 minutes until the pellet dried. The DNA was dissolved in 20 µl of TE buffer (+0.5 µg RNAse).

6 Conclusions

The wheat stem rust resistance gene *Sr44* confers resistance to the Ug99 race complex (Liu et al., 2013b). We attempted to clone this gene using MutRenSeq and to this end, we generated an EMS-mutant population of the *Sr44* introgression line IK1019 x Angas. From the suppressor screens, we identified twelve independent susceptible mutants in which ten of the mutants and the wild-type *Sr44* were subjected to NLR capture and sequencing. However, sequence analysis did not reveal a clear candidate. We speculate that this may be due to residual presence/absence heterogeneity in the M_0 seed batch used for mutagenesis. Therefore in future cloning-by-sequencing experiments, it is crucial to ensure that the initial starting material used for mutagenesis is as homogenous as possible. We also hypothesise that the gene was not efficiently captured and assembled. Thus, we are currently testing a new NLR library (V3) which carries introns and additional NLR source sequences (including *Sr44* transcriptome data) on wild-type *Sr44* and the mutants to clone the gene.

To investigate whether the mutations occur within the *Sr44* gene itself, or at a second, independent gene required for the resistance phenotype, we inter-crossed six of the susceptible mutants and phenotyped the F_1 mutants with *Pgt* race TTKSK. A network showing the relationship and possible complementation groups of the mutants indicates the presence of three complementation groups. However, the interpretation was complicated by the observation that there was some degree of non-complementation between two of the mutants but not supported by a clear complementation of the other mutants. We also do not rule out the possibility that all the mutants may carry lesions in *Sr44*, and that the resistant phenotype that we observed in some F_1 mutant intercross progeny results from the heterozygous combination of alleles with mutations in different parts of the gene. This heterozygous combination can result from intra- and inter-molecular NB-LRR interactions which may have obscured the complementation cross analysis by giving rise to allelic complementation. In contrast, we cannot also exclude the possibility of hetero-dimerisation between a mutated *Sr44* gene products with another mutated NB-LRR in an F_1 testcross. This can result in a non-activate resistance signalling complex or a reduced level of downstream signalling, which in turn gives rise to false non-complementation suggesting that the mutants are in the same gene, whereas in fact they are in different genes. To shed more light on the nature of the *Sr44* mutants, we are planning to generate testcrosses between the mutants and the susceptible recurrent parent Angas and screen the F_1 with stem rust.

To investigate meristem cell fate in hexaploid wheat, we phenotyped the sister spikes of those primary spikes from which we obtained *Sr44* mutants in the disease resistance suppressor screen. The majority of the tested sister spikes were found to be resistant suggesting that the tillers are indeed genetically distinct from the primary spike. This information can be used in the generation of mutant populations for suppressor screens by harvesting and screening of multiple individual spikes from the same plant to generate a more effective M_2 population size. This would potentially reduce the cost, space and effort needed to identify enough susceptible mutants for cloning-by-sequencing.

To improve the immunity of barley against wheat stem rust, we have transferred the cloned wheat *Sr* genes into barley by transformation. The transgenic lines carrying *Sr22*, *Sr33,* and *Sr45* expressed high-level resistance to *Pgt* indicating that wheat *Sr* genes are functional in barley. As more *Sr* genes are cloned in wheat, we could easily test those genes in barley using this strategy and open up a much-needed avenue for engineering genetic disease control by deploying multiple *Sr* genes in a stack to provide broad-spectrum resistance and reduce the risk of resistance breakdown. The next step would be to assess the agronomic value of the generated barley transgenic lines expressing *Sr* transgenes in the field.

Based on the phylogenetic organisation in the Chinese Spring wheat genome, we show that the *Sr22* locus is a simple, single gene locus. We observe some evidences of historical sequence exchange in the LRR region of some of the alleles through nucleotide alignment and sequence analysis. We also generated wheat transgenics and phenotyped the plants to confirm the gene postulation of two of the previously identified *Sr22* alleles. From a comparison of the postulated function with their predicted amino acid sequences, we did not identify a region associated with functional resistance against wheat stem rust.

Functional stacking of multiple *Sr* genes simultaneously at a single transgene locus conferring broad resistance to worldwide *Pgt* isolates could result in more durable resistance to this disease. We attempted to use CRISPR/Cas9 to repair the hygromycin phosphotransferase II (*HPTII*) gene as a proof-of-concept to stack multiple *Sr* genes. Super transformation of barley T_0 and T_1 transgenics containing an integrated landing pad with repair template DNA did not yield positive transformants. However, we identified one deletion event out of twenty-four calli derived from immature embryos of T_0 transgenics. For the next attempt, it is crucial to test the feasibility of obtaining editing events using the transient assay prior to stable transformation. The transient assay could also be used to test the functionality of protospacers and measure their cutting efficiency. One of the limiting factors to achieve gene

knock-in or gene targeting is delivering a sufficient amount of repair template (Svitashev et al., 2015). Therefore, the next logical step would be to bombard the immature embryos isolated from transgenics containing landing pad with large amounts of repair template DNA to increase the chance of obtaining editing events.

With the advent of different novel approaches for rapid gene cloning in hexaploid wheat, more wheat rust resistance genes are expected to be cloned in the next few years. As the number of cloned genes increased, there is continuing discussion on how the multigenes stack will look like. How many genes we want to put in a stack? What are the genes to combine? Would it be consists of race-specific or non-race specific genes or combination of both? Could we include other resistance genes (i.e leaf rust and stripe rust) in the stack?

The number of gene in the stack would partially depend on the ability to transfer large GM cassettes into the wheat genome as the transformation efficiency is inversely proportional to the size of the T-DNA (Park et al., 2000). Apart from our attempt to use CRISPR/Cas9, there are other strategies currently being tested to achieve this goal such as site-specific recombination approaches. Further improvement in the engineering of large and complex constructs combining with an improved wheat transformation efficiencies would put ourselves in a position where we can stack as many genes as we wish. Recently, *Agrobacterium*-mediated transformation efficiency of the spring wheat variety Fielder has been improved significantly with efficiency of 40 to 90% (Ishida et al., 2015). Application of this technique in other commercial wheat varieties will certainly attract local farmers to deploy their favourite varieties containing the stacked R genes.

Adult plant resistance (APR) genes are assumed to be non-race specific and confer partial resistance but durable and additive in their effect while seedling resistance (R) genes appear to be race-specific and confer complete resistance but less durable due to changes in pathogen *Avr* genes. Therefore, the most promising stacking strategy is to combine both types of genes to achieve a diverse level of resistance and specificity and to minimise the possibility of pathogen virulence evolution.

Once a GM wheat has been developed, it is crucial to functionally test individual genes in a stack for any suppressive interaction between transgenes and also to distinguish their resistance specificities. In the case of APR genes, biochemical assays may be required, while for race-specific seedling R genes, cloning of the corresponding effectors and developing a robust and rapid assays will be important for these tests.

To maintain the durability of the GM wheat, it is also crucial to ensure that individual resistance gene within the stack are not deployed singularly. One way to prevent deployment of single genes is by protecting the newly cloned genes with intellectual property protection.

7 References

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M., Innan, H., Cano, L., Kamoun, S. and Terauchi, R. (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* **30**, 174-178.

Acevedo-García, J., Spencer, D., Thieron, H., Reinstadler, A., Hammond-Kosack, K., Phillips, A.L. and Panstruga, R. (2017) *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant biotechnology journal* **15**, 367-378.

Ade, J., DeYoung, B.J., Golstein, C. and Innes, R.W. (2007) Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2531-2536.

Ain, Q.-u., Rasheed, A., Anwar, A., Mahmood, T., Mahmood, T., Imtiaz, M., He, Z., Xia, X. and Quraishi, U. (2015) Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Frontiers in Plant Science* **6**.

Ainley, W.M., Sastry-Dent, L., Welter, M.E., Murray, M.G., Zeitler, B., Amora, R., Corbin, D.R., Miles, R.R., Arnold, N.L., Strange, T.L., Simpson, M.A., Cao, Z., Carroll, C., Pawelczak, K.S., Blue, R., West, K., Rowland, L.M., Perkins, D., Samuel, P., Dewes, C.M., Shen, L., Sriram, S., Evans, S.L., Rebar, E.J., Zhang, L., Gregory, P.D., Urnov, F.D., Webb, S.R. and Petolino, J.F. (2013) Trait stacking via targeted genome editing. *Plant biotechnology journal* **11**, 1126-1134.

Al-Kaff, N., Knight, E., Bertin, I., Foote, T., Hart, N., Griffiths, S. and Moore, G. (2008) Detailed Dissection of the Chromosomal Region Containing the *Ph1* Locus in Wheat *Triticum aestivum*: With Deletion Mutants and Expression Profiling. *Annals of Botany* **101**, 863-872.

Allen, R.S., Nakasugi, K., Doran, R.L., Millar, A.A. and Waterhouse, P.M. (2013) Facile mutant identification via a single parental backcross method and application of whole genome sequencing based mapping pipelines. *Frontiers in Plant Science* **4**, 362.

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

Aoun, M., Breiland, M., Kathryn Turner, M., Loladze, A., Chao, S., Xu, S.S., Ammar, K., Anderson, J.A., Kolmer, J.A. and Acevedo, M. (2016) Genome-Wide Association Mapping of Leaf Rust Response in a Durum Wheat Worldwide Germplasm Collection. *The plant genome* **9**.

Arora, D., Gross, T. and Brueggeman, R. (2013) Allele characterization of genes required for *rpg4*-mediated wheat stem rust resistance identifies *Rpg5* as the *R* gene. *Phytopathology* **103**, 1153-1161.

Arruda, M.P., Brown, P., Brown-Guedira, G., Krill, A.M., Thurber, C., Merrill, K.R., Foresman, B.J. and Kolb, F.L. (2016) Genome-Wide Association Mapping of Fusarium Head Blight Resistance in Wheat using Genotyping-by-Sequencing. *The plant genome* **9**.

Austin, R.S., Vidaurre, D., Stamatiou, G., Breit, R., Provart, N.J., Bonetta, D., Zhang, J., Fung, P., Gong, Y., Wang, P.W., McCourt, P. and Guttman, D.S. (2011) Next-generation mapping of Arabidopsis genes. *The Plant journal : for cell and molecular biology* **67**, 715-725.

Autrique, E., Tanksley, S.D., Sorrells, M.E. and Singh, R.P. (1995) Molecular markers for four leaf rust resistance genes introgressed into wheat from wild relatives. *Genome* **38**, 75-83.

Ayliffe, M., Singh, R. and Lagudah, E. (2008) Durable resistance to wheat stem rust needed. *Current opinion in plant biology* **11**, 187-192.

Bai, J., Pennill, L.A., Ning, J., Lee, S.W., Ramalingam, J., Webb, C.A., Zhao, B., Sun, Q., Nelson, J.C., Leach, J.E. and Hulbert, S.H. (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. *Genome research* **12**, 1871-1884.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D. and Dangl, J.L. (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS pathogens* **7**, e1002132.

Bariana, H.S., Brown, G.N., Bansal, U.K., Miah, H., Standen, G.E. and Lu, M. (2007) Breeding triple rust resistant wheat cultivars for Australia using conventional and marker-assisted selection technologies. *Australian Journal of Agricultural Research* **58**, 576-587.

Bernardo, A.N., Ma, H., Zhang, D. and Bai, G. (2012) Single nucleotide polymorphism in wheat chromosome region harboring Fhb1 for Fusarium head blight resistance. *Molecular Breeding* **29**, 477-488.

Bhattacharya, S. (2017) Deadly new wheat disease threatens Europe's crops. *Nature* **542**, 145-146.

Binka, A., Orczyk, W. and Nadolska-Orczyk, A. (2012) The *Agrobacterium*-mediated transformation of common wheat (*Triticum aestivum L.*) and triticale (*x Triticosecale Wittmack*): role of the binary vector system and selection cassettes. *Journal of applied genetics* **53**, 1-8.

Bittner-Eddy, P.D., Crute, I.R., Holub, E.B. and Beynon, J.L. (2000) *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *The Plant journal : for cell and molecular biology* **21**, 177-188.

Blomstedt, C.K., Gleadow, R.M., O'Donnell, N., Naur, P., Jensen, K., Laursen, T., Olsen, C.E., Stuart, P., Hamill, J.D., Moller, B.L. and Neale, A.D. (2012) A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor L. Moench* resulting in acyanogenic forage production. *Plant biotechnology journal* **10**, 54-66.

Bos, J.I., Prince, D., Pitino, M., Maffei, M.E., Win, J. and Hogenhout, S.A. (2010) A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). *PLoS genetics* **6**, e1001216.

Broertjes, C. and Harten, A.M.v. (1978) *Application of mutation breeding methods in the improvement of vegetatively propagated crops* Amsterdam, The Netherlands:Elsevier Scientific Publishing Company.

Brueggeman, R., Druka, A., Nirmala, J., Cavileer, T., Drader, T., Rostoks, N., Mirlohi, A., Bennypaul, H., Gill, U., Kudrna, D., Whitelaw, C., Kilian, A., Han, F., Sun, Y., Gill, K., Steffenson, B. and Kleinhofs, A. (2008) The stem rust resistance gene *Rpg5* encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14970-14975.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B. and 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* **99**, 9328-9333.

Bryan, G.T., Wu, K.S., Farrall, L., Jia, Y., Hershey, H.P., McAdams, S.A., Faulk, K.N., Donaldson, G.K., Tarchini, R. and Valent, B. (2000) tA single amino acid difference

distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. *The Plant cell* **12**, 2033-2046.

Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. and Waugh, R. (2004) A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare L.*). *The Plant journal : for cell and molecular biology* **40**, 143-150.

Cao, A., Xing, L., Wang, X., Yang, X., Wang, W., Sun, Y., Qian, C., Ni, J., Chen, Y., Liu, D., Wang, X. and Chen, P. (2011) Serine/threonine kinase gene *Stpk-V*, a key member of powdery mildew resistance gene *Pm21,* confers powdery mildew resistance in wheat. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7727-7732.

Capal, P., Blavet, N., Vrana, J., Kubalakova, M. and Doležel, J. (2015) Multiple displacement amplification of the DNA from single flow-sorted plant chromosome. *The Plant journal : for cell and molecular biology* **84**, 838-844.

Cauderon, Y., Saigne, B. and Dauge, M. (1973) The resistance to wheat rusts of *Agropyron intermedium* and its use in wheat improvement. In: *Proceedings of the 4th International Wheat Genetics Symposium* pp. 401-407. Columbia, MO.

Cavanagh, C.R., Chao, S., Wang, S., Huang, B.E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G.L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., da Silva, M.L., Bockelman, H., Talbert, L., Anderson, J.A., Dreisigacker, S., Baenziger, S., Carter, A., Korzun, V., Morrell, P.L., Dubcovsky, J., Morell, M.K., Sorrells, M.E., Hayden, M.J. and Akhunov, E. (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 8057-8062.

Čermák, T., Baltes, N.J., Čegan, R., Zhang, Y. and Voytas, D.F. (2015) High-frequency, precise modification of the tomato genome. *Genome Biology* **16**, 232.

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

Cesari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., Shimamoto, K., Dodds, P., Terauchi, R. and Kroj, T. (2014b) The NB-LRR proteins *RGA4* and *RGA5* interact functionally and physically to confer disease resistance. *The EMBO journal* **33**, 1941-1959.

Chen, A., Li, C., Hu, W., Lau, M.Y., Lin, H., Rockwell, N.C., Martin, S.S., Jernstedt, J.A., Lagarias, J.C. and Dubcovsky, J. (2014) Phytochrome C plays a major role in the acceleration of wheat flowering under long-day photoperiod. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 10037-10044.

Chen, X., Shang, J., Chen, D., Lei, C., Zou, Y., Zhai, W., Liu, G., Xu, J., Ling, Z., Cao, G., Ma, B., Wang, Y., Zhao, X., Li, S. and Zhu, L. (2006) A B-lectin receptor kinase gene conferring rice blast resistance. *The Plant journal : for cell and molecular biology* **46**, 794-804.

Chen, X.M., Luo, Y.H., Xia, X.C., Xia, L.Q., Chen, X., Ren, Z.L., He, Z.H. and Jia, J.Z. (2005) Chromosomal location of powdery mildew resistance gene Pm16 in wheat using SSR marker analysis. *Plant Breeding* **124**, 225-228.

Chen, Y., Liu, Z. and Halterman, D.A. (2012) Molecular determinants of resistance activation and suppression by *Phytophthora infestans* effector IPI-O. *PLoS pathogens* **8**, e1002595.

Cheng, M., Fry, J.E., Pang, S., Zhou, H., Hironaka, C.M., Duncan, D.R., Conner, T.W. and Wan, Y. (1997) Genetic Transformation of Wheat Mediated by *Agrobacterium tumefaciens*. *Plant physiology* **115**, 971-980.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L., Sourdille, P., Couloux, A., Paux, E., Leroy, P., Mangenot, S., Guilhot, N., Le Gouis, J., Balfourier, F., Alaux, M., Jamilloux, V., Poulain, J., Durand, C., Bellec, A., Gaspin, C., Safar, J., Doležel, J., Rogers, J., Vandepoele, K., Aury, J.M., Mayer, K., Berges, H., Quesneville, H., Wincker, P. and Feuillet, C. (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**, 1249721.

Chung, E.H., da Cunha, L., Wu, A.J., Gao, Z., Cherkis, K., Afzal, A.J., Mackey, D. and Dangl, J.L. (2011) Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. *Cell host & microbe* **9**, 125-136.

