USING NEXT-GEN SEQUENCING TO MAP UG99 STEM RUST RESISTANCE GENES IN WHEAT BIPARENTAL AND NESTED ASSOCIATION MAPPING POPULATIONS

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

PRABIN BAJGAIN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

JAMES A. ANDERSON, ADVISER

October 2014

Acknowledgements

I would like to express my sincere appreciation towards everyone who assisted me with the projects that make up this thesis. I especially thank my adviser Dr James Anderson for his guidance, mentorship, and support throughout these years. I greatly appreciate his constant care and concern towards my personal as well as professional progress; and am forever indebted for the opportunity he provided me to work with him.

I wish to extend my gratitude to the members of my graduate research committee, Dr Brian Steffenson, Dr Kevin Smith, Dr Matthew Rouse, and Dr Robert Stupar for providing thoughtful suggestions regarding my research. I am particularly grateful towards Dr Matthew Rouse for coordinating the planting of my materials in East African stem rust nurseries, and for recording phenotypic data when I was unable to travel.

I also thank the members of the Wheat Breeding Project, Brian Seda, Kathryn Turner, Emily Conley, Catherine Springer, Jen Flor, Susan Reynolds, Roger Caspers, Dr Godwin Macharia, Dr Itamar Nava, Dr Liangliang Gao, and Dr Xiaofei Zhang and all undergraduate students (especially Brittney Johnson) associated with the breeding project from Fall 2011 to 2014. Their technical assistance and advice, as well as the friendship have been highly valuable in completing my research work.

Thanks to Justin Anderson, Lian Lian, Patrick Ewing, and Ahmad Sallam for being excellent officemates, and to fellow APS graduate students for many fruitful discussions and the camaraderie. I thank the faculty and staff in the Department of

Agronomy and Plant Genetics, whose academic knowledge, and research expertise were of great resources to me.

Lastly, I thank the Cereal Disease Laboratory, the University of Minnesota Genomics Center, Minnesota Supercomputing Institute, the Microbial & Plant Genomics Institute, the University of Minnesota Graduate School, Kenya Agricultural Research Institute, Ethiopia Institute of Agricultural Research, and the Agricultural Research Council of South Africa for their help and support during various phases of the study. Funding for most of the work discussed herein was provided by the United States Department of Agriculture, Agriculture and Food Research Initiative and 2011-68002-30029 (Triticeae Coordinated Agricultural Project, www.triticeaecap.org), and the Borlaug Global Rust Initiative Durable Rust Resistance in Wheat Project (administered by Cornell University with a grant from the Bill & Melinda Gates Foundation, and the UK Department for International Development).

Dedication

I dedicate this thesis, with much humility and love, to my parents, Prabha and Purushotam, who while living half the circumference of the earth away, always believed in me, and supported my endeavors. Apart from the encouragement they provided to pursue the highest academic degree possible, they have inspired me to always be patient and humble; and practice kindness and compassion with all living beings, which I believe, define a person more so than just formal education.

Table of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Figures

Chapter 2

Chapter 3

Chapter 4

List of Appendices

Chapter 1

A Historical Account of Resistance Breeding and Genetic Mapping of Stem Rust Resistance in Wheat

Records indicate that the documentation of stem rust of wheat (caused by *Puccinia graminis* Pers. f. sp. *tritici*) likely began with simple field observations, and can reasonably be traced back to the Late Bronze Age II (1400 – 1200 B.C.) (Arthur et al. 1929; Kislev 1982). From field observations in ancient Rome (Fantham 1998) and ancient Greece (Theophrastus) to taxonomical classification of the pathogen (Tulasne and Tulasne 1847; Tulasne 1853; Persoon 1794) and dissection of its life cycle (de Bary 1853), the disease stem rust of wheat as well as the pathogen causing the disease have been thoroughly studied. Rich historical data collected have paved the way for modern epidemiologic and genetic studies, which continue to add to the voluminous nature of research work carried out on this disease.

Stem Rust Epidemics

The destructive nature of stem rust on cultivated wheat has been recorded during epidemics throughout both ancient and modern agricultural periods. The first documented stem rust epidemic came from Europe in 1767 when two Italian scientists, Felice Fontana (1767) and Giovanni Targioni Tozzetti (1767), independently published their observations. Both authors stated that stem rust was prevalent in all wheat growing

regions of Italy, indicative of an epidemic. Other major rounds of severe epidemics in Europe occurred in 1932 in central Europe with recorded yield losses of up to 20%; and in 1951 in Scandinavia, with losses ranging 9-33% (Zadoks 1963). No major outbreaks have been reported in Europe after these reports.

Nearly around the same time the stem rust epidemic was observed in Italy in 1767, wheat stem rust may have also been a big problem in North America (Roelfs 1985a). Wheat rusts were observed in the United States as early as 1660 (Bidwell and Falconer 1925; Fulling 1943). Despite the lack of written evidence of a stem rust epidemic, Alan Roelfs hypothesizes that passing of barberry eradication laws in the Eastern United States in the 1750s and 1760s suggests localized outbreaks of stem rust in those areas. Fulling (1943) states that the wheat production losses in the Eastern United States was connected to the presence of barberry bushes, which led the states of Connecticut, Massachusetts, and Rhode Island to pass barberry eradication laws in 1726, 1764, and 1766, respectively. The first recorded stem rust epidemic in the United States was observed in 1878, affecting Northern Iowa, Southern Minnesota, and Wisconsin (Hamilton 1939). These areas experienced severe yield losses from additional wheat stem rust epidemics in 1904 (Carleton 1905), 1916, 1923, 1925, 1935, 1937, and 1953-1954 (Stakman and Harrar 1957). Of all these epidemics, the 1916 outbreak holds a significant value because of the level of awareness it raised towards resistance breeding and disease management. It was after the 1916 epidemics that barberry eradication laws were passed and implemented on the national scale in an effort to keep the pathogen population in check. This epidemic also led to identification of sources of resistance, and their

introgression into hard red spring bread wheat lines that were later released as varieties (McFadden 1930). Several rust epidemics in Canada coincided with the US outbreaks. Canadian wheat growing regions observed rust epidemics mainly in 1904, 1909, 1916, 1919, 1923, 1925, and 1927.

The rust outbreaks of the mid-1930s and mid-1950s in the United States are attributed to evolution of new pathogen races, with race 56 being prominent in the 1935-1937 epidemics, and race 15B in the 1953-1955 epidemics. No major stem rust outbreak was observed between 1938 and 1952 due to control of the pathogen population aided by barberry eradication in addition to use of resistant cultivars. There have been no widespread stem rust epidemics in the Northern Great Plains after the last epidemic observed during the 1950s, though localized epidemics have been occasionally observed in the southern states (Roelfs 1985a; Brian Steffenson, personal communication). In the past decade and a half, a highly virulent race TTTTF has been observed in North America (Jin 2005). However, this pathogen race does not pose an imminent risk to wheat crop mainly because the resistance genes *Sr24* and *Sr31*, along with other unknown genes safeguard wheat crops from this race (Olson et al. 2010b; Klindworth et al. 2011).

In addition to North American and European epidemics, several other wheat growing regions of the world also have experienced stem rust outbreaks, albeit less frequently and of less damaging nature. Stem rust epidemics in Australia have been observed in 1889 (McAlpine 1906), 1903, 1916 (Waterhouse 1929), 1947 (Butler 1948), and 1974 (Watson 1981), causing severe yield loss. Similarly, epidemics have also been observed in China in 1948, 1951-1952 and 1956 (Roelfs 1977), and intermittently and

less severe, only during unusually warm seasons in India (Joshi and Palmer 1973). The outbreaks in Australia and China are credited mainly to an unusual rise in temperature during those years. Pathogen race survey and identification projects in East Africa suggested that small, scattered wheat growing areas are prone to the attack of the pathogen (Green et al. 1969; Harder et al. 1971). In Kenya, stem rust epidemics might have been observed as early as in the 1910s and 1920s, which led to establishment of rigorous resistance breeding programs (Burton 1928; McDonald 1931).

In summary, the historical data suggest that stem rust of wheat has been a major problem in most wheat growing regions of the world. The data also document effective control of the disease by using a combination of resistance genes and other protective measures such as fungicide application, with the former being the preferred and most effective method. We can project that this disease can be effectively controlled if sources of resistance are used timely and effectively, with regular monitoring of evolution and spread of the pathogen. The latter seems to be of high importance, as manifested by the evolution and dominance of the Ug99 group of races in Africa.

The Emergence of Ug99

After decades of relative absence of epidemic proportions of the disease came the untimely reminder that this historical foe of the wheat crop is not quite ready to give up. An evolved and much more damaging isolate of the pathogen was discovered in Uganda in 1998, and was named Ug99 for the country of origin and the year it was evaluated for its virulence phenotyping (Pretorius et al. 2000). Later, the isolate was designated as race TTKS according to the North American race nomenclature system (Wanyera et al. 2006).