Cloutier, S., McCallum, B.D., Loutre, C., Banks, T.W., Wicker, T., Feuillet, C., Keller, B. and 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**, 93-106.

Cockram, J., White, J., Zuluaga, D.L., Smith, D., Comadran, J., Macaulay, M., Luo, Z., Kearsey, M.J., Werner, P., Harrap, D., Tapsell, C., Liu, H., Hedley, P.E., Stein, N., Schulte, D., Steuernagel, B., Marshall, D.F., Thomas, W.T., Ramsay, L., Mackay, I., Balding, D.J., Waugh, R. and O'Sullivan, D.M. (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21611-21616.

Coe Jr, E.H. and Neuffer, M.G. (1977) The genetics of corn. In: *Corn and Corn Improvement (G.F. Sprague, ed.)* pp. 111-223. Amer. Sot. Agron., Madison, Wise.

Cui, H., Tsuda, K. and Parker, J.E. (2015) Effector-triggered immunity: from pathogen perception to robust defense. *Annual review of plant biology* **66**, 487-511.

Cuthbert, P.A., Somers, D.J., Thomas, J., Cloutier, S. and Brule-Babel, A. (2006) Fine mapping Fhb1, a major gene controlling fusarium head blight resistance in bread wheat (Triticum aestivum L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **112**, 1465-1472.

D'Halluin, K., Vanderstraeten, C., Van Hulle, J., Rosolowska, J., Van Den Brande, I., Pennewaert, A., D'Hont, K., Bossut, M., Jantz, D., Ruiter, R. and Broadhvest, J. (2013) Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant biotechnology journal* **11**, 933-941.

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

Dawkins, R. and Krebs, J.R. (1979) Arms races between and within species. *Proceedings of the Royal Society of London. Series B, Biological sciences* **205**, 489-511.

Day, B., Dahlbeck, D., Huang, J., Chisholm, S.T., Li, D. and Staskawicz, B.J. (2005) Molecular basis for the *RIN4* negative regulation of *RPS2* disease resistance. *The Plant cell* **17**, 1292-1305.

De Wolf, E., Murray, T., Paul, P., Osborne, L. and Tenuta, A. (2011) Identification and Management of Stem Rust on Wheat and Barley. USDA-CREES Extension IPM 2009-41533-05331.

Devoto, A., Hartmann, H.A., Piffanelli, P., Elliott, C., Simmons, C., Taramino, G., Goh, C.S., Cohen, F.E., Emerson, B.C., Schulze-Lefert, P. and Panstruga, R. (2003) Molecular phylogeny and evolution of the plant-specific seven-transmembrane *MLO* family. *Journal of molecular evolution* **56**, 77-88.

Dierking, E.C. and Bilyeu, K.D. (2009) New sources of soybean seed meal and oil composition traits identified through TILLING. *BMC plant biology* **9**, 89.

Dodds, P.N., Lawrence, G.J., Catanzariti, A.-M., Ayliffe, M.A. and Ellis, J.G. (2004) The *Melampsora lini AvrL567* Avirulence Genes Are Expressed in Haustoria and Their Products Are Recognized inside Plant Cells. *The Plant Cell* **16**, 755-768.

Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Teh, T., Wang, C.I., Ayliffe, M.A., Kobe, B. and Ellis, J.G. (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8888-8893.

Dodds, P.N., Lawrence, G.J. and Ellis, J.G. (2001) Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. *The Plant cell* **13**, 163-178.

Dodds, P.N. and Rathjen, J.P. (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature reviews. Genetics* **11**, 539-548.

Doehlemann, G., van der Linde, K., Assmann, D., Schwammbach, D., Hof, A., Mohanty, A., Jackson, D. and Kahmann, R. (2009) Pep1, a secreted effector protein of *Ustilago maydis,* is required for successful invasion of plant cells. *PLoS pathogens* **5**, e1000290.

Doležel, J., Kubalakova, M., Cihalikova, J., Suchankova, P. and Simkova, H. (2011) Chromosome analysis and sorting using flow cytometry. *Methods in molecular biology (Clifton, N.J.)* **701**, 221-238.

Dong, C., Dalton-Morgan, J., Vincent, K. and Sharp, P. (2009) A Modified TILLING Method for Wheat Breeding. *The Plant Genome* **2**, 39-47.

Döring, H.-P., Lin, J., Uhrig, H. and Salamini, F. (1999) Clonal analysis of the development of the barley (*Hordeum vulgare L.*) leaf using periclinal chlorophyll chimeras. *Planta* **207**, 335-342.

Doussinault, G., Delibes, A., Sanchez-Monge, R. and Garcia-Olmedo, F. (1983) Transfer of a dominant gene for resistance to eyespot disease from a wild grass to hexaploid wheat. *Nature* **303**, 698-700.

Ellis, J., Dodds, P. and Pryor, T. (2000) The generation of plant disease resistance gene specificities. *Trends in plant science* **5**, 373-379.

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

Ellis, J.G., Lawrence, G.J., Luck, J.E. and Dodds, P.N. (1999) Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. *The Plant cell* **11**, 495-506.

Engler, C., Kandzia, R. and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PloS one* **3**, e3647.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J.D.G., Patron, N.J. and Marillonnet, S. (2014) A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synthetic Biology* **3**, 839-843.

Erickson, A.W. (1945) *McFadden's Hope: Fighting Plant Breeders Win Battle for Bread*. Minneapolis, MN:Field Notes Crop Reporting Service.

Faris, J.D., Fellers, J.P., Brooks, S.A. and Gill, B.S. (2003) A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. *Genetics* **164**, 311-321.

Faris, J.D., Zhang, Z., Lu, H., Lu, S., Reddy, L., Cloutier, S., Fellers, J.P., Meinhardt, S.W., Rasmussen, J.B., Xu, S.S., Oliver, R.P., Simons, K.J. and Friesen, T.L. (2010) A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13544-13549.

Farrell, A., Coleman, B.I., Benenati, B., Brown, K.M., Blader, I.J., Marth, G.T. and Gubbels, M.J. (2014) Whole genome profiling of spontaneous and chemically induced mutations in *Toxoplasma gondii*. *BMC Genomics* **15**, 354.

Fetch, T., Johnston, P.A. and Pickering, R. (2009) Chromosomal location and inheritance of stem rust resistance transferred from *Hordeum bulbosum* into cultivated barley (*H. vulgare*). *Phytopathology* **99**, 339-343.

Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A. and 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**, 15253-15258.

Fradin, E.F., Abd-El-Haliem, A., Masini, L., van den Berg, G.C., Joosten, M.H. and Thomma, B.P. (2011) Interfamily transfer of tomato *Ve1* mediates Verticillium resistance in Arabidopsis. *Plant physiology* **156**, 2255-2265.

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., Sela, H., Fahima, T. and Dubcovsky, J. (2009) A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science* **323**, 1357-1360.

Gaul, H. (1959) Über die Chimärenbildung in. Gerste Pflanzen nach Röntgenbestrahlung von. Samen. *Flora* **147**, 207-241.

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C.V., Sánchez-León, S., Baltes, N.J., Starker, C., Barro, F., Gao, C. and Voytas, D.F. (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *The Plant journal : for cell and molecular biology* **89**, 1251-1262.

Giorgi, D., Farina, A., Grosso, V., Gennaro, A., Ceoloni, C. and Lucretti, S. (2013) FISHIS: fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. *PloS one* **8**, e57994.

Glen, H.F. and Institute, N.B. (2002) *Cultivated Plants of Southern Africa: Botanical Names, Common Names, Origins, Literature*:Jacana.

Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U. and Stein, N. (2009) TILLING in the two-rowed barley cultivar 'Barke' reveals preferred sites of functional diversity in the gene *HvHox1*. *BMC research notes* **2**, 258.

Grando, S. and Macpherson, H.G. (2005) *Food barley: importance, uses and local knowledge. Proceedings of the International Workshop on Food Barley Improvement, Hammamet, Tunisia, 14-17 January, 2002*. Aleppo:International Center for Agricultural Research in the Dry Areas (ICARDA).

Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I. and Moore, G. (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* **439**, 749-752.

Groos, C., Robert, N., Bervas, E. and Charmet, G. (2003) Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **106**, 1032-1040.

Gu, K., Yang, B., Tian, D., Wu, L., Wang, D., Sreekala, C., Yang, F., Chu, Z., Wang, G.L., White, F.F. and Yin, Z. (2005) *R* gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122-1125.

Halterman, D., Zhou, F., Wei, F., Wise, R.P. and Schulze-Lefert, P. (2001) The *MLA6* coiled-coil, NBS-LRR protein confers *AvrMla6*-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. *The Plant journal : for cell and molecular biology* **25**, 335-348.

Harper, A.L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., Hattori, C., Werner, P. and Bancroft, I. (2012) Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nat Biotech* **30**, 798-802.

Harwood, W.A. (2014a) A protocol for high-throughput *Agrobacterium*-mediated barley transformation. *Methods in molecular biology (Clifton, N.J.)* **1099**, 251-260.

Harwood, W.A. (2014b) A Protocol for High-Throughput *Agrobacterium*-Mediated Barley Transformation. In: *Cereal Genomics: Methods and Protocols* (Henry, R.J. and Furtado, A. eds), pp. 251-260. Totowa, NJ: Humana Press.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C. and Haldar, K. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**, 1934-1937.

Huang, L., Brooks, S.A., Li, W., Fellers, J.P., Trick, H.N. and Gill, B.S. (2003) Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. *Genetics* **164**, 655-664.

Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W., Guo, Y., Lu, Y., Zhou, C., Fan, D., Weng, Q., Zhu, C., Huang, T., Zhang, L., Wang, Y., Feng, L., Furuumi, H., Kubo, T., Miyabayashi, T., Yuan, X., Xu, Q., Dong, G., Zhan, Q., Li, C., Fujiyama, A., Toyoda, A., Lu, T., Feng, Q., Qian, Q., Li, J. and Han, B. (2012a) A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**, 497-501.

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., Liu, K., Huang, T., Zhou, T., Jing, Y., Li, W., Lin, Z., Buckler, E.S., Qian, Q., Zhang, Q.-F., Li, J. and Han, B. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* **42**, 961-967.

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., Fan, D., Lu, Y., Weng, Q., Liu, K., Zhou, T., Jing, Y., Si, L., Dong, G., Huang, T., Lu, T., Feng, Q., Qian, Q., Li, J. and Han, B. (2012b) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* **44**, 32-39.

Huerta-Espino, J. (1992) Analysis of Wheat leaf rust and stem rust virulence on a worldwide basis. University of Minnesota, St. Paul.

Hulbert, S.H. (1997) Structure and evolution of the *rp1* complex conferring rust resistance in maize. *Annual review of phytopathology* **35**, 293-310.

Hulbert, S.H., Webb, C.A., Smith, S.M. and Sun, Q. (2001) Resistance gene complexes: evolution and utilization. *Annual review of phytopathology* **39**, 285-312.

Hummel, A.W., Doyle, E.L. and Bogdanove, A.J. (2012) Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *The New phytologist* **195**, 883-893.

Inoue, Y., Vy, T.T.P., Yoshida, K., Asano, H., Mitsuoka, C., Asuke, S., Anh, V.L., Cumagun, C.J.R., Chuma, I., Terauchi, R., Kato, K., Mitchell, T., Valent, B., Farman, M. and Tosa, Y. (2017) Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science* **357**, 80.

Irish, V.F. and Sussex, I.M. (1992) A fate map of the Arabidopsis embryonic shoot apical meristem. *Development* **115**, 745-753.

Ishida, Y., Tsunashima, M., Hiei, Y. and Komari, T. (2015) Wheat (*Triticum aestivum L.*) transformation using immature embryos. *Methods in molecular biology (Clifton, N.J.)* **1223**, 189-198.

Islam, M.T., Croll, D., Gladieux, P., Soanes, D.M., Persoons, A., Bhattacharjee, P., Hossain, M.S., Gupta, D.R., Rahman, M.M., Mahboob, M.G., Cook, N., Salam, M.U., Surovy, M.Z., Sancho, V.B., Maciel, J.L., NhaniJunior, A., Castroagudin, V.L., Reges, J.T., Ceresini, P.C., Ravel, S., Kellner, R., Fournier, E., Tharreau, D., Lebrun, M.H., McDonald, B.A., Stitt, T., Swan, D., Talbot, N.J., Saunders, D.G., Win, J. and Kamoun,

S. (2016) Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. *BMC biology* **14**, 84.
IWGSC (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345**, 1251788.
Jedel, P.E. (1990) A gene for resistance to *Puccinia graminis* f. sp. *tritici* in PI 382313. *Barley Genet Newsl.* **20**, 43-44.
Jedel, P.E., Metcalfe, D.R. and Martens, J.W. (1989) Assessment of Barley Accessions PI 382313, PI 382474, PI 382915, and PI 382976 for Stem Rust Resistance. *Crop Science* **29**, 1473-1477.
Jia, H., Zhang, Y., Orbovic, V., Xu, J., White, F.F., Jones, J.B. and Wang, N. (2016) Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. *Plant biotechnology journal*.
Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P. and Valent, B. (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO journal* **19**, 4004-4014.
Jiao, Y., Zhao, H., Ren, L., Song, W., Zeng, B., Guo, J., Wang, B., Liu, Z., Chen, J., Li, W., Zhang, M., Xie, S. and Lai, J. (2012) Genome-wide genetic changes during modern breeding of maize. *Nat Genet* **44**, 812-815.
Jin, Y., Steffenson, B.J. and Miller, J.D. (1994) Inheritance of resistance to pathotypes QCC and MCC of *Puccinia graminis* f. sp. *tritici* in barley line Q21861 and temperature effects on the expression of resistance. *Phytopathology* **84**, 452-455.
Jin, Y., Szabo, L.J., Rouse, M., Fetch, T.J., Pretorius, Z.A., Wanyere, R. and Njau, P. (2009) Detection of virulence to resistance gene *Sr36* within race TTKS lineage of *Puccinia graminis* f. sp. *tritici*. *Plant Disease* **93**, 367-370.
Jo, K.-R., Kim, C.-J., Kim, S.-J., Kim, T.-Y., Bergervoet, M., Jongsma, M.A., Visser, R.G., Jacobsen, E. and Vossen, J.H. (2014) Development of late blight resistant potatoes by cisgene stacking. *BMC Biotechnology* **14**, 50.
Johri, M.M. and Coe, E.H. (1983) Clonal analysis of corn plant development. *Developmental Biology* **97**, 154-172.
Jones, J.D. and Dangl, J.L. (2006) The plant immune system. *Nature* **444**, 323-329.
Joosten, M. and de Wit, P. (1999) THE TOMATO-*CLADOSPORIUM FULVUM* INTERACTION: A Versatile Experimental System to Study Plant-Pathogen Interactions. *Annual review of phytopathology* **37**, 335-367.
Jørgensen, I.H. (1992) Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. *Euphytica* **63**, 141-152.
Jupe, F., Pritchard, L., Etherington, G.J., Mackenzie, K., Cock, P.J., Wright, F., Sharma, S.K., Bolser, D., Bryan, G.J., Jones, J.D. and Hein, I. (2012a) Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics* **13**, 75.
Jupe, F., Pritchard, L., Etherington, G.J., MacKenzie, K., Cock, P.J., Wright, F., Sharma, S.K., Bolser, D., Bryan, G.J., Jones, J.D. and Hein, I. (2012b) Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics* **13**, 75.
Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G.J., Maclean, D., Cock, P.J., Leggett, R.M., Bryan, G.J., Cardle, L., Hein, I. and Jones, J.D. (2013) Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant journal : for cell and molecular biology* **76**, 530-544.
Kanzaki, H., Yoshida, K., Saitoh, H., Fujisaki, K., Hirabuchi, A., Alaux, L., Fournier, E., Tharreau, D. and Terauchi, R. (2012) Arms race co-evolution of *Magnaporthe oryzae AVR-Pik* and rice *Pik* genes driven by their physical interactions. *The Plant journal : for cell and molecular biology* **72**, 894-907.

Keet, J.H., Cindi, D.D. and du Preez, P.J. (2016) Assessing the invasiveness of *Berberis aristata* and *B. julianae* (Berberidaceae) in South Africa: Management options and legal recommendations. *South African Journal of Botany* **105**, 288-298.

Kerbach, S., Lorz, H. and Becker, D. (2005) Site-specific recombination in *Zea mays*. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **111**, 1608-1616.

Khan, I.A. (1996) Utilization of molecular markers in the selection and characterization of wheat-alien recombinant chromosomes. University of Adelaide, Adelaide, Australia.

Khan, I.A. (2000) Molecular and agronomic characterization of wheat-*Agropyron intermedium* recombinant chromosomes. *Plant Breeding* **119**, 25-29.

Khang, C.H., Berruyer, R., Giraldo, M.C., Kankanala, P., Park, S.Y., Czymmek, K., Kang, S. and Valent, B. (2010) Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. *The Plant cell* **22**, 1388-1403.

Kibirige-Sebunya, I. and Knott, D.R. (1983) Transfer of stem rust resistance to wheat from an Agropyron chromosome having a gametocidal effect. *Canadian Journal of Genetics and Cytology* **25**, 215-221.

Kleinhofs, A., Brueggeman, R., Nirmala, J., Zhang, L., Mirlohi, A., Druka, A., Rostoks, N. and Steffenson, B.J. (2009) Barley Stem Rust Resistance Genes: Structure and Function. *The Plant Genome* **2**, 109-120.