With the addition of four more lines to the standard 16 wheat stem rust differential lines set, race TTKS was subsequently keyed to TTKSK (Jin et al. 2008). At its discovery, isolate Ug99 had overcome the most important stem rust resistance gene of the time: *Sr31*. The gene *Sr31* was transferred to hexaploid wheat from the short arm of chromosome 1 of rye (*Secale cereal*e) (Zeller 1973). This gene had been widely deployed by CIMMYT (The International Maize and Wheat Improvement Center, Mexico) because of its high level of resistance to several diseases, broad adaptation, and higher yield potential (Dubin and Brennan 2009). *Sr31* remained highly effective and durable globally for more than three decades until isolate Ug99 arose. Within a few years of its discovery, variants in the Ug99 lineage were detected that carried virulence to lines with genes *Sr24* and *Sr36* in Kenya in 2006 and 2007, respectively (Jin et al. 2008; Jin et al. 2009b). Originally detected as a single race in East Africa, isolate Ug99 has now evolved to include seven races (TTKSK, TTKST, TTTSK, TTKSF, TTKSP, PTKSK, and PTKST) in its lineage, with demonstrated virulence on a number of important stem rust genes (Singh et al. 2011). These races have spread, as documented by their arrival in Iran in the north and South Africa in the south. The races not only travel long distances, but also evolve further in regions of deposition (Singh et al. 2011; Pretorius et al. 2012). However, the most disturbing aspect of this movement and evolution is that most of the existing wheat varieties and germplasm are susceptible to the Ug99 race group. During 2005-2010, more than 200,000 wheat accessions representing advanced breeding materials from several wheat breeding programs around the world, and germplasm collections were screened in Kenya and Ethiopia (Singh et al. 2006; Singh et al. 2008b).

The results showed that up to 95% of the tested materials are susceptible to the races. This is a terrifying scenario, especially considering that these new virulent African stem rust races are not yet completely under control, and are capable of traveling long distances. It might only be a matter of time before these races make into the bread baskets of Asia, Europe, and the Americas. In an effort to minimize future stem rust losses, the Durable Rust Resistance in Wheat (DRRW) project was funded as a part of the Borlaug Global Rust Initiative (BGRI). BGRI is an international collaborative community that aims to reduce the world's vulnerability to all three rust diseases of wheat (leaf, stem, and stripe) by facilitating an international system to contain their threat. Because of the initiative taken by BGRI as well as all collaborating wheat breeding programs in several countries, breeders and pathologists worldwide are now better prepared to face the disease. However, the competition between rust and the host is a constant arms' race, and as such, continual discovery and deployment of resistance genes is needed.

Gene Discovery & Resistance Breeding

Records indicate that wheat varieties resistant to the rusts were in cultivation as early as 1841 (Chester et al. 1951); however, systematic breeding for resistance and gene deployment did not become common until the early 1900s (Roelfs 1985a). The credit for identification of the earliest source of resistance to wheat stem rust goes to Herbert K. Hayes (1925). He crossed the hard red spring line 'Marquis' with Iumillo durum, and progenies of this cross protected wheat planted in the fields of the northern United States from leaf and stem rust. 'Thatcher', one of the cultivars developed from this particular cross, was a popular cultivar and has served as the source for many hard red spring wheat

lines (Roelfs 1985a). The first documented introgression and deployment of a stem rust resistance gene is that of *Sr2* (McFadden 1930). In the 1920s, Edgar S. McFadden made selections in the field from the cross he made between emmer wheat (*T. turgidum* var. *dicoccum* cv. Yaroslav) and a hexaploid wheat 'Marquis'. He labeled his selections as varieties 'Hope' and 'H44-24' (also known as 'H-44'), and these two sib lines are still being used worldwide to transfer *Sr2* resistance to modern germplasm. The gene *Sr2* has been highly durable and serves as a landmark gene by continuing to provide partial, adult plant resistance to all known races of stem rust.

The work of gene discovery and deployment that began in the 1920s has been religiously continued, with 58 stem rust resistance genes being discovered and documented to date (McIntosh et al. 1995; McIntosh et al. 2003; McIntosh et al. 2011). Most of the genes catalogued in these documents have been contributed by US or Canadian scientists. This is not surprising, as the high frequency of stem rust epidemics observed in North America called for the identification of novel sources of resistance that could protect the wheat crop from newly emerged rust pathogen races. Some of the genes discovered in North America made it to the field quicker, and helped in reducing the damage done by the frequent stem rust epidemics observed in these areas. However, many of the deployed genes had to be replaced within a few years as the selection pressure on pathogens led to evolution of new, virulent rust pathogen races that overcame the deployed genes. This cyclical process of gene discovery, deployment, and defeat by a more evolved pathogen, known famously as the 'boom and bust cycle' (Browning 1979), can be reviewed in a few case studies in North America. For example, the gene *Sr9d* was

used in bread wheat and durum wheat in North America, and was resistant to race 56 during the 1934-1937 epidemics. However, race 15B, which arose to cause the 1950s epidemics overcame *Sr9d* (McIntosh et al. 1995). The genes *Sr12* and *Sr14* also bear a similar story during the rust epidemics in North America. In fact, the boom and bust cycle continues today with the emergence of the Ug99 race group.

Types of Resistance and Disease Control

The preferred approach to combating rust pathogen races has been via the deployment of resistant cultivars. The resistance deployed against stem rust can be grouped into two broad categories: all-stage (or seedling) resistance and adult plant resistance (APR). All-stage resistance is often associated with its hypersensitive reaction upon fungal attack, and confers a high level of resistance throughout all stages of plant growth (Roelfs et al. 1992). Resistance of this type is relatively easy to phenotype in field and greenhouse settings because of the large effect on disease reduction, and have been well catalogued (McIntosh et al. 1995). While all-stage resistance can be bred into desired germplasm with higher degrees of success than APR, the major drawback of using this resistance is the possibility of the pathogen evolving to overcome the deployed resistance. As an alternative, APR, which is generally expressed during the adult growth stages of the plant, has been proposed for durable resistance. APR is sometimes also explained as race-nonspecific resistance or partial resistance, though anomalies to this definition have also been observed in the field. In the wheat rust pathosystems, partial resistance is often called slow rusting resistance, defined by Caldwell (1968) as the type of resistance that would either 1) exclude the fungus, 2) limit pustule size without

expressing hypersensitivity, or 3) reduce the overall rate of growth and development of the pathogen fungus. Parlevliet (1976) characterized slow leaf rusting in barley, and explained the observed resistance as a form of incomplete resistance exhibiting a reduced rate of disease development despite the plants developing a susceptible infection type. APR genes are considered more durable than all-stage resistance, with the stem rust resistance gene *Sr2* being durable against wheat stem rust for almost nine decades. However, phenotyping and breeding of APR genes that confer partial resistance are difficult, and provide low, generally inadequate, levels of resistance when deployed as single genes. Therefore, deployment of lines with both types of resistance genes pyramided is a more common and desired approach (Singh 2012; Evanega et al. 2014).

The disease can also be effectively controlled by the use of fungicides (Cook et al. 1999; Wanyera et al. 2009; Zhensheng et al. 2010). Foliar fungicides have been documented to reduce infection severity on plants, increase grain density as well as grain yield (Mayfield 1985; Wanyera et al. 2009). However, they may come with severe environmental burdens (Steffenson et al. 2007). High costs associated with fungicides may keep them out of reach of smallholder farmers in developing nations. Effective fungicide application largely relies on weather conditions. For example, if there is a rainfall not too long after fungicide application, another round of application might be necessary. Another concern is the development of fungicide tolerance in the pathogen, which render the chemical ineffective for rust control (Gill et al. 1985). Considering these drawbacks, discovery and mapping of genes followed by their introgression into elite

germplasm and careful deployment is the ideal strategy for sustainable and durable protection against stem rust.

Stem Rust Resistance Gene Mapping

While the discovery of resistance genes began in the 1920s, assigning the genes to chromosomes did not commence until the mid 1950s. Even so, this work was limited only to identification of genes in monosomic lines in a Chinese Spring background developed by Ernest Sears (Sears 1954). Seedlings of monosomic lines were inoculated with stem rust races, and genes were mapped (a more appropriate term may be chromosome tagging) to chromosomes through aneuploid analysis. *Sr6*, a dominant gene, and *Sr8*, a recessive gene, were among the first stem rust resistance genes assigned to a chromosome (Sears et al. 1957; Sawhney et al. 1981). The seedlings of monosomic Chinese Spring lines were inoculated with stem rust races 15C and 122 to determine the locations of *Sr6* and *Sr8*, respectively. Analysis of segregants revealed the location of *Sr6* to be on chromosome 2D, and that of *Sr8* on 6A. The development of monosomic lines enabled chromosome tagging of several other genes, namely *Sr5*, *Sr7*, *Sr9*, *Sr11*, *Sr15*, and *Sr16*, carried out by Plessers (1954) and Sears et al. (1957). Mapping stem rust resistance genes solely using aneuploid stocks continued until the early-1990s, where *Sr33* (Jones et al. 1991) and *Sr26* (Bariana and McIntosh 1993) were mapped using recombinant substitution lines obtained from crossing a double-ditelosomic line with a normal chromosome (1D), and a monosomic series of lines, respectively. A point to note here is that tagging chromosomes (using aneuploids) with genes that were discovered before DNA markers, and marker-based genetic mapping of the same genes have not

been a coherent process. In other words, genes discovered before the 1980s were named in the order they were discovered, but their genetic mapping with molecular markers has not always been carried out in the same order of gene discovery. In fact, many of the discovered genes remain unmapped to this day despite their significant contribution to resistance breeding and study of genetic resistance. This is exemplified by the use of several unmapped yet important stem rust genes in the first four sets of differentials used in the North American nomenclature system for race identification of stem rust isolates (Roelfs and Martens 1988).

Once the era of molecular markers began with the use of restriction fragment length polymorphism (RFLP) technique for genetic profiling in the late 1980s, accurate genetic mapping of causative loci was possible. Because of the availability of aneuploid lines for each chromosome in hexaploid wheat (Sears 1954), construction of genetic maps of each chromosome to isolate regions of interest was the logical approach. This approach was first implemented by Chao et al. (1989) who reported the construction of RFLP-based linkage map of all three subgenomes of wheat chromosome 7. A series of studies that followed assigned RFLPs to several wheat chromosomal arms (Anderson et al. 1992; Werner et al. 1992; Devey and Hart 1993; Deynze et al. 1995). Information from these studies would be utilized by studies in subsequent years to accurately map stem rust resistance genes, among other genes of interest, to wheat chromosomes. One of the first studies to do so was carried out by Paull et al. (1994) who identified RFLP markers associated with the stem rust resistance gene *Sr22* on chromosome 7A. The gene *Sr22* is derived from einkorn wheat (*Triticum monococcum*), and is one of the very few

genes that are effective to races in the Ug99 lineage (Kerber and Dyck 1973; Singh et al. 2011). In their study, Paull et al. crossed einkorn wheat *T. boeoticum* with the bread wheat variety 'Schomburgk' to generate F_2 lines. Using RFLP markers, the researchers were able to identify lines with low levels of recombination between chromosome 7A of cultivated and chromosome $7A^m$ of einkorn wheat. The study also investigated the amount of linkage drag in the progeny, and demonstrated reduced alien introgression in progeny. Linkage drag leads to fitness reduction when undesired genes get introduced along with beneficial genes during gene introgression from wild species.

Nelson et al. (1995) identified the locus *Xbcd1095* on the long arm of chromosome 2B that provided major resistance against stem rust of wheat in field conditions. This study hypothesized that the locus could be linked to *Sr16*, a resistance gene possibly contributed by the line 'Thatcher' in the population under study. Shortly after, *Sr2* was mapped to chromosome 3B by Bariana et al. (1998) using short tag sequence (STS) markers and also by Johnston et al. (1998) using RFLP markers in an RIL population of 141 lines. Because of the high importance of the gene *Sr2*, multiple attempts have been made to discover tightly linked markers (Spielmeyer et al. 2003; Hayden et al. 2004; Mago et al. 2011a). Spielmeyer et al. (2003) reported that *Xgwm533* was tightly linked with *Sr2* (at an approximate distance of 2cM) and produced a 120 bp product in PCR. Later, Mago et al. (2011a) reported an improved marker *csSr2* which had an accuracy of 95%, compared to 84% of the marker *Xgwm533*.

Mapping of several other resistance genes that are effective against the Ug99 race group have also been accomplished, in addition to mapping of *Sr2* and *Sr22*. *Sr13* is

resistant against races TTKSK, TTKST, and TTTSK, and was mapped in tetraploid durum wheat (*Triticum turgidum* ssp. *durum* L.) (Simons et al. 2011). *Sr47*, a gene transferred from *Aegilops speltoides* to durum wheat, confers high level of resistance to Ug99 stem rust and was mapped by Klindworth et al. (2012). However, markers for both *Sr13* and *Sr47* are not diagnostic in all genetic backgrounds. Diagnostic markers for the Ug99-effective genes *Sr24*, *Sr25*, and *Sr26* have been developed (Mago et al. 2005; Liu et al. 2010). Mapping and marker development for the genes *Sr28* (Rouse et al. 2012), *Sr39* (Mago et al. 2009), and *Sr45* (Periyannan et al. 2014) have also carried out, but the markers liked to these genes are not reported to be diagnostic.

To date, only two stem rust resistance genes that are also effective against the Ug99 race group have been cloned: *Sr33*, derived from *Aegilops tauschii*, was cloned by Periyannan et al. (2013); and *Sr35*, derived from *Triticum monococcum*, was cloned by Saintenac et al. (2013b). Despite one of the goals of QTL mapping being map based cloning of the gene of interest, it is not an easy task due to environmental conditions confounding QTL detection, and lack of completely linked markers detected during QTL mapping (Remington et al. 2001; Salvi and Tuberosa 2005). In addition, as one of the culturally most important crops in the world, wheat faces political and behavioral challenges in the areas of cloning and gene transformation. Therefore, candidate gene discovery and recurrent selection using diagnostic molecular markers are still the best means of obtaining superior varieties with enhanced trait values.

In addition to the mapped genes, several uncharacterized sources of resistance in common wheat and its diploid progenitors and wild relatives were also documented as

early as 1970 (Roelfs and McVey 1979). These resistant sources have not been assigned a gene number because genetic information such as chromosome location and/or tests of allelism with previously designated genes are required to confirm the identity of the gene. This problem exists mainly because gene discovery and mapping have not always been coherent, as mentioned earlier. One method to determine the identity of a gene is by mapping the segregating resistance, commonly known as gene or quantitative trait loci (QTL) mapping. Comparing the marker loci linked with the resistance to that of a mapped gene, predictions about gene identity can be made. The task of mapping unknown as well as novel sources of resistance did not begin until after chromosome maps were published in the 1990s. Several gene mapping studies have been conducted to identify QTL associated with stem rust resistance in wheat, including those that confer resistance to the Ug99 race group. The majority of these studies have aimed at mapping both all-stage resistance and APR that provide high level of resistance against the disease. Using different mapping population types, QTL mapping is usually carried out by testing the populations in different environments to get exposure of the wide diversity of the pathogen population. These mapping studies characterize the resistance, map the gene location, and identify markers that can be used in marker assisted selection for resistance breeding. The markers identified in a QTL mapping study are not always tightly linked with the gene of interest, and therefore require fine mapping. However, it must be noted that the availability and usability of these markers have been, and will continue to be a vital resources in resistance breeding against stem rust of wheat worldwide. Development

of tightly linked markers to these QTLs as well as characterization of the discovered QTL will play important roles in fighting the spread of rust in the coming decades.

Gene Mapping Tools & Strategies

Of different strategies implemented in mapping, early-generation (F_2, $F_{2:3}$) mapping (examples: Hiebert et al. 2011; Liu et al. 2011), mapping in recombinant inbred line (RIL) populations (examples: Bansal et al. 2008; Singh et al. 2013d), and in association mapping panels (examples: Yu et al. 2012; Letta et al. 2014) populate the literature. With regard to stem rust QTL mapping, biparental RIL populations seem to be preferred, because of 1) more power, and 2) the possibility of detecting rare alleles. The drawback of this approach, however, is that it takes time to develop the population, and unless the population size is adequately large, mapping resolution is typically low. A relatively new approach for mapping genes and QTL in several populations connected by a common parent has been developed in maize (Yu et al. 2008; McMullen et al. 2009). Coined as the nested mapping association (NAM) approach, this method scans for putative QTL in all biparental RIL populations with the genome of the common parent serving as a uniform background. The NAM approach combines the mapping power of a RIL population and higher mapping resolution of an association mapping panel, and therefore is able to discover both major and minor-effect QTL in a large population. Coupled with high throughput genotyping approaches, the NAM design has the potential to dissect marker-trait associations in several populations across several environments.