Knott, D.R. (1984) The genetic nature of mutations of a gene for yellow pigment linked to *Lr19* in 'Agatha' wheat. *Canadian Journal of Genetics and Cytology* **26**, 392-393.

Konishi, S., Sasakuma, T. and Sasanuma, T. (2010) Identification of novel *Mlo* family members in wheat and their genetic characterization. *Genes & genetic systems* **85**, 167-175.

Krasileva, K.V., Dahlbeck, D. and Staskawicz, B.J. (2010) Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *The Plant cell* **22**, 2444-2458.

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. and Dubcovsky, J. (2017) Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E913-e921.

Krattinger, S.G., Lagudah, E.S., Spielmeyer, W., Singh, R.P., Huerta-Espino, J., McFadden, H., Bossolini, E., Selter, L.L. and Keller, B. (2009) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* **323**, 1360-1363.

Krattinger, S.G., Sucher, J., Selter, L.L., Chauhan, H., Zhou, B., Tang, M., Upadhyaya, N.M., Mieulet, D., Guiderdoni, E., Weidenbach, D., Schaffrath, U., Lagudah, E.S. and Keller, B. (2016) The wheat durable, multipathogen resistance gene *Lr34* confers partial blast resistance in rice. *Plant biotechnology journal* **14**, 1261-1268.

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

Kuang, H., Woo, S.S., Meyers, B.C., Nevo, E. and Michelmore, R.W. (2004) Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. *The Plant cell* **16**, 2870-2894.

Kunzel, G., Korzun, L. and Meister, A. (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics* **154**, 397-412.

Kurowska, M., Labocha-Pawłowska, A., Gnizda, D., Maluszynski, M. and Szarejko, I. (2012) Molecular analysis of point mutations in a barley genome exposed to MNU and

gamma rays. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **738–739**, 52-70.

Lababidi, S., Mejlhede, N., Rasmussen, S.K., Backes, G., Al-Said, W., Baum, M. and Jahoor, A. (2009) Identification of barley mutants in the cultivar 'Lux' at the *Dhn* loci through TILLING. *Plant Breeding* **128**, 332-336.

Labuschagne, M.T., Pretorius, Z.A. and Grobbelaar, B. (2002) The influence of leaf rust resistance genes Lr29, Lr34, Lr35 and Lr37 on breadmaking quality in wheat. *Euphytica* **124**, 65-70.

Lacombe, S., Rougon-Cardoso, A., Sherwood, E., Peeters, N., Dahlbeck, D., van Esse, H.P., Smoker, M., Rallapalli, G., Thomma, B.P., Staskawicz, B., Jones, J.D. and Zipfel, C. (2010) Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. *Nature biotechnology* **28**, 365-369.

Leonard, K.J. and Szabo, L.J. (2005) Stem rust of small grains and grasses caused by *Puccinia graminis*. *Molecular plant pathology* **6**, 99-111.

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., Liu, J., Warburton, M.L., Cheng, Y., Hao, X., Zhang, P., Zhao, J., Liu, Y., Wang, G., Li, J. and Yan, J. (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat Genet* **45**, 43-50.

Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J., Li, J. and Gao, C. (2016) Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nature plants* **2**, 16139.

Li, T., Liu, B., Spalding, M.H., Weeks, D.P. and Yang, B. (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature biotechnology* **30**, 390-392.

Li, Z., Xing, A., Moon, B.P., McCardell, R.P., Mills, K. and Falco, S.C. (2009) Site-specific integration of transgenes in soybean via recombinase-mediated DNA cassette exchange. *Plant physiology* **151**, 1087-1095.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., Wang, X., Huang, Z., Li, J., Zhang, C., Wang, T., Zhang, Y., Wang, A., Zhang, Y., Lin, K., Li, C., Xiong, G., Xue, Y., Mazzucato, A., Causse, M., Fei, Z., Giovannoni, J.J., Chetelat, R.T., Zamir, D., Stadler, T., Li, J., Ye, Z., Du, Y. and Huang, S. (2014) Genomic analyses provide insights into the history of tomato breeding. *Nat Genet* **46**, 1220-1226.

Liu, S., Sehgal, S.K., Li, J., Lin, M., Trick, H.N., Yu, J., Gill, B.S. and Bai, G. (2013a) Cloning and characterization of a critical regulator for preharvest sprouting in wheat. *Genetics* **195**, 263-273.

Liu, S., Yeh, C.T., Tang, H.M., Nettleton, D. and Schnable, P.S. (2012) Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PloS one* **7**, e36406.

Liu, W., Danilova, T.V., Rouse, M.N., Bowden, R.L., Friebe, B., Gill, B.S. and Pumphrey, M.O. (2013b) Development and characterization of a compensating wheat-*Thinopyrum intermedium* Robertsonian translocation with *Sr44* resistance to stem rust (Ug99). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **126**, 1167-1177.

Liu, W., Frick, M., Huel, R., Nykiforuk, C.L., Wang, X., Gaudet, D.A., Eudes, F., Conner, R.L., Kuzyk, A., Chen, Q., Kang, Z. and Laroche, A. (2014) The stripe rust resistance gene *Yr10* encodes an evolutionary-conserved and unique CC-NBS-LRR sequence in wheat. *Molecular plant* **7**, 1740-1755.

Loutre, C., Wicker, T., Travella, S., Galli, P., Scofield, S., Fahima, T., Feuillet, C. and 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 : for cell and molecular biology* **60**, 1043-1054.

Louwerse, J.D., van Lier, M.C., van der Steen, D.M., de Vlaam, C.M., Hooykaas, P.J. and Vergunst, A.C. (2007) Stable recombinase-mediated cassette exchange in Arabidopsis using *Agrobacterium tumefaciens*. *Plant physiology* **145**, 1282-1293.

Lu, G., Harper, A.L., Trick, M., Morgan, C., Fraser, F., O'Neill, C. and Bancroft, I. (2014) Associative transcriptomics study dissects the genetic architecture of seed glucosinolate content in *Brassica napus*. *DNA research : an international journal for rapid publication of reports on genes and genomes* **21**, 613-625.

Luck, J.E., Lawrence, G.J., Dodds, P.N., Shepherd, K.W. and Ellis, J.G. (2000) Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *The Plant cell* **12**, 1367-1377.

Luo, M., Gilbert, B. and Ayliffe, M. (2016) Applications of CRISPR/Cas9 technology for targeted mutagenesis, gene replacement and stacking of genes in higher plants. *Plant cell reports* **35**, 1439-1450.

Maccaferri, M., Zhang, J., Bulli, P., Abate, Z., Chao, S., Cantu, D., Bossolini, E., Chen, X., Pumphrey, M. and 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 (Bethesda, Md.)* **5**, 449-465.

MacQueen, A. and Bergelson, J. (2016) Modulation of *R*-gene expression across environments. *Journal of experimental botany* **67**, 2093-2105.

Mago, R., Bariana, H.S., Dundas, I.S., Spielmeyer, W., Lawrence, G.J., Pryor, A.J. and Ellis, J.G. (2005a) Development of PCR markers for the selection of wheat stem rust resistance genes *Sr24* and *Sr26* in diverse wheat germplasm. *Theoretical and Applied Genetics* **111**, 496-504.

Mago, R., Lawrence, G.J. and Ellis, J.G. (2011) The application of DNA marker and doubled-haploid technology for stacking multiple stem rust resistance genes in wheat. *Mol Breeding* **27**, 329-335.

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

Mago, R., Spielmeyer, W., Lawrence, G.J., Ellis, J.G. and Pryor, A.J. (2004) Resistance genes for rye stem rust (*SrR*) and barley powdery mildew (*Mla*) are located in syntenic regions on short arm of chromosome. *Genome* **47**, 112-121.

Mago, R., Zhang, P., Vautrin, S., Simkova, H., Bansal, U., Luo, M.C., Rouse, M., Karaoglu, H., Periyannan, S., Kolmer, J., Jin, Y., Ayliffe, M.A., Bariana, H., Park, R.F., McIntosh, R., Doležel, J., Berges, H., Spielmeyer, W., Lagudah, E.S., Ellis, J.G. and Dodds, P.N. (2015) The wheat *Sr50* gene reveals rich diversity at a cereal disease resistance locus. *Nature plants* **1**, 15186.

Marais, G.F. (1992) The modification of a common wheat-*Thinopyrum distichum* translocated chromosome with a locus homoeoallelic to *Lr19*. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **85**, 73-78.

Marcussen, T., Sandve, S.R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K.S., Wulff, B.B., Steuernagel, B., Mayer, K.F. and Olsen, O.A. (2014) Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**, 1250092.

Marti, M., Good, R.T., Rug, M., Knuepfer, E. and Cowman, A.F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**, 1930-1933.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D. and Tanksley, S.D. (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**, 1432-1436.

Mascher, M., Jost, M., Kuon, J.-E., Himmelbach, A., Aßfalg, A., Beier, S., Scholz, U., Graner, A. and Stein, N. (2014) Mapping-by-sequencing accelerates forward genetics in barley. *Genome Biology* **15**, R78.

Mayer, K.F., Waugh, R., Brown, J.W., Schulman, A., Langridge, P., Platzer, M., Fincher, G.B., Muehlbauer, G.J., Sato, K., Close, T.J., Wise, R.P. and Stein, N. (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491**, 711-716.

Mayer, K.F.X., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., Sourdille, P., Endo, T.R., Kubaláková, M., Číhalíková, J., Dubská, Z., Vrána, J., Šperková, R., Šimková, H., Febrer, M., Clissold, L., McLay, K., Singh, K., Chhuneja, P., Singh, N.K., Khurana, J., Akhunov, E., Choulet, F., Alberti, A., Barbe, V., Wincker, P., Kanamori, H., Kobayashi, F., Itoh, T., Matsumoto, T., Sakai, H., Tanaka, T., Wu, J., Ogihara, Y., Handa, H., Maclachlan, P.R., Sharpe, A., Klassen, D., Edwards, D., Batley, J., Olsen, O.-A., Sandve, S.R., Lien, S., Steuernagel, B., Wulff, B., Caccamo, M., Ayling, S., Ramirez-Gonzalez, R.H., Clavijo, B.J., Wright, J., Pfeifer, M., Spannagl, M., Martis, M.M., Mascher, M., Chapman, J., Poland, J.A., Scholz, U., Barry, K., Waugh, R., Rokhsar, D.S., Muehlbauer, G.J., Stein, N., Gundlach, H., Zytnicki, M., Jamilloux, V., Quesneville, H., Wicker, T., Faccioli, P., Colaiacovo, M., Stanca, A.M., Budak, H., Cattivelli, L., Glover, N., Pingault, L., Paux, E., Sharma, S., Appels, R., Bellgard, M., Chapman, B., Nussbaumer, T., Bader, K.C., Rimbert, H., Wang, S., Knox, R., Kilian, A., Alaux, M., Alfama, F., Couderc, L., Guilhot, N., Viseux, C., Loaec, M., Keller, B. and Praud, S. (2014) A chromosome-based draft sequence of the hexaploid bread wheat *Triticum aestivum* genome. *Science* **345**.

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (2000) Targeted screening for induced mutations. *Nat Biotech* **18**, 455-457.

McDaniel, C.N. and Poethig, R.S. (1988) Cell-lineage patterns in the shoot apical meristem of the germinating maize embryo. *Planta* **175**, 13-22.

McFadden, E.S. (1930a) A successful transfer of emmer characters to vulgare wheat. *Journal of the American Society of Agronomy* **22**, 1020-1034.

McFadden, E.S. (1930b) A Successful Transfer of Emmer Characters to Vulgare Wheat1. *Agron. J.* **22**, 1020-1034.

McGrann, G.R.D., Smith, P.H., Burt, C., Mateos, G.R., Chama, T.N., MacCormack, R., Wessels, E., Agenbag, G. and Boyd, L.A. (2014) Genomic and genetic analysis of the wheat race-specific yellow rust resistance gene *Yr5*. *Journal of Plant Science and Molecular Breeding* **3**.

McIntosh, R., Dubcovsky, J., Rogers, W., Morris, C. and Xia, X. (2017) Catalogue of gene symbols for wheat: 2017 supplement. *Komugi Wheat Genet. Resour. Database Comm. Natl. BioResour. Proj., Kyoto, Japan, updated Jan. 11.* *http://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp*

McIntosh, R., Wellings, C. and Park, R. (1995) *Wheat Rusts: An Atlas of Resistance Genes*:CSIRO PUBLISHING.

McIntosh, R.A. and Gyarfas, J. (1971) *Triticum timopheevi* as a source of resistance to wheat stem rust. *Zeitschrift für Pflazenzüchtung* **66**, 240-248.

Mendes, B.M.J., Cardoso, S.C., Boscariol-Camargo, R.L., Cruz, R.B., Mourão Filho, F.A.A. and Bergamin Filho, A. (2010) Reduction in susceptibility to *Xanthomonas axonopodis* pv. *citri* in transgenic Citrus sinensis expressing the rice *Xa21* gene. *Plant Pathology* **59**, 68-75.

Mestre, P. and Baulcombe, D.C. (2006) Elicitor-mediated oligomerization of the tobacco *N* disease resistance protein. *The Plant cell* **18**, 491-501.

Meyers, B.C., Chin, D.B., Shen, K.A., Sivaramakrishnan, S., Lavelle, D.O., Zhang, Z. and Michelmore, R.W. (1998a) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *The Plant cell* **10**, 1817-1832.

Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral, B.W. and Young, N.D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant journal : for cell and molecular biology* **20**, 317-332.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *The Plant cell* **15**, 809-834.

Meyers, B.C., Shen, K.A., Rohani, P., Gaut, B.S. and Michelmore, R.W. (1998b) Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. *The Plant cell* **10**, 1833-1846.

Michelmore, R.W. and Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome research* **8**, 1113-1130.

Moffett, P., Farnham, G., Peart, J. and Baulcombe, D.C. (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *The EMBO journal* **21**, 4511-4519.

Montenegro, J.D., Golicz, A.A., Bayer, P.E., Hurgobin, B., Lee, H., Chan, C.K., Visendi, P., Lai, K., Doležel, J., Batley, J. and Edwards, D. (2017) The pangenome of hexaploid bread wheat. *The Plant journal : for cell and molecular biology* **90**, 1007-1013.

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. and Lagudah, E. (2015) A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat Genet* **47**, 1494-1498.

Morita, R., Kusaba, M., Iida, S., Yamaguchi, H., Nishio, T. and Nishimura, M. (2009) Molecular characterization of mutations induced by gamma irradiation in rice. *Genes & genetic systems* **84**, 361-370.

Munns, R. and Tester, M. (2008) Mechanisms of salinity tolerance. *Annual review of plant biology* **59**, 651-681.

Muth, J., Hartje, S., Twyman, R.M., Hofferbert, H.R., Tacke, E. and Prufer, D. (2008) Precision breeding for novel starch variants in potato. *Plant biotechnology journal* **6**, 576-584.

Nandy, S., Zhao, S., Pathak, B.P., Manoharan, M. and Srivastava, V. (2015) Gene stacking in plant cell using recombinases for gene integration and nucleases for marker gene deletion. *BMC biotechnology* **15**, 93.

Nanto, K., Sato, K., Katayama, Y. and Ebinuma, H. (2009) Expression of a transgene exchanged by the recombinase-mediated cassette exchange (RMCE) method in plants. *Plant cell reports* **28**, 777-785.

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D. and Kamoun, S. (2017) Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Scientific reports* **7**, 482.

Niu, Z., Klindworth, D.L., Yu, G., L Friesen, T., Chao, S., Jin, Y., Cai, X., Ohm, J.-B., Rasmussen, J.B. and Xu, S.S. (2014) Development and characterization of wheat lines carrying stem rust resistance gene *Sr43* derived from *Thinopyrum ponticum*. *Theoretical and Applied Genetics* **127**, 969-980.