The advent of better marker technologies such as randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and

simple sequence repeat (SSR) markers in the late 1990s paved the way for efficient mapping of several *Sr* genes. As a result, several stem rust resistance genes have been linked to diagnostic markers since then such as *Sr39* (Gold et al. 1999), *Sr24* and *Sr26* (Mago et al. 2005), *Sr9a* (Tsilo et al. 2007), and *Sr38* (Seah et al. 2000; Sharp et al. 2001). Diagnostic markers for several other important stem rust genes have been published (MAS Wheat 2014, http://maswheat.ucdavis.edu/protocols/), aiding to marker assisted gene pyramiding for resistance breeding. As the marker technology has evolved, SNP-based mapping approaches will likely dominate the former technologies in the future. Because of its improved cost/data point value, mapping genes using SNPs obtained from restriction-digestion based sequencing procedures have been demonstrated well in barley (Liu et al. 2014), and wheat (Saintenac et al. 2013a). Also known as genotyping by sequencing (GBS) or restriction site associated DNA (RAD) sequencing, this approach of obtaining population specific SNPs at a cheaper, faster rate can be used to enrich QTL regions linked to traits of interest (Chapters 2, 3 of this thesis). The flexibility offered by the GBS approach expands to imputation of missing haplotypes in related individuals due to low coverage in sequencing. Recent studies in several crop species including wheat have shown that imputing markers to construct missing haplotypes can be done with high accuracy (Rutkoski et al. 2013; Fu 2014). Thus, in mapping populations, where the individuals are either linked in families (F_2, RIL populations) or via population structure (association mapping panels), the utilization of GBS SNPs appears promising.

Chapter 2

Mapping Putatively Novel QTL Conferring Adult Plant Resistance to Ug99 in the Biparental Population RB07/MN06113-8

The emergence and spread of the Ug99 race group of the stem rust pathogen in the past decade has exposed the vulnerability of wheat to this disease. Discovery of novel and effective sources of resistance is vital to reduce losses. The experimental breeding line MN06113-8 developed by the University of Minnesota wheat breeding program exhibits a high level of adult plant resistance (APR) to Ug99 race group. The University of Minnesota wheat cultivar RB07 exhibits a medium level of APR to Ug99 race group. Both of these lines are susceptible at seedling stages to the Ug99 races. To dissect the genetic mechanism of resistance present in these lines, MN06113-8 was crossed to RB07 to generate 141 F_6 recombinant inbred lines (RILs). The RIL population was evaluated for APR to Ug99 in Kenya and Ethiopia over three seasons; and for APR to North American stem rust races in St. Paul, MN in one season. Composite interval mapping detected six quantitative trait loci (QTL) involved in APR to African stem rust pathogen races and three QTL involved in APR to North American stem rust pathogen races. One QTL located on the short arm of chromosome 2B was observed in all environments and provided resistance to all stem rust pathogen races, and could be a novel source of APR to stem rust pathogen races, including the Ug99 race group. Development of diagnostic

markers linked to this gene will facilitate will assist in marker assisted selection of resistant lines to generate varieties with elevated levels of APR to Ug99.

INTRODUCTION

One of the primary objectives of resistance breeding is the effective employment of resistance sources that promise durability. Constantly evolving pathogen populations challenge the effectiveness and durability of deployed resistance genes. The disease stem rust of wheat, caused by the fungal pathogen *Puccinia graminis* Pers. f. sp. *tritici*, is one such example where the pathogen is persistently evolving to overcome host resistance. One of the oldest plant diseases known to mankind (Kislev 1982), stem rust of wheat is highly destructive, and bears potential to completely destroy small scale farm plots to millions of hectares of susceptible varieties (Roelfs 1985b).

The severe threat that stem rust has historically posed to global wheat production has been magnified in recent decades by the evolution of a highly virulent race, TTKSK. The isolate of this pathogen race was first observed in Uganda in 1998 and was named Ug99 for the country of origin and the year it was evaluated for its virulence phenotyping (Pretorius et al. 2000). Ug99 was found to overcome lines with *Sr31*, a widely deployed gene that provided resistance against stem rust of wheat at the time. Within a few years of its discovery, this race spread toward North Africa, West Africa, and the Middle East, and has potential to travel to West-South Asia as global wind patterns may transport the fungal spores over long distance (Singh et al. 2008a; Hodson et al. 2011). Ug99, later named as TTKSK after characterization using North American stem rust differential sets (Jin et al. 2008), along with its six other related races are virulent to 85–95% of breeding materials worldwide (Wanyera et al. 2006; Singh et al. 2011). The evolution of race

TTKSK and related races meant that the pathogen is capable of defeating multiple important rust genes that had been protecting wheat from stem rust for a long time, as evident by the breakdown of genes such as *Sr31*, *Sr24*, and *Sr36* (Jin et al. 2008; Singh et al. 2008a; Jin et al. 2009a). Furthermore, many of the effective rust genes in bread wheat and its wild relatives were either ineffective against these races or unusable in the field with regards to grain quality and/or field performance (Singh et al. 2011). While this unexpected evolution of the Ug99 race group exposed the high vulnerability of wheat crops grown worldwide to stem rust, it also drew attention to the fact that the efforts in global rust monitoring and resistance breeding were not adequate to contain stem rust, making a search for durable resistance even more urgent.

Durability of resistance in wheat is considered to be contributed by slow-rusting genes that protect the host based on the adult plant resistance (APR) mechanism. APR is assumed to act in ways that are nonspecific to rust races, and is generally distinguished by low infection frequency, reduced size of urediniospores, and overall diminished urediniospore production (Stuthman et al. 2007). In the case of wheat stem rust, APR has been further described as detected in mature plants, and mostly associated with absence of hypersensitive response (Hare and McIntosh 1979), and is quantitatively inherited (Knott 1982). APR genes are known to perform marginally when deployed individually. Singh et al. (2005) discuss that pyramiding a few APR genes can confer near immunity against diseases but may be difficult to accomplish because of the difficulties in phenotyping and identifying diagnostic markers associated with the resistance loci (Singh 2012). Combining multiple race-specific or major genes with or without APR genes has

also been proposed and utilized to obtain durable resistance against the disease (Ayliffe et al. 2008; Mago et al. 2011b; Evanega et al. 2014). Only a few stem rust APR genes have been discovered in wheat: *Sr2* (Knott 1968), *Sr55* (*Lr67/Yr46/Pm46*) (Herrera-Foessel et al. 2014), *Sr56* (Bansal et al. 2014), *Sr57* (*Lr34/Yr18/Pm38*) (Lagudah et al. 2006), and *Sr58* (*Lr46/Yr29/Pm39*) (Singh et al. 2013c). Continual discovery of new sources that confer APR and promise durable resistance is vital for crop protection.

The threat from Ug99 and its lineal races has been responded to, at least partly, by wheat research teams throughout the world by screening for resistance, and using the identified genes in their breeding programs. Many of these genes effective to Ug99 have been identified in non-bread wheat species. Examples include *Sr32* which was identified in *Aegilops speltoides* (McIntosh et al. 1995), *Sr37* in *Triticum timopheevi* (McIntosh and Gyarfas 1971), *Sr39* in *Aegilops speltoides* (Kerber and Dyck 1990), *Sr40* in *Triticum araritum* (Dyck 1992), *Sr44* in *Thinopyrum intermedium*, and *Sr53* in *Aegilops geniculata* (Liu et al. 2011). While wild relatives of bread wheat are excellent sources of previously undiscovered resistance genes, the issue of linkage drag, which occurs from introgression of genes derived from non-elite germplasm to elite breeding material, is a challenge. Breeders are usually hesitant to utilize such wild sources of resistance in their materials because of the time and effort it takes to select for lines with desired agronomic traits. Hence, discovery of resistance material in existing breeding programs would be a clear advantage, as crossing advanced lines with novel resistance to other elite lines would introduce the resistance while also preserving desired agronomic qualities.

The wheat breeding program at the University of Minnesota develops hard-red spring wheat varieties with superior agronomic performance and disease resistance. To safeguard the released varieties against a potential Ug99 epidemic, the program has routinely screened dozens of advanced experimental lines each year for resistance since 2005 in a stem rust nursery coordinated by USDA-ARS and the Kenya Agricultural Research Institute (KARI) in Njoro, Kenya. The advanced experimental line MN06113-8, despite being susceptible to Ug99 races at the seedling stage, was found to exhibit APR to Ug99 races the in Njoro stem rust nursery in Kenya. RB07 is moderately resistant – moderately susceptible (MRMS) to Ug99 races in the field, susceptible to Ug99 at the seedling stage. In order to understand the genetic mechanism of stem rust resistance in these lines, they were crossed (RB07/MN06113-8) to generate a biparental RIL population segregating for APR to Ug99 races. As both parent lines are highly inbred advanced lines, they possess several desirable agronomic traits. For the same reason, the resistance genes discovered in this RIL population can be used to develop varieties with elevated resistance to Ug99 races.

In this study, we map the resistance segregating in the RB07/MN06113-8 RIL population. We also estimate the genetic effects of the detected loci and trace their origin. We hypothesize that some of the QTL regions detected in this study are putatively novel.