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

Nordstrom, K.J., Albani, M.C., James, G.V., Gutjahr, C., Hartwig, B., Turck, F., Paszkowski, U., Coupland, G. and Schneeberger, K. (2013) Mutation identification by direct

comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. *Nature biotechnology* **31**, 325-330.
O'Connell, R.J. and Panstruga, R. (2006) Tete a tete inside a plant cell: establishing compatibility between plants and biotrophic fungi and oomycetes. *The New phytologist* **171**, 699-718.
Okmen, B., Etalo, D.W., Joosten, M.H., Bouwmeester, H.J., de Vos, R.C., Collemare, J. and de Wit, P.J. (2013) Detoxification of α-tomatine by *Cladosporium fulvum* is required for full virulence on tomato. *The New phytologist* **198**, 1203-1214.
Olivera Firpo, P.D., Newcomb, M., Flath, K., Sommerfeldt-Impe, N., Szabo, L.J., Carter, M., Luster, D.G. and Jin, Y. (2017) Characterization of *Puccinia graminis* f. sp. *tritici* isolates derived from an unusual wheat stem rust outbreak in Germany in 2013. *Plant Pathology*, n/a-n/a.
Olivera, P., Newcomb, M., Szabo, L.J., Rouse, M., Johnson, J., Gale, S., Luster, D.G., Hodson, D., Cox, J.A., Burgin, L., Hort, M., Gilligan, C.A., Patpour, M., Justesen, A.F., Hovmoller, M.S., Woldeab, G., Hailu, E., Hundie, B., Tadesse, K., Pumphrey, M., Singh, R.P. and Jin, Y. (2015) Phenotypic and Genotypic Characterization of Race TKTTF of *Puccinia graminis* f. sp. *tritici* that Caused a Wheat Stem Rust Epidemic in Southern Ethiopia in 2013-14. *Phytopathology* **105**, 917-928.
Olivera, P.D., Kolmer, J.A., Anikster, Y. and Steffenson, B.J. (2007) Resistance of Sharon goatgrass (*Aegilops sharonensis*) to fungal diseases of wheat. *Plant Disease* **91**, 942-950.
Olsen, O., Wang, X. and von Wettstein, D. (1993) Sodium azide mutagenesis: preferential generation of A.T-->G.C transitions in the barley *Ant18* gene. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8043-8047.
Ori, N., Eshed, Y., Paran, I., Presting, G., Aviv, D., Tanksley, S., Zamir, D. and Fluhr, R. (1997) The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *The Plant cell* **9**, 521-532.
Pankin, A., Campoli, C., Dong, X., Kilian, B., Sharma, R., Himmelbach, A., Saini, R., Davis, S.J., Stein, N., Schneeberger, K. and von Korff, M. (2014) Mapping-by-sequencing identifies *HvPHYTOCHROME C* as a candidate gene for the *early maturity 5* locus modulating the circadian clock and photoperiodic flowering in barley. *Genetics* **198**, 383-396.
Park, S.H., Lee, B.-M., Salas, M.G., Srivatanakul, M. and Smith, R.H. (2000) Shorter T-DNA or additional virulence genes improve Agrobactrium-mediated transformation. *Theoretical and Applied Genetics* **101**, 1015-1020.
Parniske, M., Hammond-Kosack, K.E., Golstein, C., Thomas, C.M., Jones, D.A., Harrison, K., Wulff, B.B. and Jones, J.D. (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* **91**, 821-832.
Parniske, M. and Jones, J.D. (1999) Recombination between diverged clusters of the tomato *Cf-9* plant disease resistance gene family. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 5850-5855.
Patpour, M., Hovmøller, M.S., Shahin, A.A., Newcomb, M., Olivera, P., Jin, Y., Luster, D., Hodson, D., Nazari, K. and Azab, M. (2015) First Report of the Ug99 Race Group of Wheat Stem Rust, *Puccinia graminis* f. sp. *tritici*, in Egypt in 2014. *Plant Disease* **100**, 863-863.
Patterson, F.L., Shands, R.G. and Dickson, J.G. (1957) Temperature and seasonal effects on seedling reactions of barley varieties to three races of *Puccinia graminis* f. sp. *tritici*. *Phytopathology* **47**, 395-402.

Peart, J.R., Mestre, P., Lu, R., Malcuit, I. and Baulcombe, D.C. (2005) *NRG1*, a CC-NB-LRR protein, together with *N*, a TIR-NB-LRR protein, mediates resistance against tobacco mosaic virus. *Current biology : CB* **15**, 968-973.

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. and Lagudah, E. (2013) The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. *Science* **341**, 786-788.

Poland, J.A., Bradbury, P.J., Buckler, E.S. and Nelson, R.J. (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6893-6898.

Popp, J. and Hantos, K. (2011) The impact of crop protection on agricultural production. *Studies in Agricultural Economics* **113**, 47-66.

Powers, L. and Hines, L. (1933) Inheritance of reaction to stem rust and barbing awns in barley crosses. *J. Agr. Res.* **46**, 1121-1129.

Pretorius, Z.A., Bender, C.M., Visser, B. and Terefe, T. (2010) First report of a *Puccinia graminis* f. sp. *tritici* race virulent to the *Sr24* and *Sr31* wheat stem rust resistance genes in South Africa. *Plant Disease* **94**, 784.

Pretorius, Z.A., Singh, R.P., Wagoire, W.W. and Payne, T.S. (2000) Detection of Virulence to Wheat Stem Rust Resistance Gene *Sr31* in *Puccinia graminis*. f. sp. *tritici* in Uganda. *Plant Disease* **84**, 203-203.

Qiu, J.W., Schurch, A.C., Yahiaoui, N., Dong, L.L., Fan, H.J., Zhang, Z.J., Keller, B. and Ling, H.Q. (2007) Physical mapping and identification of a candidate for the leaf rust resistance gene *Lr1* of wheat. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **115**, 159-168.

Rairdan, G.J. and Moffett, P. (2006) Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. *The Plant cell* **18**, 2082-2093.

Ramírez-Gonzàlez, R.H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., Jack, P., Caccamo, M. and Uauy, C. (2015) RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant biotechnology journal* **13**, 613-624.

Ramsay, L., Comadran, J., Druka, A., Marshall, D.F., Thomas, W.T.B., Macaulay, M., MacKenzie, K., Simpson, C., Fuller, J., Bonar, N., Hayes, P.M., Lundqvist, U., Franckowiak, J.D., Close, T.J., Muehlbauer, G.J. and Waugh, R. (2011) *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. *Nat Genet* **43**, 169-172.

Ravensdale, M., Bernoux, M., Ve, T., Kobe, B., Thrall, P.H., Ellis, J.G. and Dodds, P.N. (2012) Intramolecular interaction influences binding of the Flax *L5* and *L6* resistance proteins to their *AvrL567* ligands. *PLoS pathogens* **8**, e1003004.

Rawat, N., Pumphrey, M.O., Liu, S., Zhang, X., Tiwari, V.K., Ando, K., Trick, H.N., Bockus, W.W., Akhunov, E., Anderson, J.A. and Gill, B.S. (2016) Wheat *Fhb1* encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nat Genet* **48**, 1576-1580.

REX-Consortium (2016) Combining selective pressures to enhance the durability of disease resistance genes. *Frontiers in plant science* **7**, 1916.

Risk, J.M., Selter, L.L., Chauhan, H., Krattinger, S.G., Kumlehn, J., Hensel, G., Viccars, L.A., Richardson, T.M., Buesing, G., Troller, A., Lagudah, E.S. and Keller, B. (2013) The wheat *Lr34* gene provides resistance against multiple fungal pathogens in barley. *Plant biotechnology journal* **11**, 847-854.

Roelfs, A.P., Singh, R.P. and Saari, E.E. (1992) *Rust diseases of wheat: concepts and methods of disease management*:Centro Internacional de Mejoramiento de Maiz y Trigo.

Römer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U. and Lahaye, T. (2007) Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. *Science* **318**, 645-648.

Romer, P., Recht, S. and Lahaye, T. (2009) A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 20526-20531.

Rooney, H.C., Van't Klooster, J.W., van der Hoorn, R.A., Joosten, M.H., Jones, J.D. and de Wit, P.J. (2005) Cladosporium *Avr2* inhibits tomato *Rcr3* protease required for *Cf-2*-dependent disease resistance. *Science* **308**, 1783-1786.

Rothe, N. (2010) Validation of tilling populations in diploid and hexaploid wheat. M.Sc. thesis. Kansas State University

Rouse, M.N. and Jin, Y. (2011a) Genetics of resistance to race TTKSK of *Puccinia graminis* f. sp. *tritici* in *Triticum monococcum*. *Phytopathology* **101**, 1418-1423.

Rouse, M.N. and Jin, Y. (2011b) Stem Rust Resistance in A-Genome Diploid Relatives of Wheat. *Plant Disease* **95**, 941-944.

Russell, J., Mascher, M., Dawson, I.K., Kyriakidis, S., Calixto, C., Freund, F., Bayer, M., Milne, I., Marshall-Griffiths, T., Heinen, S., Hofstad, A., Sharma, R., Himmelbach, A., Knauft, M., van Zonneveld, M., Brown, J.W.S., Schmid, K., Kilian, B., Muehlbauer, G.J., Stein, N. and Waugh, R. (2016) Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat Genet* **48**, 1024-1030.

Sadat, M.A. and Choi, J. (2017) Wheat Blast: A New Fungal Inhabitant to Bangladesh Threatening World Wheat Production. *The plant pathology journal* **33**, 103-108.

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

Salmeron, J.M., Oldroyd, G.E., Rommens, C.M., Scofield, S.R., Kim, H.S., Lavelle, D.T., Dahlbeck, D. and Staskawicz, B.J. (1996) Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* **86**, 123-133.

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. and Wulff, B.B.H. (2016) Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biology* **17**, 221.

Sarma, D. and Knott, D.R. (1966) THE TRANSFER OF LEAF-RUST RESISTANCE FROM AGROPYRON TO TRITICUM BY IRRADIATION. *Canadian Journal of Genetics and Cytology* **8**, 137-143.

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

Schlegel, R. and Meinel, A. (1994) A Quantitative Trait Locus (QTL) on Chromosome Arm 1RS of Rye and its Effect on Yield Performance of Hexaploid Wheat. *Cereal Research Communications* **22**, 7-13.

Schlegel, R.H.J. (2013) *Rye: Genetics, Breeding, and Cultivation*:CRC Press.

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A.H., Nielsen, K.L., Jorgensen, J.-E., Weigel, D. and Andersen, S.U. (2009) SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat Meth* **6**, 550-551.

Schoonbeek, H.J., Wang, H.H., Stefanato, F.L., Craze, M., Bowden, S., Wallington, E., Zipfel, C. and Ridout, C.J. (2015a) Arabidopsis EF-Tu receptor enhances bacterial disease resistance in transgenic wheat. *The New phytologist* **206**, 606-613.

Schoonbeek, H.J., Wang, H.H., Stefanato, F.L., Craze, M., Bowden, S., Wallington, E., Zipfel, C. and Ridout, C.J. (2015b) Arabidopsis EF-Tu receptor enhances bacterial disease resistance in transgenic wheat. *New Phytologist* **206**, 606-613.

Schumann, G.L. and Leonard, K.J. (2000) Stem rust of wheat (black rust). The Plant Health Instructor.

Seeholzer, S., Tsuchimatsu, T., Jordan, T., Bieri, S., Pajonk, S., Yang, W., Jahoor, A., Shimizu, K.K., Keller, B. and Schulze-Lefert, P. (2010) Diversity at the *Mla* powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Molecular plant-microbe interactions : MPMI* **23**, 497-509.

Shabab, M., Shindo, T., Gu, C., Kaschani, F., Pansuriya, T., Chintha, R., Harzen, A., Colby, T., Kamoun, S. and van der Hoorn, R.A. (2008) Fungal effector protein *AVR2* targets diversifying defense-related cys proteases of tomato. *The Plant cell* **20**, 1169-1183.

Shamanin, V., Salina, E., Wanyera, R., Zelenskiy, Y., Olivera, P. and Morgounov, A. (2016) Genetic diversity of spring wheat from Kazakhstan and Russia for resistance to stem rust Ug99. *Euphytica* **212**, 287-296.

Sharma, R.K., Singh, P.K., Vinod, Joshi, A.K., Bhardwaj, S.C., Bains, N.S. and Singh, S. (2013) Protecting South Asian Wheat Production from Stem Rust (Ug99) Epidemic. *Journal of Phytopathology* **161**, 299-307.

Shen, Q.H., Zhou, F., Bieri, S., Haizel, T., Shirasu, K. and Schulze-Lefert, P. (2003) Recognition specificity and RAR1/SGT1 dependence in barley *Mla* disease resistance genes to the powdery mildew fungus. *The Plant cell* **15**, 732-744.

Shirasawa, K., Hirakawa, H., Nunome, T., Tabata, S. and Isobe, S. (2016) Genome-wide survey of artificial mutations induced by ethyl methanesulfonate and gamma rays in tomato. *Plant biotechnology journal* **14**, 51-60.

Shukla, V.K., Doyon, Y., Miller, J.C., DeKelver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X., Choi, V.M., Rock, J.M., Wu, Y.-Y., Katibah, G.E., Zhifang, G., McCaskill, D., Simpson, M.A., Blakeslee, B., Greenwalt, S.A., Butler, H.J., Hinkley, S.J., Zhang, L., Rebar, E.J., Gregory, P.D. and Urnov, F.D. (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* **459**, 437-441.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D. and Higgins, D.G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**, 539.

Singh, R.P. (2012) Pros and cons of utilizing major, race-specific resistance genes versus partial resistance in breeding rust resistant wheat. pp. 57-65. Ithaca: Borlaug Global Rust Initiative.

Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Njau, P., Wanyera, R., Herrera-Foessel, S.A. and Ward, R.W. (2008a) Will Stem Rust Destroy the World's Wheat Crop? pp. 271-309.

Singh, R.P., Hodson, D.P., Huerta-Espino, J., Yue Jin; Njau, P., Wanyera, R., Herrera-Foessel, S.A. and Ward, R.W. (2008b) Will stem rust destroy the world's wheat crop? *Advances in Agronomy* **98**, 271-309.

Singh, R.P., Hodson, D.P., Jin, Y., Huerta-Espino, J., Kinyua, M.G., Wanyera, R., Njau, P. and Ward, R.W. (2006) Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* **1**.

Singh, R.P., Hodson, D.P., Jin, Y., Lagudah, E.S., Ayliffe, M.A., Bhavani, S., Rouse, M.N., Pretorius, Z.A., Szabo, L.J., Huerta-Espino, J., Basnet, B.R., Lan, C. and Hovmoller, M.S. (2015) Emergence and Spread of New Races of Wheat Stem Rust Fungus: Continued Threat to Food Security and Prospects of Genetic Control. *Phytopathology* **105**, 872-884.

Singh, R.P., William, H.M., Huerta-Espino, J. and Rosewarne, G. (2004) Wheat Rust in Asia: Meeting the Challenges with Old and New Technologies. In: *Proceedings of the 4thInternational Crop Science Congress* pp. 1-13. Brisbane, Australia.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. and Facciotti, D. (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nature biotechnology* **23**, 75-81.

Slade, A.J., McGuire, C., Loeffler, D., Mullenberg, J., Skinner, W., Fazio, G., Holm, A., Brandt, K.M., Steine, M.N., Goodstal, J.F. and Knauf, V.C. (2012) Development of high amylose wheat through TILLING. *BMC plant biology* **12**, 69.

Srichumpa, P., Brunner, S., Keller, B. and Yahiaoui, N. (2005) Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. *Plant physiology* **139**, 885-895.

Srivastava, V., Ariza-Nieto, M. and Wilson, A.J. (2004) Cre-mediated site-specific gene integration for consistent transgene expression in rice. *Plant biotechnology journal* **2**, 169-179.

Stadler, L.J. (1928) MUTATIONS IN BARLEY INDUCED BY X-RAYS AND RADIUM. *Science* **68**, 186-187.

Stahl, E.A., Dwyer, G., Mauricio, R., Kreitman, M. and Bergelson, J. (1999) Dynamics of disease resistance polymorphism at the *Rpm1* locus of Arabidopsis. *Nature* **400**, 667-671.

Stakman, E.C. (1957) Problems in Preventing Plant Disease Epidemics. *American Journal of Botany* **44**, 259-267.

Steffenson, B.J. (1992) Analysis of durable resistance to stem rust in barley. *Euphytica* **63**, 153-167.

Steffenson, B.J., Case, A.J., Pretorius, Z.A., Coetzee, V., Kloppers, F.J., Zhou, H., Chai, Y., Wanyera, R., Macharia, G., Bhavani, S. and Grando, S. (2017) Vulnerability of Barley to African Pathotypes of *Puccinia graminis* f. sp. *tritici* and Sources of Resistance. *Phytopathology* **107**, 950-962.

Steffenson, B.J., Jin, Y., Brueggeman, R.S., Kleinhofs, A. and Sun, Y. (2009) Resistance to stem rust race TTKSK maps to the *rpg4/Rpg5* complex of chromosome 5H of barley. *Phytopathology* **99**, 1135-1141.

Steffenson, B.J., Wilcoxson, R.D. and Roelfs, A.P. (1984) Inheritance of resistance to *Puccinia graminis* f. sp. *secalis* in barley. *Plant Disease* **68**, 762-763.

Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and 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**, 13436-13441.

Stergiopoulos, I., Collemare, J., Mehrabi, R. and De Wit, P.J. (2013) Phytotoxic secondary metabolites and peptides produced by plant pathogenic Dothideomycete fungi. *FEMS microbiology reviews* **37**, 67-93.

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

Strauss, T., van Poecke, R.M., Strauss, A., Romer, P., Minsavage, G.V., Singh, S., Wolf, C., Strauss, A., Kim, S., Lee, H.A., Yeom, S.I., Parniske, M., Stall, R.E., Jones, J.B., Choi, D., Prins, M. and Lahaye, T. (2012) RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19480-19485.

Sucher, J., Boni, R., Yang, P., Rogowsky, P., Buchner, H., Kastner, C., Kumlehn, J., Krattinger, S.G. and Keller, B. (2016) The durable wheat disease resistance gene *Lr34* confers common rust and northern corn leaf blight resistance in maize. *Plant biotechnology journal*.

Sun, Q., Collins, N.C., Ayliffe, M., Smith, S.M., Drake, J., Pryor, T. and Hulbert, S.H. (2001) Recombination between paralogues at the *Rp1* rust resistance locus in maize. *Genetics* **158**, 423-438.

Sun, Y. and Steffenson, B.J. (2005) Reaction of barley seedlings with different stem rust resistance genes to *Puccinia graminis* f. sp. *tritici* and *Puccinia graminis* f. sp. *secalis*. *Canadian Journal of Plant Pathology* **27**, 80-89.

Sun, Y., Steffenson, B.J. and Jin, Y. (1996) Genetics of resistance to *Puccinia graminis* f. sp. *secalis* in barley line Q21861. *Phytopathology* **86**, 1299-1302.

Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C. and Cigan, A.M. (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant physiology* **169**, 931-945.

Tai, T.H., Dahlbeck, D., Clark, E.T., Gajiwala, P., Pasion, R., Whalen, M.C., Stall, R.E. and Staskawicz, B.J. (1999) Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. *Proceedings of the National Academy of Sciences* **96**, 14153-14158.

Takagi, H., Tamiru, M., Abe, A., Yoshida, K., Uemura, A., Yaegashi, H., Obara, T., Oikawa, K., Utsushi, H., Kanzaki, E., Mitsuoka, C., Natsume, S., Kosugi, S., Kanzaki, H., Matsumura, H., Urasaki, N., Kamoun, S. and Terauchi, R. (2015) MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nature biotechnology* **33**, 445-449.

Talame, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U. and Salvi, S. (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant biotechnology journal* **6**, 477-485.

Tanaka, S., Brefort, T., Neidig, N., Djamei, A., Kahnt, J., Vermerris, W., Koenig, S., Feussner, K., Feussner, I. and Kahmann, R. (2014) A secreted *Ustilago maydis* effector promotes virulence by targeting anthocyanin biosynthesis in maize. *eLife* **3**, e01355.

The, T.T., Gupta, R.B., Dyck, P.L., Appels, R., Hohmann, U. and McIntosh, R.A. (1991) Characterization of stem rust resistant derivatives of wheat cultivar Amigo. *Euphytica* **58**, 245-252.

The, T.T., Latter, B.D.H., McIntosh, R.A., Ellison, F.W., Brennan, P.S. and Fisher, J. (1988) Grain yield of near-isogenic lines with added genes for stem rust resistance. In: *7th International Wheat Genetics Symposium* (E., M.T. and M., K.R. eds). Bath Press.

Thind, A.K., Wicker, T., Simkova, H., Fossati, D., Moullet, O., Brabant, C., Vrana, J., Doležel, J. and Krattinger, S.G. (2017) Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nature biotechnology*.

Tian, D., Wang, J., Zeng, X., Gu, K., Qiu, C., Yang, X., Zhou, Z., Goh, M., Luo, Y., Murata-Hori, M., White, F.F. and Yin, Z. (2014) The rice TAL effector-dependent resistance protein *XA10* triggers cell death and calcium depletion in the endoplasmic reticulum. *The Plant cell* **26**, 497-515.

Tian, F., Bradbury, P.J., Brown, P.J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T.R., McMullen, M.D., Holland, J.B. and Buckler, E.S. (2011) Genome-wide association

study of leaf architecture in the maize nested association mapping population. *Nat Genet* **43**, 159-162.

Trick, M., Adamski, N.M., Mugford, S.G., Jiang, C.-C., Febrer, M. and Uauy, C. (2012) Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biology* **12**, 14.

Tripathi, J.N., Lorenzen, J., Bahar, O., Ronald, P. and Tripathi, L. (2014) Transgenic expression of the rice *Xa21* pattern-recognition receptor in banana (*Musa* sp.) confers resistance to *Xanthomonas campestris* pv. *musacearum*. *Plant biotechnology journal* **12**, 663-673.

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C., Tran, R.K., Khan, A.A., Filkov, V., Tai, T.H., Dubcovsky, J. and Comai, L. (2011) Discovery of Rare Mutations in Populations: TILLING by Sequencing. *Plant Physiology* **156**, 1257-1268.

Tseng, T.T., Tyler, B.M. and Setubal, J.C. (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology* **9 Suppl 1**, S2.

Uauy, C., Brevis, J.C., Chen, X., Khan, I., Jackson, L., Chicaiza, O., Distelfeld, A., Fahima, T. and 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**, 97.

Uauy, C., Brevis, J.C. and Dubcovsky, J. (2006a) The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. *Journal of experimental botany* **57**, 2785-2794.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A. and Dubcovsky, J. (2006b) A *NAC* Gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* **314**, 1298-1301.

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R.K., Tsai, H., Berardi, S., Comai, L. and Dubcovsky, J. (2009) A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC plant biology* **9**, 115.

Uauy, C., Wulff, B.B.H. and Dubcovsky, J. (2017) Combining Traditional Mutagenesis with New High-Throughput Sequencing and Genome Editing to Reveal Hidden Variation in Polyploid Wheat. *Annual Review of Genetics* **51**, null.

van der Hoorn, R.A. and Kamoun, S. (2008) From Guard to Decoy: a new model for perception of plant pathogen effectors. *The Plant cell* **20**, 2009-2017.

Vleeshouwers, V.G., Raffaele, S., Vossen, J.H., Champouret, N., Oliva, R., Segretin, M.E., Rietman, H., Cano, L.M., Lokossou, A., Kessel, G., Pel, M.A. and Kamoun, S. (2011) Understanding and exploiting late blight resistance in the age of effectors. *Annual review of phytopathology* **49**, 507-531.

Wang, C., Zhang, X., Fan, Y., Gao, Y., Zhu, Q., Zheng, C., Qin, T., Li, Y., Che, J., Zhang, M., Yang, B., Liu, Y. and Zhao, K. (2015) *XA23* is an executor R protein and confers broad-spectrum disease resistance in rice. *Molecular plant* **8**, 290-302.

Wang, M., Li, Z., Matthews, P.R., Upadhyaya, N.M. and Waterhouse, P.M. (1998) Improved vectors for *Agrobacterium Tumefaciens*-mediated transformation of monocot plants. pp. 401-408. International Society for Horticultural Science (ISHS), Leuven, Belgium.

Wang, M.B., Upadhyaya, N.M., Brettell, R.I.S. and Waterhouse, P.M. (1997) Intron-mediated improvement of a selectable marker gene for plant transformation using *Agrobacterium tumefaciens*. *Journal of Genetics and Breeding* **51**, 325-334.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M., Salvi, S., Milner, S.G., Cattivelli, L., Mastrangelo, A.M., Whan, A., Stephen, S., Barker, G.,

Wieseke, R., Plieske, J., Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A.R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K.J., Hayden, M. and Akhunov, E. (2014a) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant biotechnology journal* **12**, 787-796.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.L. (2014b) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature biotechnology* **32**, 947-951.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191.

Watson, A., Ghosh, S., Williams, M., Cuddy, W.S., Simmonds, J., Rey, M.-D., Hatta, M.A.M., Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N., Breakspear, A., Korolev, A., Rayner, T., Dixon, L.E., Riaz, A., Martin, W., Ryan, M., Edwards, D., Batley, J., Raman, H., Rogers, C., Domoney, C., Moore, G., Harwood, W., Nicholson, P., Dieters, M.J., DeLacy, I.H., Zhou, J., Uauy, C., Boden, S.A., Park, R.F., Wulff, B.B.H. and Hickey, L.T. (2017) Speed breeding: a powerful tool to accelerate crop research and breeding. *bioRxiv*.

Weber, E., Engler, C., Gruetzner, R., Werner, S. and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PloS one* **6**, e16765.

Wicker, T., Narechania, A., Sabot, F., Stein, J., Vu, G.T., Graner, A., Ware, D. and Stein, N. (2008) Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. *BMC Genomics* **9**, 518.

Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E. and Keller, B. (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum L.*) reveals multiple mechanisms of genome evolution. *The Plant journal : for cell and molecular biology* **26**, 307-316.

Williams, S.J., Sohn, K.H., Wan, L., Bernoux, M., Sarris, P.F., Segonzac, C., Ve, T., Ma, Y., Saucet, S.B., Ericsson, D.J., Casey, L.W., Lonhienne, T., Winzor, D.J., Zhang, X., Coerdt, A., Parker, J.E., Dodds, P.N., Kobe, B. and Jones, J.D. (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* **344**, 299-303.

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

Worland, A.J., Law, C.N., Hollins, T.W., Koebner, R.M.D. and Giura, A. (1988) Location of a Gene for Resistance to Eyespot (Pseudocercosparella herpotrichoides) on Chromosome 7D of Bread Wheat. *Plant Breeding* **101**, 43-51.

Wu, C.H., Belhaj, K., Bozkurt, T.O., Birk, M.S. and Kamoun, S. (2016) Helper NLR proteins *NRC2a/b* and *NRC3* but not *NRC1* are required for Pto-mediated cell death and resistance in *Nicotiana benthamiana*. *The New phytologist* **209**, 1344-1352.

Wulff, B.B., Heese, A., Tomlinson-Buhot, L., Jones, D.A., de la Pena, M. and Jones, J.D. (2009) The major specificity-determining amino acids of the tomato *Cf-9* disease resistance protein are at hypervariable solvent-exposed positions in the central leucine-rich repeats. *Molecular plant-microbe interactions : MPMI* **22**, 1203-1213.

Wulff, B.B., Horvath, D.M. and Ward, E.R. (2011) Improving immunity in crops: new tactics in an old game. *Current opinion in plant biology* **14**, 468-476.

Wulff, B.B. and Moscou, M.J. (2014) Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. *Frontiers in plant science* **5**, 692.

Wulff, B.B., Thomas, C.M., Parniske, M. and Jones, J.D. (2004) Genetic variation at the tomato *Cf-4/Cf-9* locus induced by EMS mutagenesis and intralocus recombination. *Genetics* **167**, 459-470.

Wulff, B.B., Thomas, C.M., Smoker, M., Grant, M. and Jones, J.D. (2001) Domain swapping and gene shuffling identify sequences required for induction of an Avr-dependent hypersensitive response by the tomato *Cf-4* and *Cf-9* proteins. *The Plant cell* **13**, 255-272.

Xin, Z., Wang, M.L., Barkley, N.A., Burow, G., Franks, C., Pederson, G. and Burke, J. (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC plant biology* **8**, 103.

Xu, Y. (2010) Molecular Plant Breeding. CABI Publishing.

Yahiaoui, N., Brunner, S. and Keller, B. (2006) Rapid generation of new powdery mildew resistance genes after wheat domestication. *The Plant journal : for cell and molecular biology* **47**, 85-98.

Yahiaoui, N., Srichumpa, P., Dudler, R. and Keller, B. (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. *The Plant journal : for cell and molecular biology* **37**, 528-538.

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S. and Dubcovsky, J. (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19581-19586.

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V. and Dubcovsky, J. (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* **303**, 1640-1644.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 6263-6268.

Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P.-c., Hu, L., Yamasaki, M., Yoshida, S., Kitano, H., Hirano, K. and Matsuoka, M. (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet* **48**, 927-934.

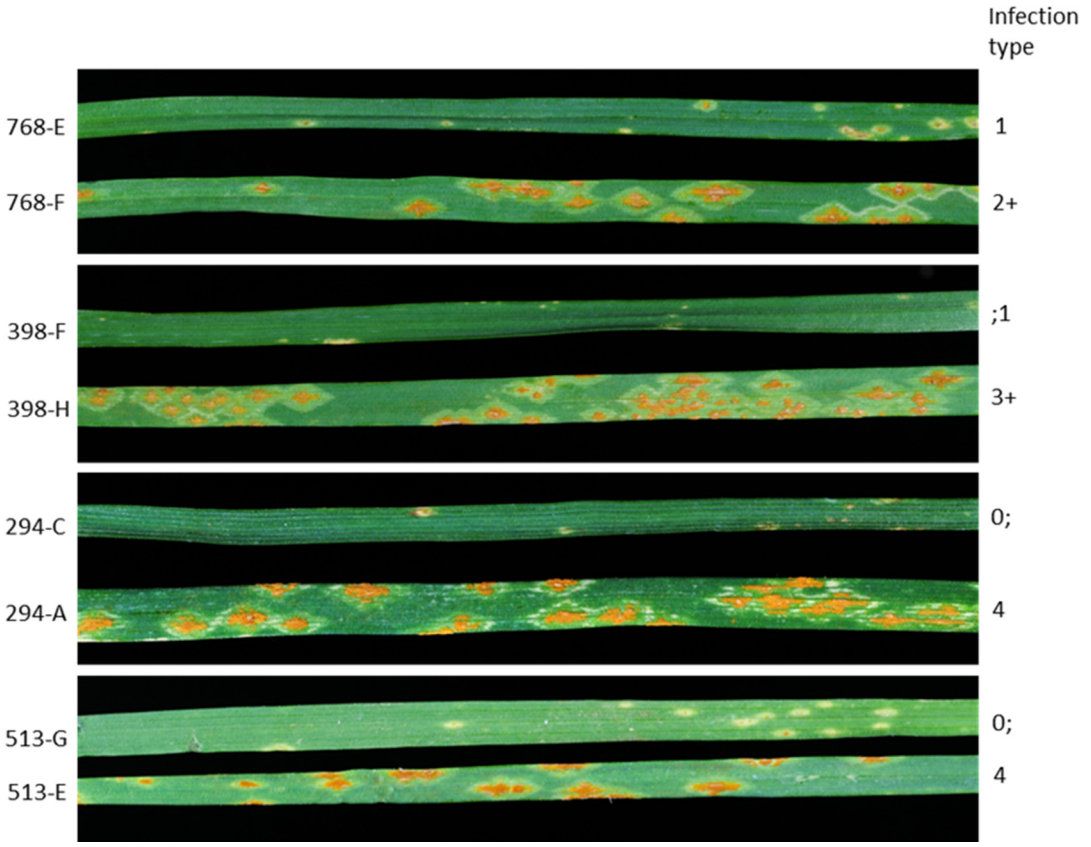
Yoshida, K., Schuenemann, V.J., Cano, L.M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F.N., Kamoun, S., Krause, J., Thines, M., Weigel, D. and Burbano, H.A. (2013) The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. *eLife* **2**, e00731.

Zeng, X., Tian, D., Gu, K., Zhou, Z., Yang, X., Luo, Y., White, F.F. and Yin, Z. (2015) Genetic engineering of the *Xa10* promoter for broad-spectrum and durable resistance to *Xanthomonas oryzae* pv. *oryzae*. *Plant biotechnology journal* **13**, 993-1001.

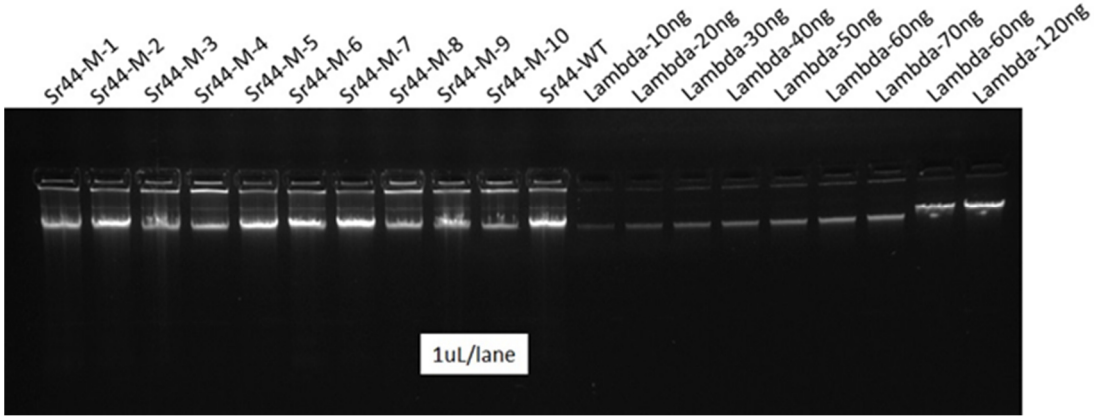
Zhao, B., Lin, X., Poland, J., Trick, H., Leach, J. and Hulbert, S. (2005) A maize resistance gene functions against bacterial streak disease in rice. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15383-15388.

Zhao, K., Tung, C.-W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., McClung, A.M., Bustamante, C.D. and McCouch, S.R. (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nature Communications* **2**, 467.

8 Appendices

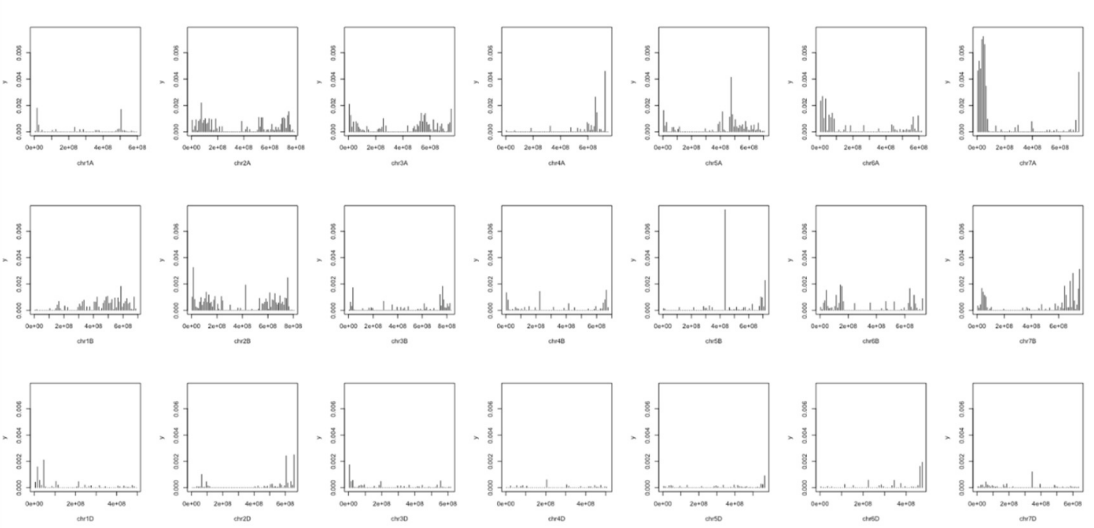


Supplementary Figure 1 Stem rust infection phenotype of representative susceptible mutants and resistant siblings from segregating M_2 families. The seedlings were screened with *Pgt* race TTKSK.