MATERIALS AND METHODS

Plant Materials

A mapping population of 141 recombinant inbred lines (RILs) was developed via the single seed descent method by crossing 'MN06113-8' to 'RB07', both of hard red spring wheat growth habit. The line RB07 has the pedigree Norlander (PI 591623)/HJ98 and was developed by the University of Minnesota Agricultural Experiment Station. RB07 was released as a cultivar in 2007 on the basis of its high and consistent grain yield, earliness, resistance to wheat leaf rust, moderate resistance to Fusarium head blight, and good grain end-use quality (Anderson et al. 2009). The F_6-derived line MN06113-8 has the pedigree MN97695-Lr52/HJ98-Fhb1 and is a breeding line at the University of Minnesota Wheat Breeding Program that advanced to second year yield trials before being discontinued for consideration as a new cultivar candidate due to low test weight and soft texture endosperm. Both parents are seedling susceptible to TTKSK, and TTKST races; and MN06113-8 is susceptible to race TTTSK of the Ug99 lineage (no data on RB07), and confer medium to high levels of resistance in the field in Kenya (Table 1). Seedling tests for reaction to the Ug99 races was carried out following the protocol outlined by Rouse and Jin (2011) at the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) Cereal Disease Laboratory. Infection types (ITs) were recorded on a 0-4 scale according to Stakman et al. (1962).

Field Stem Rust Evaluation

The $F_{6:7}$ and $F_{6:8}$ populations, along with the parents, were evaluated for their field response to African stem rust races at two locations in East-Africa over three seasons: at Njoro, Kenya during the main season from May - October of 2012 and the off-season from November 2012 to April 2013 (referred as Ken12 and Ken13 hereafter, respectively), and at Debre Zeit, Ethiopia during the off-season from March - June of 2013 (hereafter referred as Eth13). The population was also evaluated in St. Paul, MN, USA during May – August 2013 (hereafter referred as StP13) for response to N. American stem rust races.

In the Njoro nursery, lines were planted in an augmented design with 1 check, 'Red Bobs'. Each line was sown in double 70 cm long rows, 20 cm apart. On each side of the plot, and in the middle of the plots, a twin-row of susceptible spreader wheat cultivar 'Cacuke' was sown. The field was also surrounded by a border of several spreader rows comprised of susceptible wheat varieties which were artificially inoculated using a bulk inoculum of *Pgt* urediniospores collected at the Njoro field site; however wheat stem rust differential lines with known stem rust resistance genes indicated that the predominant, if not only, race present in the nursery since 2008 was race TTKST (avirulence/virulence formula on the wheat stem rust differential panel: *Sr36*, *SrTmp*/*Sr5*, *Sr6*, *Sr7b*, *Sr8a*, *Sr9a*, *Sr9b*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr10*, *Sr11*, *Sr17*, *Sr21*, *Sr24*, *Sr30*, *Sr31*, *Sr38*, *SrMcN*) (Njau et al. 2010).

In the Debre Zeit nursery, lines were planted in 1 m long twin rows that were flanked by spreader rows comprised of a mixture of susceptible wheat varieties

'PBW343', 'Morocco', and 'Local Red'. Spreader rows were artificially inoculated with bulk inoculum of *Pgt* urediniospores to initiate the disease. At both locations, fresh urediniospores collected from susceptible varieties in the field were suspended in water followed by syringe-inoculation into susceptible spreader plants prior to booting stage (growth stages Z35 – Z37; Zadoks et al. 1974) at an approximate distance of 1 m.

In the St. Paul nursery, lines were planted in hill-plots with 20 cm distance between the hills. The population was planted in an augmented design with 4 check varieties 'Oklee' (Anderson et al. 2005), 'Thatcher' (Hayes et al. 1936), 'Tom' (Anderson et al. 2012), and 'Verde' (Busch et al. 1996) planted after every 30 entries. A mixture of susceptible lines 'Morocco' and 'LMPG-6' were planted perpendicularly to surround the lines on all sides. To initiate disease, spreader rows were syringe-injected with a mixture of North American stem rust races MCCFC (isolate 59KS19), QFCSC (isolate 03ND76C), QTHJC (isolate 75ND717C), RCRSC (isolate 77ND82A), RKQQC (isolate 99KS76A), and TPMKC (isolate 74MN1409) at the jointing stage. The spreader rows were sprayed with a bulked mixture of *Pgt* races suspended in a light mineral oil suspension using an Ulva+ sprayer (Micron Sprayers Ltd., Bromyard, UK) after heading.

Phenotyping and Data Analysis

Field reaction of the RILs to stem rust were recorded as disease severity on the 0-to-100 modified Cobb scale (Peterson et al. 1948), and infection response, based on the size of pustules and amount of chlorosis and necrosis visible on the stem (Roelfs et al. 1992). Disease phenotyping of the population segregating for resistance was carried out after the susceptible check varieties in each trial had attained maximum severity.

Following Stubbs et al (1986), the severity response value was multiplied with the infection response to obtain coefficient of infection values. Using a mixed model (*lme4* package in R 3.0.3, R Development Core Team, 2013), the growth stages of the lines were fitted to the coefficient of infection values to obtain fixed-effect estimates for each line, which were used for QTL mapping. Growth stages of each line in Kenya were determined mainly by assessing grain development stages such as watery, milky, soft dough, and hard dough; and also for stages of booting and flowering, as explained by Zadoks et al. (1974).

Analysis of variance (ANOVA) was performed using function PROC GLM in SAS 9.1 (SAS Institute Inc., Cary, NC, USA) with genotype as a fixed effect, and environments and a combination of locations and years as random effects. Replicated checks were used to calculate the pooled error mean square value. Pearson correlation coefficients among the trials were calculated using the function PROC CORR in SAS 9.1.

Molecular Marker Assay

Genomic DNA was extracted from ground seeds of the parents and $F_{6:7}$ RILs using a modified cetyltrimethylammonium bromide (CTAB) protocol (Kidwell and Osborn 1992). The extracted DNA was quantified using an ND 1000 Spectrophotometer (NanoDrop Technologies, Delaware, USA). The population was genotyped using SNP markers obtained from two approaches: 1) the 9000 Infinium iSelect SNP assay (9K) (Cavanagh et al. 2013) and 2) genotyping by sequencing (GBS) (Elshire et al. 2011).

For genotyping using the Infinium iSelect assay, DNA suspended in ddH_2O at approximately 80 ng/µl was submitted to the USDA-ARS Small Grain Genotyping Center, Fargo, ND, USA. The data generated was manually called using Illumina's GenomeStudio 2011.1 (Illumina Inc., Hayward, CA). Briefly, monomorphic markers, markers with the same calls for the entire population, markers with more than 10% missing data, and markers that deviated from a 1:1 segregation ratio were discarded. Markers with 5% or less heterozygous calls were retained to avoid false purging of heterozygous loci. This resulted in 1,050 high quality markers that were retained for linkage mapping.

To increase mapping resolution, and partly to investigate the feasibility of genome mapping using markers obtained from next-gen sequencing, the population was also genotyped using the GBS method (Elshire et al. 2011). In the GBS approach, a double-digested library was created using the restriction enzymes *Pst*I and *Msp*I on 200 ng of DNA/sample, following Poland et al. (2012a) with modifications. Each library was 76-plexed, with the parents repeated six times each, and the libraries were sequenced in two lanes of Illumina HiSeq 2000, generating 100 bp paired-end sequences. The sequences were processed using the UNEAK pipeline (Singh et al. 2013b) using the parameters -c 10 –e 0.025 to obtain *de novo* SNPs. Reads containing SNPs were used as query sequences and blastn-searched against the wheat chromosome survey sequences (CSS) to assign SNPs to unique chromosomes. The wheat CSS sequences are obtained by assembling reads obtained from sequencing flow-sorted wheat chromosomes from the 'Chinese Spring' variety (International Wheat Genome Sequencing Consortium,

http://wheaturgi.versailles.inra.fr/Seq-Repository/). To ensure that correct SNPs were obtained, only the full length alignment of a query sequence with the survey sequences allowing either one base mismatch or one gap was permitted. To circumvent retaining redundant SNPs on paralog sequences and duplicated regions among the A, B, and D subgenomes, SNPs thus obtained were filtered to remove those that mapped more than once to multiple chromosomes. SNPs that were monomorphic, had no allele calls for >10 individuals (>7% missing data), and were heterozygous in >10 individuals (7% heterozygosity) were also discarded. This process resulted in 932 high quality SNP loci that were retained for linkage mapping.

Linkage Map Construction and QTL Mapping

SNPs obtained from both genotyping approaches (9K, *de novo*) were combined to assign markers to linkage groups. Linkage groups were constructed using Mapdisto V1.7.7.0.1 (Lorieux 2012) using a minimum logarithm of odds (LOD) value of 3.0. Genetic distances between the markers were calculated based on the Kosambi mapping function (Kosambi 1943). The program Windows QTL Cartographer 2.5_011, which implements composite interval mapping (CIM) to identify QTL, was used to analyze marker-trait associations (Wang et al. 2012b). The LOD threshold for declaring a significant QTL was calculated by 1,000 permutations at α=0.05. A walk speed of 1 cM was used for QTL detection. QTL effects were estimated as the proportion of phenotypic variance (R^2) explained by the QTL. If multiple QTL were detected in an environment, digenic additive x additive epistatic interactions were tested among the detected QTL using the multiple interval mapping (MIM) algorithm available in the same program.

Using multiple marker intervals simultaneously, the MIM procedure fits multiple putative QTL directly in the QTL mapping model, and estimates several genetic architecture parameters including the effects of and interactions among significant QTL. Epistatic interaction among all SNPs, irrespective of their association with the detected QTL, was also carried out using the MIM algorithm. A QTL x QTL relationship was declared significant if LOD threshold was ≥ 1.0.

RESULTS

Disease Evaluation

The disease pressure observed in all environments was adequate for good discrimination among stem rust phenotypes of the RILs, and the mapping of loci associated with quantitative resistance. The disease severity distributions skewed towards the lower percent severity responses overall (Figure 1). The highest disease pressure was observed in Ethiopia with lines showing up to 70% severity. Disease scores recorded for the RIL population along with the parents in all four environments are presented in Table 1. There was no significant difference between MN06113-8 and RB07 for their average stem rust responses across all environments (t-test P-value of 0.25 at α=0.05). This is not completely unexpected, as both parent lines exhibit similar levels of APR in the field (Table 1, Figure 1). Disease severity distributions for the RILs across all environments were continuous, suggestive of quantitative and polygenic resistance. Disease reactions of RILs between Kenya and Ethiopia nurseries were strongly correlated whereas correlations with the St. Paul reactions were lower, yet still significant (Table 2). The lower correlation coefficient between these environments suggests the presence of high genotype by environment (GxE) interaction, corroborated by the significant F-test value of 3.1 (significant at p-value <0.001). Some of the factors that could have contributed toward GxE could be the differences between *Pgt* races, inoculum load, and differences in environmental conditions.

Construction of Linkage Maps of the RIL Population

SNPs obtained from both genotyping methods – the 9K SNP chip and *de novo* GBS SNPs – were combined to develop linkage maps to represent wheat chromosomes. Of the 1,982 markers used for linkage mapping, 1,899 (972 9K, 927 GBS) were assigned to 29 linkage groups. All 21 wheat chromosomes were represented by the linkage groups, with 8 chromosomes represented by two linkage groups. These markers covered 1,950 cM of the genome, with an average interval of 1.03 cM. Most SNPs were assigned to the A and B genomes, with 855 and 909 SNPs, respectively, while the D genome harbored 135 SNPs (Figure 2).

Quantitative Mapping of Resistance to Stem Rust

The CIM method of QTL mapping detected six significant QTL involved in resistance to African stem and four QTL involved in APR to North American *Pgt* races (Table 3) with one QTL significant in all four environments. The detected QTL ranged in their LOD scores from 2.6 to 16, with four QTL contributed by MN06113-8 and five contributed by RB07. The QTL *QSr.umn-2B.2* was detected in all environments (Figure 3), and explained 31.4%, 27.1%, 46.9%, and 8.7% of observed phenotypic variation (R^2) in the Ken12, Ken13, Eth13, and StP13 environments, respectively. This QTL, derived from MN06113-8, was also the largest QTL in all environments except in StP13 where the QTL *QSr.umn-4B.2* was the largest in terms of both R^2 and allelic effect (Table 3). Other QTL explained 5% to 13% of the resistance variation, but were significant only in one of the four environments. Multiple QTL on a single chromosome were detected in

two environments: Ken12 with two QTL on chromosome 2B; and StP13 with two QTL on chromosome 4B.

Test for additive x additive epistasis among the detected QTL revealed no significant interactions. Test for epistasis among all loci (significant or not) also detected no significant interactions. All two-QTL models were generated to estimate the average reduction in severity (Table 4). The results indicate that combinations of at least two QTL in a gene-pyramiding scheme can reduce disease severity values from 9.2% to 52.2% in the environments where the QTL were detected.

DISCUSSION

In this study, we report several genomic regions associated with APR to African *Pgt* races in the Ug99 lineage as well as North American *Pgt* races in lines derived from the University of Minnesota wheat breeding program. While disease pressure was higher in Ethiopia than in Kenya and St. Paul, it was adequate in all environments, allowing for high-quality phenotypic assessment of the mapping population. In the Ethiopia trials, the distributions skewed towards higher percent severity, most likely due to high disease pressure observed during the season. The lowest disease severity was observed in St. Paul, as both parents used to generate the population were selected for resistance to the same *Pgt* races. Analysis of the segregating population in four environments led to identification of nine QTL on chromosomes 1A, 2B, 2D, 4A, 4B, 6D, and 7A, explaining from 4.5% to 46.9% of resistance variation in the field.

The QTL *QSr.umn-2B.2* was significant in all four environments, and was contributed by the resistant parent MN06113-8 (Table 3). A small-effect QTL, *QSr.umn-2B.1*, was also detected on chromosome 2B in the Ken12 environment. Based on mapping of significant SNP sequences to the WCSS and also the published 9K map (Cavanagh et al. 2013), the chromosomal locations of *QSr.umn-2B.1* and *QSr.umn-2B.2* were determined to be on the long arm and short arm of 2B, respectively. QTL located on 2B that provide resistance to African stem rust races have been reported previously. Singh et al. (2013a) mapped the QTL *QSr.cim-2BS* on 2BS between the DArT markers *wPt-9230* and *wPt-744022* that explained 3.2 - 6.2 % of the resistance expressed in the CIMMYT population PBW343/Muu.. Bhavani et al. (2011) also detected a QTL on 2B

between the DArT markers *wPt-7829* and *wPt-2266* that provided moderate to low levels of resistance to Ug99 races in the CIMMYT population PBW343/Juchi. Additionally, in their association analysis of CIMMYT spring wheat germplasm, Yu et al. (2011) reported that the markers *wPt-7750*, *wPt-8460*, and *wPt-7200* on chromosome 2B were significantly associated with resistance to races of the Ug99 lineage. None of these reported markers overlap or flank the QTL region on 2B reported in this study; neither do the seedling genes (*Sr28*, *Sr32*, *Sr39*, *Sr40*, *Sr47*, *SrWeb*) that provide resistance to African stem rust races including the Ug99 race group. No known major genes effective to Ug99 are postulated in the population as the parents are susceptible to the races TTKSK, TTKST, and TTTSK (only MN06113-8) with infection types of 3+ to these races (data not shown). Therefore, the QTL *QSr.umn-2B.2* on 2BS could be a novel APR gene with a large-effect on disease reduction. More importantly, this QTL provided field resistance against all *Pgt* races we used to screen the population. Hence, its use in breeding for APR to stem rust globally could be a major advantage in the fight against wheat stem rust. In addition, some wheat APR genes are known to be effective against more than one pathogen, either due to pleiotropy or co-localization with genes resistant to other pathogens (Suenaga et al. 2003; William et al. 2003; Risk et al. 2013). The short arm of chromosome 2B reportedly contains APR QTL conferring resistance to stripe rust (Carter et al. 2009; Prins et al. 2011b) and leaf rust (Tsilo et al. 2014). Therefore, it is worth investigating if the QTL *QSr.umn-2B.2* also provides resistance to other pathogens such as leaf rust, and stripe rust.

The minor effect QTL *QSr.umn-1A* and *QSr.umn-2D* were both detected in the Ken13 environment. Loci significantly associated with resistance to African stem rust races have been reported in CIMMYT's winter wheat breeding germplasm (Yu et al. 2012), durum wheat (*Triticum durum* Desf.) association mapping panel (Pozniak et al. 2008), and a doubled-haploid durum wheat population Sachem/Strongfield (Singh et al. 2013b). The only QTL located on 2D providing APR to the African stem rust races was reported by Bhavani et al. (2011) in the CIMMYT population PBW343/Kiritati, which is located 14 cM from the position of *QSr.umn-2D*.