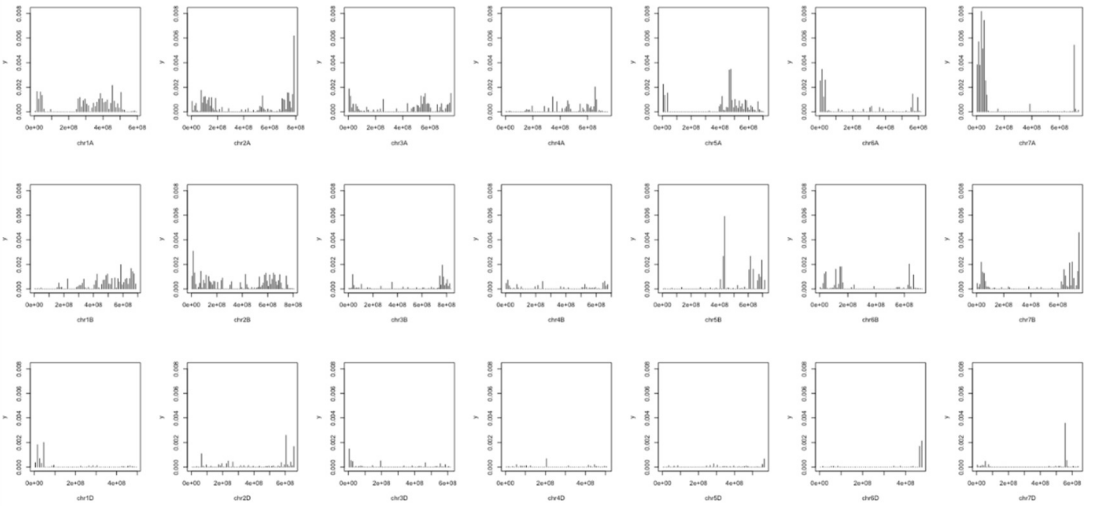


Supplementary Figure 2 DNA concentration estimation using gel electrophoresis. The concentrations of genomic DNA extracted from *Sr44* susceptible mutants and wildtype lines were estimated by comparing to lambda phage DNA with a known concentration. 1 µl of DNA prep was loaded per well. Based on the gel electrophoresis picture, the degree of degradation and contamination with residual RNA was also estimated.

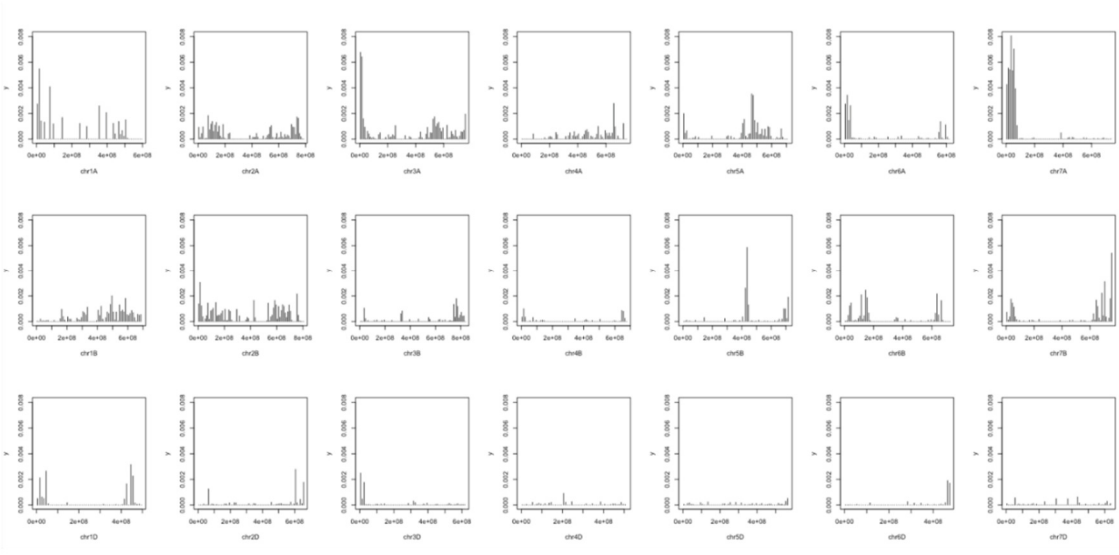
a)


chr1A
chr2A
chr3A
chr4A
chr5A
chr6A
chr7A
chr1B
chr2B
chr3B
chr4B
chr5B
chr6B
chr7B
chr1D
chr2D
chr3D
chr4D
chr5D
chr6D
chr7D

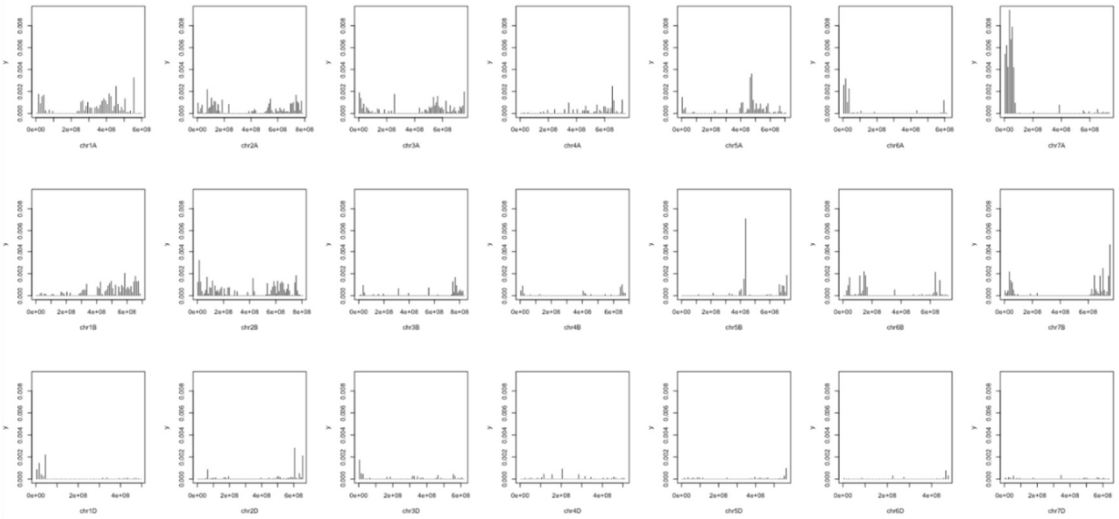
b)

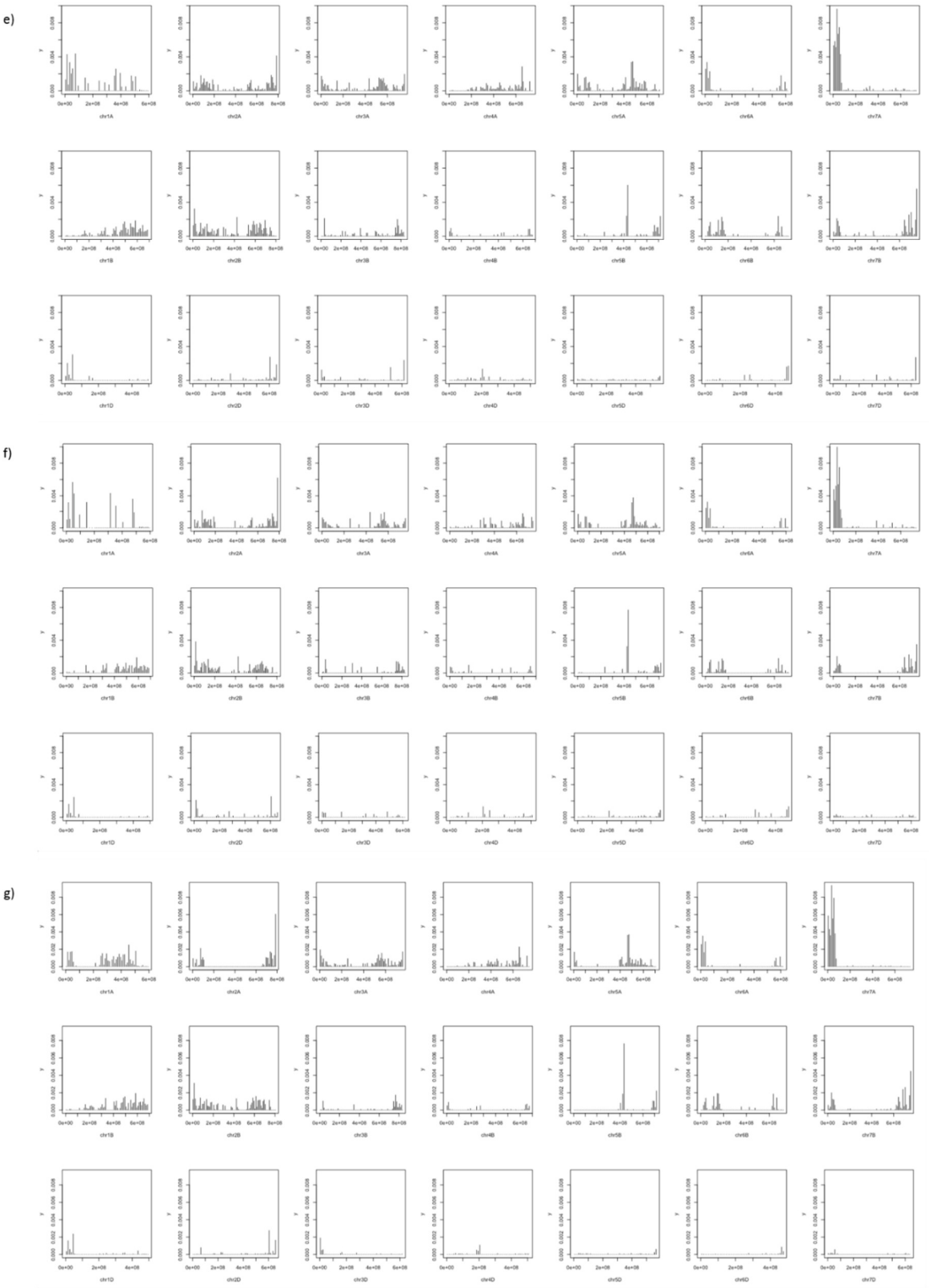

chr1A
chr2A
chr3A
chr4A
chr5A
chr6A
chr7A
chr1B
chr2B
chr3B
chr4B
chr5B
chr6B
chr7B
chr1D
chr2D
chr3D
chr4D
chr5D
chr6D
chr7D

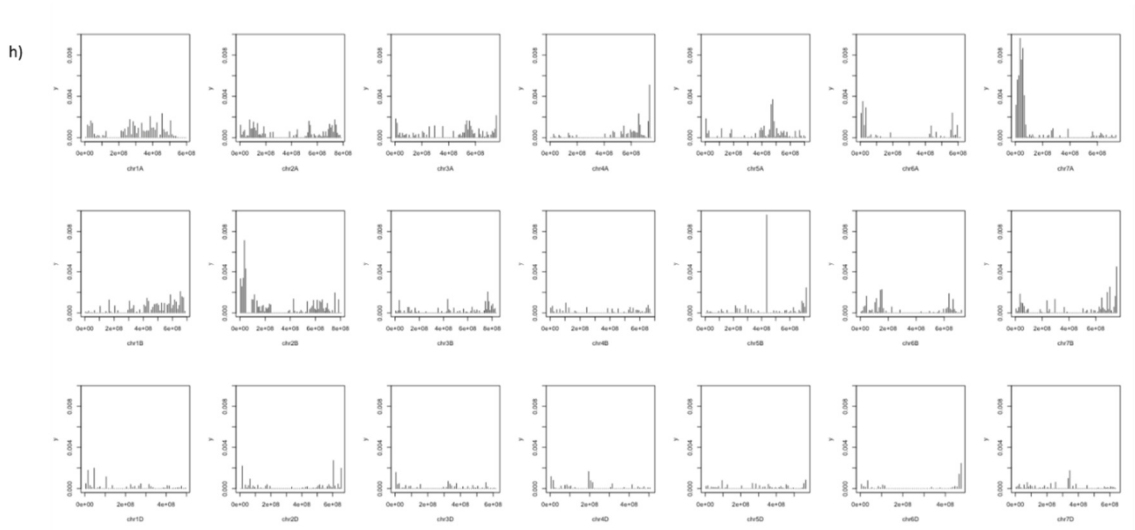
c)


chr1A
chr2A
chr3A
chr4A
chr5A
chr6A
chr7A
chr1B
chr2B
chr3B
chr4B
chr5B
chr6B
chr7B
chr1D
chr2D
chr3D
chr4D
chr5D
chr6D
chr7D

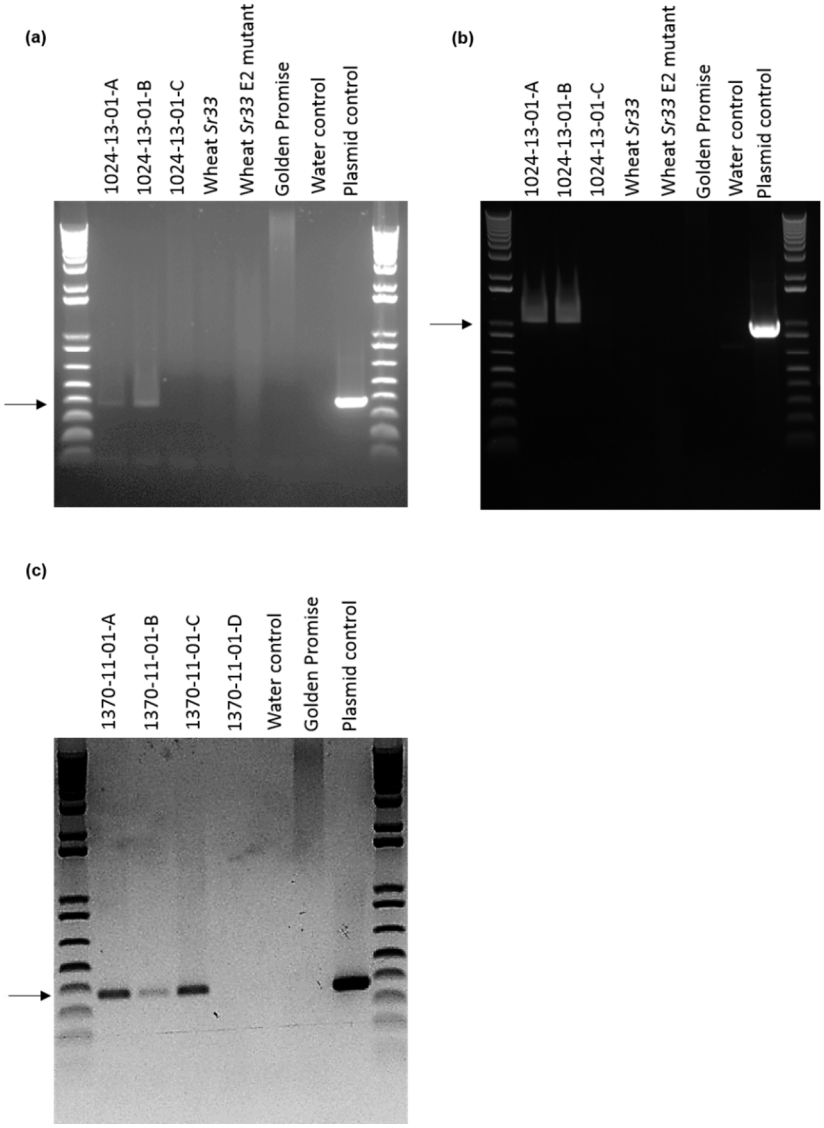
d)


chr1A
chr2A
chr3A
chr4A
chr5A
chr6A
chr7A
chr1B
chr2B
chr3B
chr4B
chr5B
chr6B
chr7B
chr1D
chr2D
chr3D
chr4D
chr5D
chr6D
chr7D

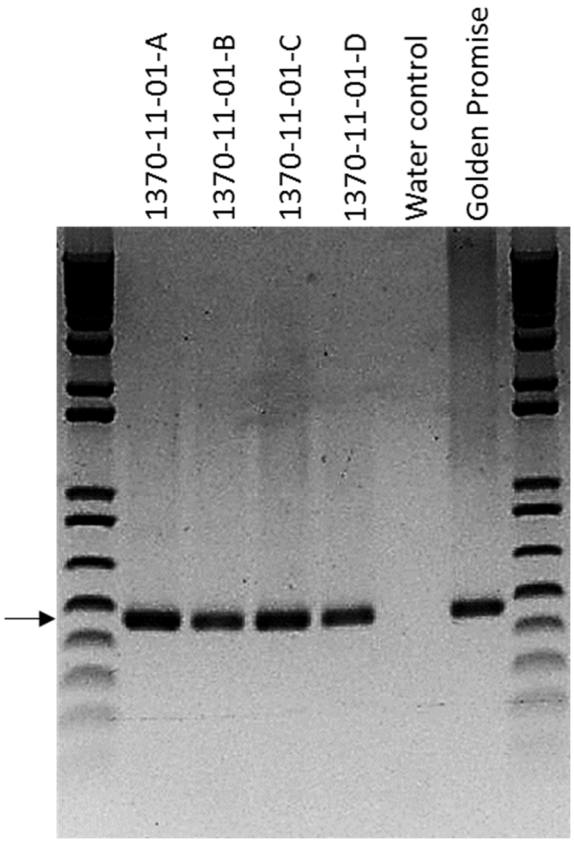

e)
f)
g)



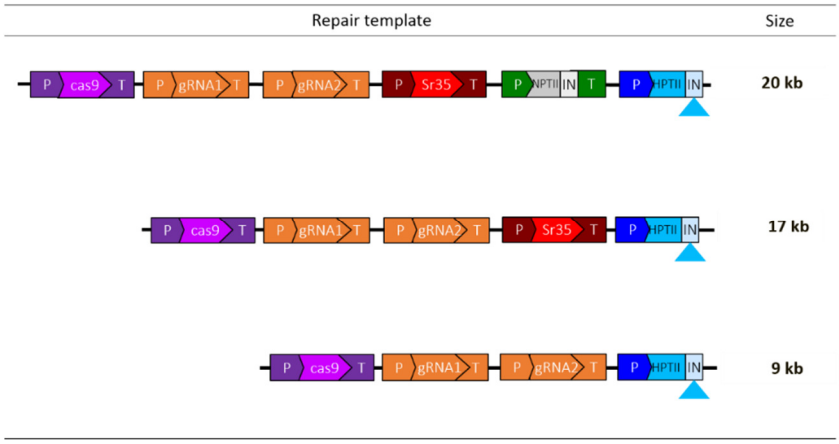
Supplementary Figure 3 GBS analysis on *Sr44* (a) mutant M3 (b) mutant M4 (c) mutant M5 (d) mutant M6 (e) mutant M7 (f) mutant M8 (g) mutant M9 (h) mutant M10.