Another minor effect QTL observed in our study was *QSr.umn-4A*, located on the short arm of chromosome 4A. Previous reports of QTL located on 4A include the mapping study by Bhavani et al. (2011) where a QTL providing resistance to African stem rust in the CIMMYT population PBW343/Juchi was mapped at position 123 cM of 4A. Yu et al. (2011) also reported a QTL on 4A between positions 68.1 cM and 102.7 cM that were significantly linked with APR to African races. The QTL *QSr.umn-4B.1* and *QSr.umn-4B.2* were both detected in the StP13 environment, yet the former was contributed by RB07 and the latter was contributed by MN06113-8. In the PBW343/Kingbird population, Bhavani et al. (2011) detected a QTL on the proximal end of 4B (4 cM) that provided APR to African stem rust. We detected two QTL on chromosome 4B - *QSr.umn-4B.1* (contributed by RB07) and *QSr.umn-4B.2* (contributed by MN06113-8) that provided resistance against stem rust in the StP13 environment. Previous reports of QTL on 4B include a minor effect QTL detected in the Canadian

cultivar ‘Carberry’ (Singh et al. 2013a); and an APR QTL contributed by the Indian cultivar ‘WL711’ in the RIL population HD2009/WL711 (Kaur et al. 2009).

We propose that the QTL *QSr.umn-6D*, detected in the Ken12 environment, is putatively novel as no QTL effective to races of the Ug99 lineage have been reported in this region. This QTL originates from RB07, and is located on the short arm of the chromosome. The All-stage gene *Sr42*, is also located on 6DS, and provides resistance to the African races TTKSK, TTKST, and TTTSK (Ghazvini et al. 2012). Since we found the parents to show susceptible infection types to these three races during seedling and marker screening, it is unlikely that *Sr42* is present in either parent. Another All-stage gene present on 6DS, *Sr5*, is not effective against races in the Ug99 lineage (Singh et al. 2011) and does not provide APR to stem rust when deployed singly, but it may be involved in APR when used with other genes (Nazareno and Roelfs 1981; Knott 2001). We do not have sufficient information to investigate the relationship between *QSr.umn-6D* and *Sr5*.

The QTL *QSr.umn-7A*, also contributed by RB07, mapped on the distal end of chromosome 7AS (0.8 cM), and explained 4.5% of the phenotypic variance. Singh et al. (unpublished; see Yu et al. 2014) report a QTL located between 2.9 cM and 5.6 cM on 7AS in the PBW343/Kenya Nyangumi population, which could be the same as the QTL detected in our study. Further studies including fine mapping of the region is essential to establish the novelty of this resistance locus. Also, given that we implemented SNP markers and other studies have used either DArT or SSR markers, a satisfactory direct comparison with published QTL cannot be made with the existing data. Characterizing

the QTL detected in our study in relation to QTL identified in other studies can be helpful in terms of sustainable gene deployment strategies. However, the use of different marker systems in different studies makes it difficult to do so. Regardless, addition of these QTL to the breeding germplasm can aid significantly in the fight against the disease.

One interesting discovery made in this study is the lack of detection of QTL on chromosomes with genes exhibiting APR to stem rust. No QTL were detected on chromosomes 1B (location of *Sr58*), 3B (location of *Sr2*), 4D (location of *Sr55*), 5B (location of *Sr56*), and 7D (location of Sr57). The lack of QTL detection on these chromosomes suggests that a putatively novel APR gene could be present in the RB07/MN06113-8 population. Our discovery of *QSr.umn-2B.2* in all four environments, as well as other QTL in previously unreported chromosomal locations provide a strong case for the existence of novel QTL in the RB07/MN06113-8 population that confer APR against African stem rust races. Identification of novel APR genes is significant because it would provide breeders new and likely durable tools to develop varieties resistant to highly virulent *Pgt* races, such as the Ug99 race group. The discovery of novel APR genes and their deployment can also alleviate the selection pressure put on the pathogen by newly deployed race-specific resistance genes (Singh et al. 2008a; Evanega et al. 2014), and therefore prolong their durability. Further field screening and tests are required to confirm the novelty of these discovered regions.

The emergence of widely virulent African stem rust races, as discussed above, is often credited to the pathogen's ability to defeat singly deployed resistance genes in areas conducive to the pathogen's growth, development, and evolution. To slow down the

pathogen from developing virulence from single-step mutation for deployed resistance genes, the scheme of resistance breeding by pyramiding multiple genes has been proposed (Knott 1989). As pyramiding of multiple genes in a single line can be resource-intensive given the large number of progeny needed for screening (Bonnett et al. 2005), combining fewer genes may be desirable to obtain immediate resistance to the disease. The combination of only two QTL detected in our study show clear reduction in disease severity (Table 4), and can be implemented to lower the disease pressure. Combining multiple QTL as the long-term breeding goal could be of higher interest, especially in disease hotspots, as pyramiding APR genes with major genes may also increase the durability of major genes (Burdon et al. 2014; Mundt 2014). Further work to fine map the regions to identify diagnostic markers is required to accomplish successful pyramiding of multiple QTL.

CONCLUSION

It is generally considered that widely virulent African stem rust races reaching the breadbaskets of Asia and the Americas is a real possibility (Hodson et al. 2011). Understanding the threat of their possible arrival, several mapping studies in different types of mapping populations have been conducted to discover resistance loci effective against the race group. Here, we report a putatively novel QTL *QSr.umn-2B.2* and other sources of resistance that are effective against African and North American *Pgt* races. The RIL mapping population was developed using advanced breeding lines from the University of Minnesota wheat breeding program. An advantage of mapping resistant QTL in a population obtained from crossing two elite parent lines is that the resistant line can be used as a donor parent in a recurrent breeding scheme without linkage drag. The discovery of these QTL in our elite parent lines should offer value to their utilization for APR to the Ug99 race group in other breeding programs globally.

Table 1: Mean values and range of stem rust severity (%) in the RIL mapping population derived from RB07/MN06113-8 in four field environments at Njoro, Kenya, Debre Zeit, Ethiopia, and St. Paul, MN, USA.

Environments	Parents (Mean ± SD)		Parent Growth Stage[a]		RIL Population	
	MN06113-8	RB07	MN06113-8	RB07	Mean ± SD	Range
Kenya 2012	1.0 ± 0.0	14.4 ± 3.8	SD-HD	SD	23.8 ± 13.8	1.0 – 50.0
Kenya 2013	20.8 ± 8.8	23.3 ± 1.4	M	F	30.5 ± 8.9	12.5 – 45.0
Ethiopia 2013	42.5 ± 2.9	51.3 ± 4.8	-	-	47.2 ± 12.2	15.0 – 70.0
St. Paul 2013	16.7 ± 2.9	16.7 ± 2.9	SD	SD	20.6 ± 7.4	5.0 – 40.0
Mean	20.3 ± 3.6	26.4 ± 3.2			30.5 ± 10.6	8.4 – 51.3

[a] F = flowering; M = milky; SD = soft dough; HD = hard dough; SD-HD = soft dough progressing towards hard dough. Growth stages were assigned according to Zadoks et al. (1974).

Table 2: Pearson correlation coefficients of stem rust severity observed in the RB07/MN06113-8 population in four field environments at Njoro, Kenya, Debre Zeit, Ethiopia, and St. Paul, MN, USA.

	Ken12	Ken13	Eth13	StP13
Ken12	1.00			
Ken13	0.40	1.00		
Eth13	0.55	0.66	1.00	
StP13	0.32	0.29	0.48	1.00

Ken12 = Njoro, Kenya 2012; Ken13 = Njoro, Kenya 2013; Eth13 = Debre Zeit, Ethiopia 2013; StP13 = St. Paul, MN 2013.

All correlations are significant at $p < 0.001$.

Table 3: Quantitative trait loci (QTL) for APR to stem rust detected in the RB07/MN06113-8 population of 141 recombinant inbred lines by composite interval mapping in four environments.

			Flanking Markers[b]					
Environment	QTL[a]	Chr	Left	Right	Pos (cM)	LOD[c]	R^{2d}	Add[e]
Ken12	QSr.umn-2B.1	2BL	TP3836	TP10865	15.3	4.7	7.6	-3.0
	QSr.umn-2B.2	2BS	wsnp_Ex_rep_c71023_69867676	TP38912	96.9	16.0	31.4	6.1
	QSr.umn-6D	6DS	wsnp_BE445201D_Ta_1_1	TP16823	0.01	3.4	5.4	-2.5
	QSr.umn-7A	7AS	TP2402	TP2424	0.8	2.9	4.5	-2.3
Ken13	QSr.umn-1A	1AS	wsnp_Ku_c5756_10191339	TP33660	98.3	2.9	12.5	3.6
	QSr.umn-2B.2	2BS	wsnp_Ex_rep_c71023_69867676	TP38912	99.3	4.6	27.1	5.2
	QSr.umn-2D	2DS	TP8148	TP26641	33.6	3.1	13.1	-3.6
Eth13	QSr.umn-2B.2	2BS	wsnp_Ex_rep_c71023_69867676	TP38912	97.9	13.5	46.9	10.0
StP13	QSr.umn-2B.2	2BS	wsnp_Ex_rep_c71023_69867676	TP38912	104.7	4.1	8.7	2.5
	QSr.umn-4A	4AL	TP10795	wsnp_Ex_c2352_4405961	20.9	2.7	5.6	2.0
	QSr.umn-4B.1	4BS	wsnp_Ex_c6739_11646407	wsnp_Ku_c8075_13785546	21.7	2.8	5.9	-2.2
	QSr.umn-4B.2	4BL	wsnp_Ku_c8075_13785546	TP4428	43.1	5.6	12.3	3.4

[a] QTL were named according to McIntosh et al. (2003)

[b] Markers beginning with 'TP' are *de novo* SNPs discovered from GBS approach.