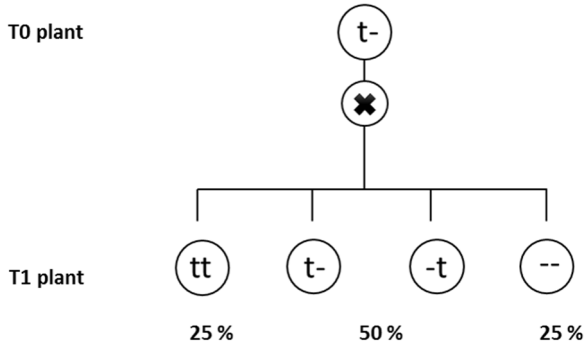
Supplementary Figure 4 Resistance to MCCFC in barley *Sr33* transgenic T_2 families segregating for the transgene correlates with the presence of the transgene. PCR amplification products of the transgenes using (a) Amplification products of the *HPTII* transgene specific marker (Partial_Hyg_F and Partial_Hyg_R) using genomic DNA (*Sr33* transgenics and wheat controls) and plasmid DNA (positive control) as a template. The amplification product of 373 bp (arrowed) is present only in the *HPTII* postulated genotypes (1024-13-01-A and 1024-13-01-B). (b) Amplification products of the *NPTII* transgene specific marker (NPTII_F and NPTII_R) using genomic DNA (*Sr33* transgenics and wheat controls) and plasmid DNA (positive control) as a template. The amplification product of 1004 bp (arrowed) is present only in the *NPTII* postulated genotypes (1024-13-01-A and 1024-13-01-B). (c) Amplification products of the *HPTII* transgene specific marker (Partial_Hyg_F and Partial_Hyg_R) using genomic DNA (*Sr22* transgenics) and plasmid DNA (positive control) as a template. The amplification product of 373 bp (arrowed) is present only in the *HPTII* postulated genotypes (1370-11-01-A, 1370-11-01-B, and 1370-11-01-C).



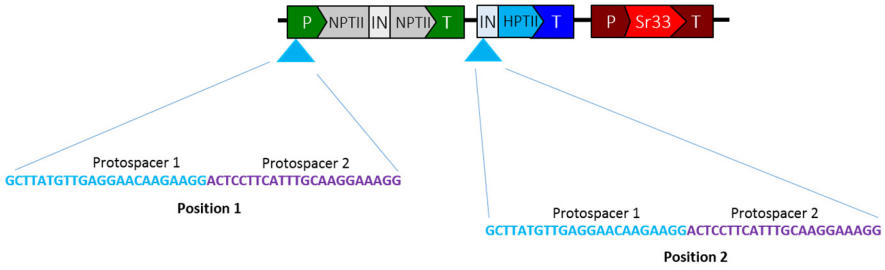
Supplementary Figure 5 Amplification of the barley endogenous *CONSTANS* gene using genomic DNA of *Sr22* transgenics (1370-11-01-A, B, C, D) and Golden Promise (positive control) as a template. The amplification product of ~480 bp (arrowed) is present in all samples except water control.



Supplementary Figure 6 Different version of repair template with approximate size of T-DNA.



Supplementary Figure 7 Transgene inheritance in barley transgenic. Selfing of primary transformants (T_0) results in 25% of T_1 transformants carrying homozygous transgene, 50% carrying hemizygous transgene, and 25% carrying null transgene.



Supplementary Figure 8 Position of protospacer 1 and 2 in landing pad and possible removal of repaired *HPTII* gene after editing event. First excision might occur between protospacer 2 in position 1 and protospacer 1 in positon 2, resulting in deletion of the *NPTII* gene and repair of *HPTII* (the other half of the repair template) by NHEJ-mediated repair. Then, second excision might occur between protospacer 1 in position 1 and protospacer 2 in position 2, resulting in deletion of the repaired *HPTII* gene.

Supplementary Table 1 Stem rust infection assays with *Pgt* race TTKSK on *Sr44* M_2 and M_3 families.

Sequential Number	Plant ID	ID	M2 Reaction	Plant 1	Plant 2	Plant 3
33	a	BW_07362	0;	0;	;1	;1
	d		3	3+	3+	4
	g		;1	;1	1	;1
40	a	BW_07370	2	1+	1+,2-	NA
	d		0;	1	1	1
	e		3	2+	2	2
	f		4	3	3+	3
	g		2-	2	1++	1+
56	a	BW_07386	;1	2	1	1
	d		4	3	3+	3
	g		1	1	1	1
67	b	BW_07397	0;	NA	NA	NA
	d		;1	;1	;1	1
	e		4	4	4	4

155	a	BW_07494	3	3+	4	4
	e		;1	;1	1	1
	h		;1	1	;1	1
176	a	BW_07535	0;	NA	NA	NA
	d		3+	4	4	3+
	e		;1	NA	NA	NA
222	a	BW_07586	1+	NA	NA	NA
	d		;1	NA	NA	NA
252	a	BW_07628	0;	0;	0;	;1
	b		0;	0;	0;	0;
	e		4	3	3+	3
294	b	BW_07628	;1	NA	NA	NA
	c		0;	NA	NA	NA
302	c	BW_07684	;1	1	;1	1
	d		;1	;1	1	1
	e		3	2++	2+,3-	3--
353	a	BW_07748	3	3	3+	3
	b		0;	0;	;1	;1
	g		1	1	1	1
398	a	BW_07797	too small	0;	NA	NA
	c		;1	;1	1	;1
	f		;1	;1	1	1
	h		3+	3	3+	3
414	b	BW_07822	1	;1	;1	1
	c		0;	;1	1	;1
	d		2+,3-	2+	2	1++
	e		3+	2+	2+,3-	2+,3-
	f		2	2+,3-	2+	3
	g		too small	;1	1	;1
484	a	BW_07910	3+	2+,3-	3	3--
	c		1	1	;1	1+
	d		1	1	1	1
	e		;1	1	1	1
498	a	BW_07925	3	3	3+	3
	c		3+	3	3+	2++,3-
	d		;1	;1	1	1
	e		1	1	1	1
	f		;1	1	1	1
	g		;1	1	1	1
513	a	BW_07910	0;	3	3+	2
	b		2+	3	3	NA
	c		0;	0;	0;	;1
	d		2+	1++	2+,3-	2
	g		1	1	1	1
559	b	BW_07996	3+	2+, (3--)	NA	NA
	e		0;	1	1	1
	g		1	1	1	1

560	c	BW_07997	0;	0;	0;	0;
	e		4	2+,3-	2++	3
	h		;1	1	1	1
627	a	BW_07997	3	2+,3-	2	2
	b		too small	NA	NA	NA
	e		too small	NA	NA	NA
	f		0;	0;	0;	0;
	g		1 (c)	;1	1	1
674	a	BW_08139	1	;1	1	1
	b		1	1	1	1
	e		;1	NA	NA	NA
	e		3+	NA	NA	NA
754	a	BW_08228	1	0;	0;	;1
	b		1	1	1	1
	c		3 (4 tip)	3	3+	3
768	a	BW_08245	3	3	3+	3
	f		2+	2+,3-	2	3--
	g		0;	1++	1+,2-	1
1050	a	BW_08680	1	NA	NA	NA
	b		1+	NA	NA	NA
	c		1	NA	NA	NA
	d		1+	NA	NA	NA
	e		1	0;	0;	0;
	f		3	2+	NA	NA
	g		1+	NA	NA	NA
	h		1	NA	NA	NA
1061	a	BW_08692	1	0;	0;	0;
	b		3	NA	NA	NA
	c		1+	1	1	1
	d		1	1	1	1
	e		1	1	1	1
	f		1+	1	1	1
	g		2	1	1+	1
	h		NA	NA	NA	NA

Supplementary Table 2 DNA concentration of *Sr44* susceptible mutants and wildtype.

Mutant	Plant	ID	Concentration (ng/µL)[a]	260:280 ratio
M2	56d	BW_07386	408	1.78
M3	155a	BW_07494	562	1.84
M4	176d	BW_07535	217	1.81
M5	252e	BW_07628	511	1.81
M6	353a	BW_07748	216	1.77
M7	398h	BW_07797	646	1.86
M8	498a	BW_07925	432	1.76
M9	754c	BW_08228	1474	1.82
M10	768a	BW_08245	373	1.75
Sr44 WT			944	1.86

[a]Concentration was measured using a Nanodrop

Supplementary Table 3 Stem rust infection assays with *Pgt* race TTKSK on sister spike of *Sr44* M_2 families.

Mutant	Spike ID	Sister spike ID	Line	No. of seeds available	Susceptible (%)	Plant															
						1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
M2	BW_07386	BW_07386_2	Sr44_M2	8	0	1	2+	2	2-	1+	1	1-	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
		BW_07386_3	Sr44_M2	7	0	2	2	2-	2	2-	1+	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
		BW_07386_bulked	Sr44_M2	5	0	1	1+	1	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
M3	BW_07494	BW_07494_2	Sr44_M2	19	11.7	1-	3	2+	2-	2-	2-	2	2+	2+	2	1	1+	3+	2	1+	2	1+
M4	BW_07535	BW_07535_2	Sr44_M2	5	0	1+	1+	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
		BW_07535_3	Sr44_M2	6	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
		BW_07535_bulked	Sr44_M2	6	0	1-	1	1+	1-	1-	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
M5	BW_07628	BW_07628_1	Sr44_M2	2	0	1-	1-	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
M6	BW_07748	BW_07748_2	Sr44_M2	41[a]	0	1	1	1+	1+	1	1	1-	1-	1	1	2-	1	1-	1	1-	1-	1
M7	BW_07797	BW_07797_1	Sr44_M2	4	0	1-	1-	1-	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
M8	BW_07925	BW_07925_bulked	Sr44_M2	8	0	1	1+	1-	1	1+	1	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
M9	BW_08228	BW_08228_bulked	Sr44_M2	9	0	2+	1+	1	1	1-	2	1-	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
M10	BW_08245	BW_08245_bulked	Sr44_M2	44	0	1	2	1+	1-	1	1-	1-	1	1-	2-	1-	2	1	1	1-	1	1-
	McNair		Susceptible check		100	4	4	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Sr44		Resistant check		0	1-1	1-	1-	11-	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Angas		Susceptible check		100	3+	4	3+	4	4	4	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Supplementary Table 4 Stem rust infection assays with *Pgt* race MCCFC on *Sr22* T_2 families.

Plant ID	Line	Plant A	Plant B	Plant C	Plant D	Plant E	Plant F	Plant G	Plant H
1370-01-01	Sr22_T1	0;	0;	0;	0;	0;	0;	0;	0;
1370-08-01	Sr22_T1	3-2	3-2	3-2	3-2	2	3-2	3-2	3-2
1370-11-01	Sr22_T1	3	3	0;	0;	0;	0;	0;	0;
1370-13-01	Sr22_T1	3-2	3-2	3-2	3-2	3-	3-	3-	3-
1370-17-01	Sr22_T1	3	3	0;	0;	3-	3-	3-2	0;
1370-19-01	Sr22_T1	3	3-	3-2	3	3-	2-2	0	0
1372-05-01	Sr22_T1	3-2	3	3	3	3-	3-	3-	3
1372-07-01	Sr22_T1	3	3	3	3-	3	3	3	3
1372-08-01	Sr22_T1	3-	0	0	0	0	0	0	0
1372-19-01	Sr22_T1	3	3	3	3	NA	3	3	3
GP-2015	Golden Promise	3-2	NA	NA	NA	NA	NA	NA	NA
GP-Saint Paul 2012	Golden Promise	3-2	NA	NA	NA	NA	NA	NA	NA
McNair	Susceptible check	4	NA	NA	NA	NA	NA	NA	NA

Supplementary Table 5 Stem rust infection assays with *Pgt* race MCCFC on *Sr33* T_2 families.

Plant ID	Line	Rep 1								Rep 2							
		Plant A	Plant B	Plant C	Plant D	Plant E	Plant F	Plant G	Plant H	Plant I	Plant J	Plant K	Plant L	Plant M	Plant N	Plant O	Plant P
1023-02-01	Sr33_T1	$2^{+}3^{-}$	1^{+}	2^{+}	;1	1^{+}	$1^{+}2^{-}$	2	2	$2^{-}1^{+}$	1^{+}	1	1	2	1	1^{+}	2
1023-06-01	Sr33_T1	1	;1	$2^{+}3^{-}$	2	1	;1	1^{+}	1	;1 (1^{+} tip)	0;	0;	0;	12	0;	1	$1^{+}2^{-}$
1023-07-03	Sr33_T1	1	1	1^{+}	1	2	1	$1^{+}2^{-}$	1	1	1^{+}	2	0;	1	2^{+}	$2^{+}3^{-}$	0;
1023-14-01	Sr33_T1	2	2^{+}	$1^{+}2^{-}$	2	2	$1^{+}2^{-}$	1^{+}	1^{+}	1^{+}	$1^{+}2^{-}$	2	2^{-}	2^{+}	2	$1^{+}2^{-}$	1^{+}
1024-01-01	Sr33_T1	2	1^{+}	1^{+}	12	1	$1^{+}2^{-}$	12	1	1	$1^{+}2^{-}$	2^{+}	1^{+}	12	$1^{+}2^{-}$	2	2
1024-04-02	Sr33_T1	;1	2^{-}	;1	1	;1	1	1^{+}	1^{+}	0;	1^{+}	2^{-}	1	1	1^{+}	1	1
1024-06-01	Sr33_T1	1	$2^{+}3^{-}$	1	;1	1	;1	1	$1^{+}2^{-}$	0;	0;	2^{+}	1^{+}	2^{+}	;1	$1^{+}2^{-}$	12
1024-07-01	Sr33_T1	2	1	1	$2^{-}1^{+}$	;1	23^{-}	1	1	1	1	;1	12	$2^{+}3^{-}$	;1	2	1
1024-11-01	Sr33_T1	2	1	1	;1	$2^{+}3^{-}$	$2^{-}1^{+}$	1	1	0;	0;	2	NA	;1	1^{+}	1	;1
1024-13-01	Sr33_T1	0;	0;1	32^{-}	2	0;	0;	0;	0;	;1	32^{-}	0;	0;	1	1^{+}	$1^{+}2^{-}$	1
1033-05-01	Sr33_T1	0;	0;	;1	0;	0;	0;	1^{+}	0;	$1^{+}2^{-}$	32^{-}	0;	0;	1	1	;1	1
1046-02-01	Sr33_T1	$3^{-}2^{+}$	$3^{-}2^{+}$	3	$3^{-}2^{+}$	3	$3^{-}2^{+}$	3	3	1	1^{+}	2		$1^{+}2^{-}$	2	2	2^{+}
GP-2015	Golden Promise	$3^{-}2^{+}$	$3^{-}2^{+}$	NA	NA	NA	NA	NA	NA	$1^{+}2^{-}$	3	2	2	2^{-}	$1^{+}2^{-}$	$3^{-}2^{+}$	2^{-}
TA_01046	Sr33	NA	NA	NA	NA	NA	NA	NA	NA	1 C	1 C	1^{+}	1	1^{+}	1^{+}	1	1
TA_01048	Sr33_E2	NA	NA	NA	NA	NA	NA	NA	NA	3	3^{+}	4	3	3^{+}	3	3^{+}	3
TA_01056	Sr33_E6	NA	NA	NA	NA	NA	NA	NA	NA	1^{+}	2^{+}	3^{--}	1	1^{+}	$1^{+}2^{-}$	$2^{-}1^{+}$	1^{+}

Supplementary Table 6 Stem rust infection assays with *Pgt* race MCCFC on *Sr45* T_2 families.

Plant ID	Line	Plant A	Plant B	Plant C	Plant D	Plant E	Plant F	Plant G	Plant H
1613-01-01	Sr45_T1	32	23	23	23	23	23	23	NA
1613-02-01	Sr45_T1	0	23	23	0	0	0	23	23
1613-03-01	Sr45_T1	23	23	32	23	32	32	32	32
1613-04-01	Sr45_T1	0	0	2	0	0	0	0	0
1613-05-01	Sr45_T1	0	23	0	0	0	2	0	0
1613-09-01	Sr45_T1	23	23	2	2	2	2	2	2
1613-10-01	Sr45_T1	23	23	23	23	2	2	23	2
1613-11-01	Sr45_T1	2	2	23	23	32	32	32	23
1613-14-01	Sr45_T1	2	2	2	2	23	2	2	2
1613-17-01	Sr45_T1	0	0	0	0	0	23	0	NA
1613-19-01	Sr45_T1	23	23	23	23	2	2	23	2
GP-2015	Golden Promise	32	32	3	3	3	32	23	2
McNair	Susceptible control	3	3^+	3^+	3	3^+	NA	NA	NA
Q21861	Resistant control	0	0	0	0	0	NA	NA	NA

Supplementary Table 7 PCR primers used in this study.

Name	Sequence 5' to 3'	Product size (bp)
NPTII_F	GAG GCT ATT CGG CTA TGA CTG G	1004
NPTII_R	ATC GGG AGC GGC GAT ACC GTA	
Partial_Hyg_F	GCA AAC TGT GAT GGA CGA CA	373
Partial_Hyg_R	TGC ATC ATC GAA ATT GCC GT	
Sr35_F2	AGG ACA TCG TTG ATG CCT TC	391
Sr35_R2	GTC TTG CCT AAC CCA CCA AA	
Sr22_F	AAT ACA GCC CGG CAA CAT AG	309
Sr22_R	TCC AAC TGA TCG CAG TCT TG	
Sr45_F	GGG AGA TCT ACC GTC ACT GG	362
Sr45_R	AAC TGT GTA CGC GGT CAC TA	
CON2F1	ATT GTG CCA ACA AGA TAG ATC G	480
CON2R1	AAA GGC AAA TAA TCT GGT CTG C	
S45F1	AGT ACT GTA ATA ATT GAT TCC GTC G	6481
S45R5	GAA ATT CCT GCT GCA TTG C	

Supplementary Table 8 List of binary constructs carrying *Sr* gene.

Name	Gene	Binary vector	Bacterial Resistance	Plant Resistance	Regulatory elements
pBW_0065	*Sr22*	pVec8	Spectinomycin	*HPTII*	Native
pBW_0041	*Sr33*	pAGM4723	Kanamycin	*NPTII*	Native
pBW_0059	*Sr35*	pAGM4723	Kanamycin	*HPTII*	Native
pBW_0141	*Sr45*	pVec8	Spectinomycin	*HPTII*	*Sr33* promoter and terminator
PC110	*Sr45*	pVecBARII	Spectinomycin	Phosphinothricin	Native
PC147	*Sr45*	pVecBARII	Spectinomycin	Phosphinothricin	*Sr33* promoter and terminator

Supplementary Table 9 Leaf rust infection assays with race 4 on *Sr33* T_2 families.

Line ID	Plant ID	Line	Infection type
1023-02-01	D	Sr33_T1_leafrust	2+3-
1023-02-01	E	Sr33_T1_leafrust	22+,3-
1023-02-01	F	Sr33_T1_leafrust	22+,3-
1023-02-01	G	Sr33_T1_leafrust	22+,3-
1023-06-01	A	Sr33_T1_leafrust	22+,3-
1023-06-01	B	Sr33_T1_leafrust	2+,3-
1023-06-01	C	Sr33_T1_leafrust	NA
1023-06-01	D	Sr33_T1_leafrust	22+
1023-06-01	F	Sr33_T1_leafrust	NA
1024-04-02	A	Sr33_T1_leafrust	2+3
1024-04-02	B	Sr33_T1_leafrust	2+3
1024-04-02	C	Sr33_T1_leafrust	2+3-

1024-04-02	E	Sr33_T1_leafrust	22+
1024-06-01	B	Sr33_T1_leafrust	NA
1024-06-01	D	Sr33_T1_leafrust	23-,1-0;
1024-06-01	F	Sr33_T1_leafrust	22+,3,1+
1024-06-01	H	Sr33_T1_leafrust	213-
1024-07-01	B	Sr33_T1_leafrust	213-
1024-07-01	D	Sr33_T1_leafrust	22+,13-
1024-07-01	E	Sr33_T1_leafrust	22+,3-
1024-07-01	F	Sr33_T1_leafrust	22+,3-
1024-11-01	D	Sr33_T1_leafrust	22+,3-
1024-11-01	E	Sr33_T1_leafrust	22+,3-
1024-11-01	B	Sr33_T1_leafrust	22+,3-1
1024-11-01	A	Sr33_T1_leafrust	22+,3-1
1024-13-01	A	Sr33_T1_leafrust	22+
1024-13-01	B	Sr33_T1_leafrust	22+1
1024-13-01	C	Sr33_T1_leafrust	2+3-
1024-13-01	D	Sr33_T1_leafrust	22+,3-
1024-13-01	E	Sr33_T1_leafrust	NA
1033-05-01	A	Sr33_T1_leafrust	22+,3-
1033-05-01	B	Sr33_T1_leafrust	22+,3-
1033-05-01	C	Sr33_T1_leafrust	22+,3-
1033-05-01	G	Sr33_T1_leafrust	22+,3-
Moore		Susceptible check	2+,3-
PI531901-4		Resistant check	0;1-
Golden Promise			3-,2+

Supplementary Table 10 List of constructs used in this study.

ID	Common name	Backbone vector	Description	Bacterial resistance	Plant resistance
pB214::ZmUbi::NPTII	Ubi-kan	pB214	ZmUbi-P_NPTII-CR_35S-T::35S-P_HPTII-intron-CR_Nos-T	Kanamycin	*NPTII*
pB214::ZmUbi::NPTII-intron	Ubi-kan-intron	pB214	ZmUbi-P_NPTII-intron-CR_35S-T::35S-P_HPTII-intron-CR_Nos-T	Kanamycin	*NPTII*
pBW_0036	L1P2 kan-intron	pICH47742	OsActin-P_NPTII-intron-CR_35S-T	Carbenicillin	*NPTII*
pBW_0143	Cas9-YFP	pICH47732	ZmUbi-P_Cas9-CR-YFP-C terminal tag_Nos-T	Carbenicillin	NA
pBW_0140	Non-binary Cas9-gRNA1	pICSL22055	ZmUbi-P_Cas9-CR_Ocs-T::TaU6-P_gRNA1	Spectinomycin	NA
pBW_0032	L1P4 Sr33	pUC57	Sr33-P_Sr33-CR_Sr33-T	Carbenicillin	NA
pBW_0034	L0 kan-intron	pICH41308	NPTII-intron	Spectinomycin	NA
pBW_0035	L0 hyg-intron	pICH41308	HPTII-intron (+2 CRISPR sites in intron)	Spectinomycin	NA
pBW_0036	L1P2 kan-intron	pICH47742	OsActin-P_NPTII-intron-CR_35S-T	Carbenicillin	*NPTII*
pBW_0037	L1P1 full Hyg-intron	pICH47732	35S-P_HPTII-intron-CR_Nos-T (+2 CRISPR sites in intron)	Carbenicillin	*HPTII*
pBW_0038	L1P3 partial Hyg-intron	pICH47751	partial HPTII-intron-CR_Nos-T (+2 CRISPR sites in intron)	Carbenicillin	NA
pBW_0040	L1P1 sgRNA	pICH47732	protospacer1-protospacer2	Carbenicillin	NA
pBW_0044	L1P1 ZmUbi10-Cas9	pICH47732	ZmUbi-P_Cas9-CR_Ocs-T	Carbenicillin	NA
pBW_0045	L1P2 gRNA1	pICH47742	TaU6-P_gRNA1	Carbenicillin	NA
pBW_0046	L1P3 gRNA2	pICH47751	TaU6-P_gRNA2	Carbenicillin	NA
pBW_0048	L1P5 partial Hyg-intron	pICH47772	double CaMV 35S-P_partial HPTII-intron-CR	Carbenicillin	NA
pBW_0041	L2 landing pad (kan-int)	pAGM4723	gRNA1-gRNA2::OsActin-P_NPTII-intron-CR_35S-T::partial HPTII-intron-CR_Nos-T::Sr33	Kanamycin	*NPTII*
pBW_0057	L2 repair template	pAGM4723	ZmUbi-P_Cas9-CR_Ocs-T::TaU6-P_gRNA1::TaU6-P_gRNA2::dummy::double CaMV 35S-P_partial HPTII-intron NGG-CR	Kanamycin	NA
pBW_0139	L2 repair template gRNA1	pAGM4723	ZmUbi-P_Cas9-CR_Ocs-T::TaU6-P_gRNA1::dummy::dummy::double CaMV 35S-P_partial HPTII-intron NGG-CR	Kanamycin	NA

Supplementary Table 11 List of primers and PCR conditions.

Name	Gene to amplify	Sequence 5' to 3'	Thermal cycles	Product size (bp)
NPTII_F NPTII_R	*NPTII*	GAGGCTATTCGGCTATGACTG G ATCGGGAGCGGCGATACCGTA	1 cycle: 5min 94˚C 35 cycles: 30s at 94˚C, 30s at 60˚C, 72˚C at 90s 1 cycle: 72˚C at 10min	1,004 (including intron)
Partial_Hyg_F Partial_Hyg_R	*HPTII*	GCAAACTGTGATGGACGACA TGCATCATCGAAATTGCCGT	1 cycle: 5min 95˚C 33 cycles: 30s at 94˚C, 30s at 54˚C, 72˚C at 30s 1 cycle: 72˚C at 5min	373
VIRD2 F VIRD2 R	*virD2*	TCAAGTAATCATTCGCATTGTGCC GCCGTGACGAAGTGAAATCTC	30 cycles: 30s at 94˚C, 1min at 55˚C 1 cycle: 1min at 72˚C	487
F_del_LP R_del_LP	deleted *NPTII*	ACTGGGGTTGAACACTCTGT GACCGGCTGCAGTTATCATC	1 cycle: 30s 98˚C 30 cycles: 10s at 98˚C, 30s at 62˚C, 72˚C at 2min 1 cycle: 72˚C at 7min	192 (deletion) 3,051 (no deletion)
Hyg_int_span_For Hyg_int_span_Rev_	*HPTII*	AATTCAGCGAGAGCCTGACC CCGTCAGGACATTGTTGGAG	1 cycle: 5min 95˚C 33 cycles: 30s at 90˚C, 40s at 61˚C, 72˚C at 30s 1 cycle: 72˚C at 5min	588
For Kan_Int_Frag1 Rev Kan_Int_Frag1	domesticated *NPTII*	AGAAGACAAAATGATTGAACAAGATGGATTGCACGC AGAAGACAATAGATCTAAGATTAACAGAATCTAAACC	1 cycle: 3min 95˚C 33 cycles: 30s at 95˚C, 40s at 57˚C, 72˚C at 1min 1 cycle: 72˚C at 5min	650
For Kan_Int_Frag2 Rev Kan_Int_Frag2	domesticated *NPTII*	AGAAGACAATCTAAGACGATTTTCTGGGTTTGATCGTTAG AGAAGACAAAAGCTCAGAAGAACTCGTCAAGAAGGCGATAG	1 cycle: 3min 95˚C 33 cycles: 30s at 95˚C, 40s at 57˚C, 72˚C at 1min 1 cycle: 72˚C at 5min	500

For_hyg_int Rev_hyg_int	*HPTII*	AGAAGACAACAGCCGGTCGCGGAGGCCATGGATGCGATCG AGAAGACAAACCTGCAGAACAGCGGGCAGTTCGGTTTCAGGC	1 cycle: 30s 98˚C 30 cycles: 10s at 98˚C, 30s at 72˚C, 72˚C at 2min 1 cycle: 72˚C at 7min	3,500
For_partial_hyg_int Rev_partial_hyg_int	partial *HPTII*	AGGTCTCAGGAGAAATTTCTAGTTTTGCTTATGTTGAG AGGTCTCAAGCGTCGATCTAGTAACATAGATGACAC	1 cycle: 30s 98˚C 30 cycles: 10s at 98˚C, 30s at 72˚C, 72˚C at 1min 1 cycle: 72˚C at 3min 30s	1,200
gRNA1_F gRNA1_R	sgRNA1	TGTGGTCTCACTTGCTTATGTTGAGGAACAAGAGTTTTAGAGCTAGAAATAGCAAG TGTGGTCTCAAGCGTAATGCCAACTTTGTAC	1 cycle: 30s 98˚C 30 cycles: 10s at 98˚C, 30s at 72˚C, 72˚C at 30s 1 cycle: 72˚C at 30s	200
gRNA2_F gRNA1_R	sgRNA2	TGTGGTCTCACTTGACTCCTTCATTTGCAAGGAAGTTTTAGAGCTAGAAATAGCAAG TGTGGTCTCAAGCGTAATGCCAACTTTGTAC	1 cycle: 30s 98˚C 30 cycles: 10s at 98˚C, 30s at 72˚C, 72˚C at 30s 1 cycle: 72˚C at 30s	200
For_partial_hyg_int_5’ 2Rev_partial_hyg_int_5’_NGG	partial *HPTII*	AGGTCTCAGGAGGTCAACATGGTGGAGCACGACACTC AGGTCTCAAGCGCCTTTCCTTGCAAATGAAGGAGTCC	1 cycle: 30s 98˚C 30 cycles: 10s at 98˚C, 30s at 72˚C, 72˚C at 1min 1 cycle: 72˚C at 3min 30s	1,300

Supplementary Table 12 Copy number analysis using q-PCR.

Sample	*NPTII* copies
01033-07-01	25
01033-03-02	14
01033-08-04	12
01033-08-01	11
01023-11-02	10
01033-02-02	9
01023-08-03	8
01024-02-01	7
01047-12-01	7
01024-12-01	7
01023-05-01	6
01023-11-01	6
01024-10-01	6
01023-01-01	6
01023-04-01	6
01033-03-01	5
01024-03-01	4
01047-14-01	4
01047-10-01	4
01024-01-02	4
01023-07-02	3
01033-04-01	3
01033-09-01	3
01023-02-02	2
01024-09-01	2
01023-05-01	2
01023-10-02	2
01023-10-01	2
01033-02-01	2
01023-10-03	2
01023-02-04	2
01047-13-01	2
01023-08-01	2
01023-12-01	2
01047-11-01	2
01023-14-04	1 to 2 (prob. 1)
01024-07-01	1
01024-11-01	1
01033-05-01	1
01023-02-03	1
01024-13-01	1
01023-07-03	1

01023-06-01	1
01024-04-02	1
01024-06-01	1
01023-02-01	1
01024-01-01	1
01024-08-01	1
01032-03-01	1
01047-09-01	1
01024-04-03	<1
01024-04-01	0
01046-01-01	0
01046-03-01	0
01023-03-01	0
01023-07-04	0
01023-15-01	0
01024-05-01	0
01047-04-01	21
01024-06-02	11
01032-02-02	8
01024-16-01	7
01024-08-02	6
01047-01-01	6
01023-09-01	5
01047-08-01	5
01046-10-01	5
01046-13-01	5
01046-08-02	4
01047-05-01	4
01046-07-01	4
01046-05-01	4
01024-16-02	4
01047-03-01	4
01047-07-01	4
01032-01-01	4
01023-09-02	3
01046-12-01	3
01046-11-01	3
01024-04-04	3
01046-16-02	2 to 3
01046-16-01	2 to 3 (prob. 2)
01046-06-01	2
01031-02-01	2
01046-17-01	2
01024-17-01	2
01033-01-01	1

01046-14-01	1
01046-08-02	1
01023-08-02	1
01046-15-02	1
01046-15-01	1
01031-01-01	1
01023-14-03	1
01023-13-01	1
01046-02-02	1
01046-02-01	1
01032-02-01	1
01023-14-01	1
01023-14-02	1
01047-02-01	1
01047-06-01	1
01046-01-02	0
01024-15-01	0
01046-04-01	0
01046-09-01	0
01116-12-02	16
01088-02-01	16
01116-12-01	16
01115-24-03	16
01115-17-01	7
01116-08-01	7
01116-01-03	6
01115-11-01	6
01116-19-01	5
01116-01-04	5
01116-01-01	5
01115-06-01	5
01088-01-01	4 to 5
01089-01-01	4
01115-01-01	4
01115-24-02	4
01089-01-02	4
01088-01-02	3
01115-16-03	3
01115-15-01	3
01115-16-04	2
01116-18-01	2
01088-03-01	2
01089-03-02	2
01088-04-03	2
01089-05-01	2

01115-02-01	2
01088-02-02	2
01088-04-04	2
01116-02-01	2
01088-04-01	2
01089-03-01	2
01089-06-02	2
01116-09-01	2
01088-04-02	2
01115-12-01	2
01089-04-01	2
01116-21-03	2
01115-13-02	2
01115-03-03	2
01089-06-01	2
01116-10-01	2
01115-19-02	2
01115-23-03	2
01116-11-03	2
01116-15-01	2
01116-04-01	2
01115-22-02	2
01115-20-01	1 to 2
01115-03-02	1
01089-07-01	1
01116-05-01	1
01116-06-01	1
01115-13-01	1
01089-08-01	1
01089-02-01	1
01115-04-03	1
01115-04-01	1
01115-14-02	1
01115-03-01	1
01116-19-02	1
01116-21-02	1
01116-17-01	1
01116-13-01	1
01116-11-01	1
01115-23-01	1
01116-16-01	1
01116-14-01	1
01116-21-01	1
01116-12-03	1
01115-18-01	<1

01115-21-01	<1
01115-22-01	0
01088-05-01	0
01115-04-02	0
01115-05-01	0
01115-08-01	0
01115-16-01	0
01115-16-02	0
01115-24-01	0
01115-24-04	0
01116-01-02	0
01116-07-01	0
01116-11-02	0
01116-20-01	0