[c] LOD values are the peak logarithm of odds score for the given QTL

[d] Value indicates the phenotypic variation explained by the QTL

[e] Value indicates the estimated additive effect of the QTL; negative value means that the allele was contributed by RB07

Table 4: Reduced stem rust severity in the RB07/MN06113-8 RIL population by combinations of QTL n pairs in each environment.

		Stem Rust Severity (%)			
Environment	QTL Combinations	No QTL[a]	2 QTL Combination[a]	Difference in Severity	Disease Reduction (%)
Ken12	2B.1 + 2B.2	29.7	16.6	13.1	44.2
	2B.1 + 6D	29.1	17.6	11.5	39.6
	2B.1 + 7A	28.5	15.7	12.9	45.1
	2B.2 + 6D	28.3	13.5	14.8	52.2
	2B.2 + 7A	29.4	14.7	14.7	50.0
	6D + 7A	31.8	17.9	13.9	43.8
Ken13	1A + 2B.2	34.3	24.7	9.6	28.0
	1A + 2D	34.1	21.5	12.6	37.0
	2B.2 + 2D	35.5	24.0	11.5	32.4
StP13	2B.2 + 4A	23.4	17.4	6.0	25.8
	2B.2 + 4B.1	21.2	17.9	3.3	15.6
	2B.2 + 4B.2	23.3	18.1	5.2	22.2
	4A + 4B.1	24.4	17.9	6.5	26.6
	4A + 4B.2	20.6	18.7	1.9	9.2
	4B.1 + 4B.2	25.2	19.1	6.1	24.1
Africa[b]	2B.2	33.3	25.6	7.7	30.0
USA[b]	2B.2	21.6	19.4	2.2	10.4
All Environments[c]	2B.2	29.4	23.9	5.5	18.7

[a] Numbers represent the average disease severity (%) of genotypes upon absence of QTL and combining two QTL

[b] The effect of presence/absence of the uniform QTL *QSr.umn-2B.2* on stem rust severity in Africa and USA environments

[c] The effect of presence/absence of the uniform QTL *QSr.umn-2B.2* on stem rust severity in all environments. Stem rust severity scores for the genotypes were averaged across the four environments

The environment Ethiopia 2013 is not shown because only 1 QTL was detected in this environment.

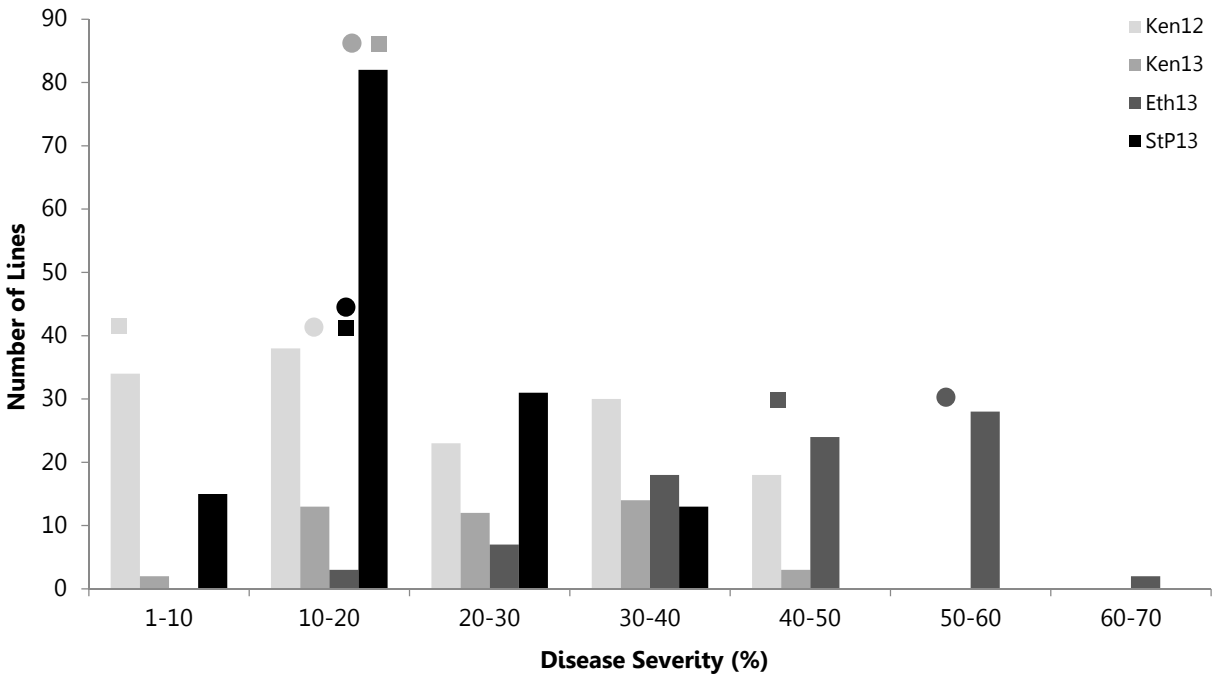


Figure 1: Frequency distribution of stem rust severity (%) for the RB07/MN06113-8 population comprised of 141 recombinant inbred lines evaluated in four environments – two at Njoro, Kenya, one at Debre Zeit, Ethiopia, and one at St. Paul, MN, USA. Square and circle symbols (with the same colors as the bars) represent the average disease severity of the parents MN06113-8 and RB07, respectively, in each environment.

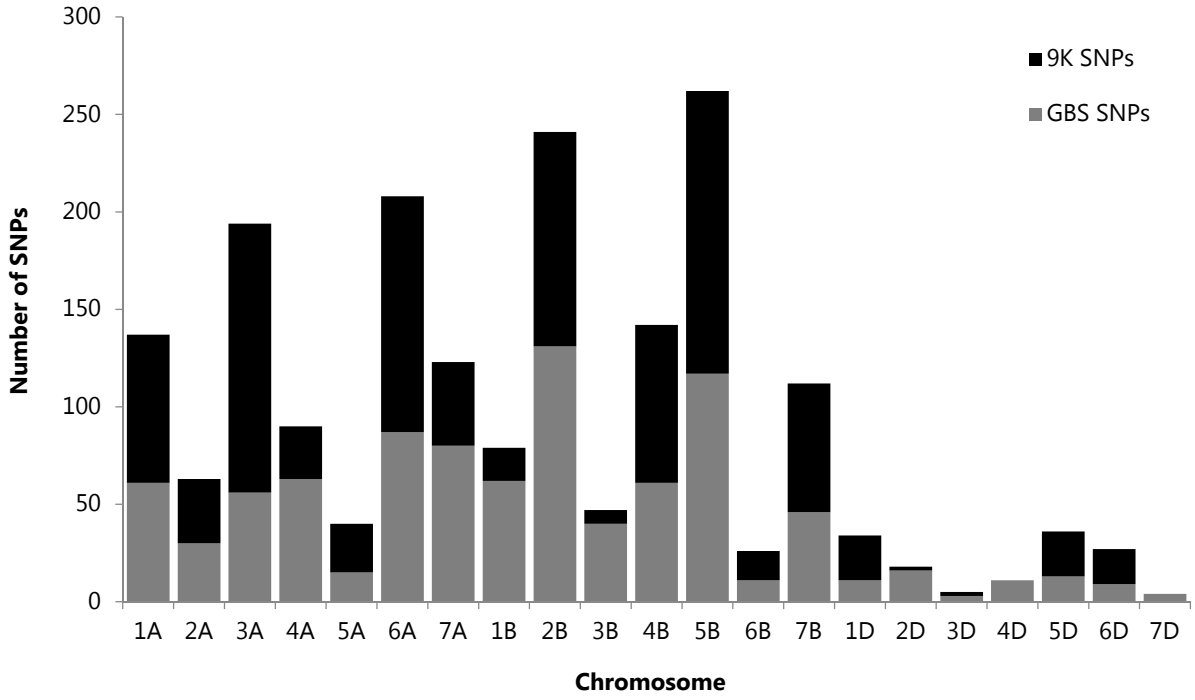


Figure 2: Distribution of markers (both 9K SNPs and GBS *de novo* SNPs) across the 21 wheat chromosomes. Markers from multiple linkage groups were pooled if a chromosome had more than one linkage group.

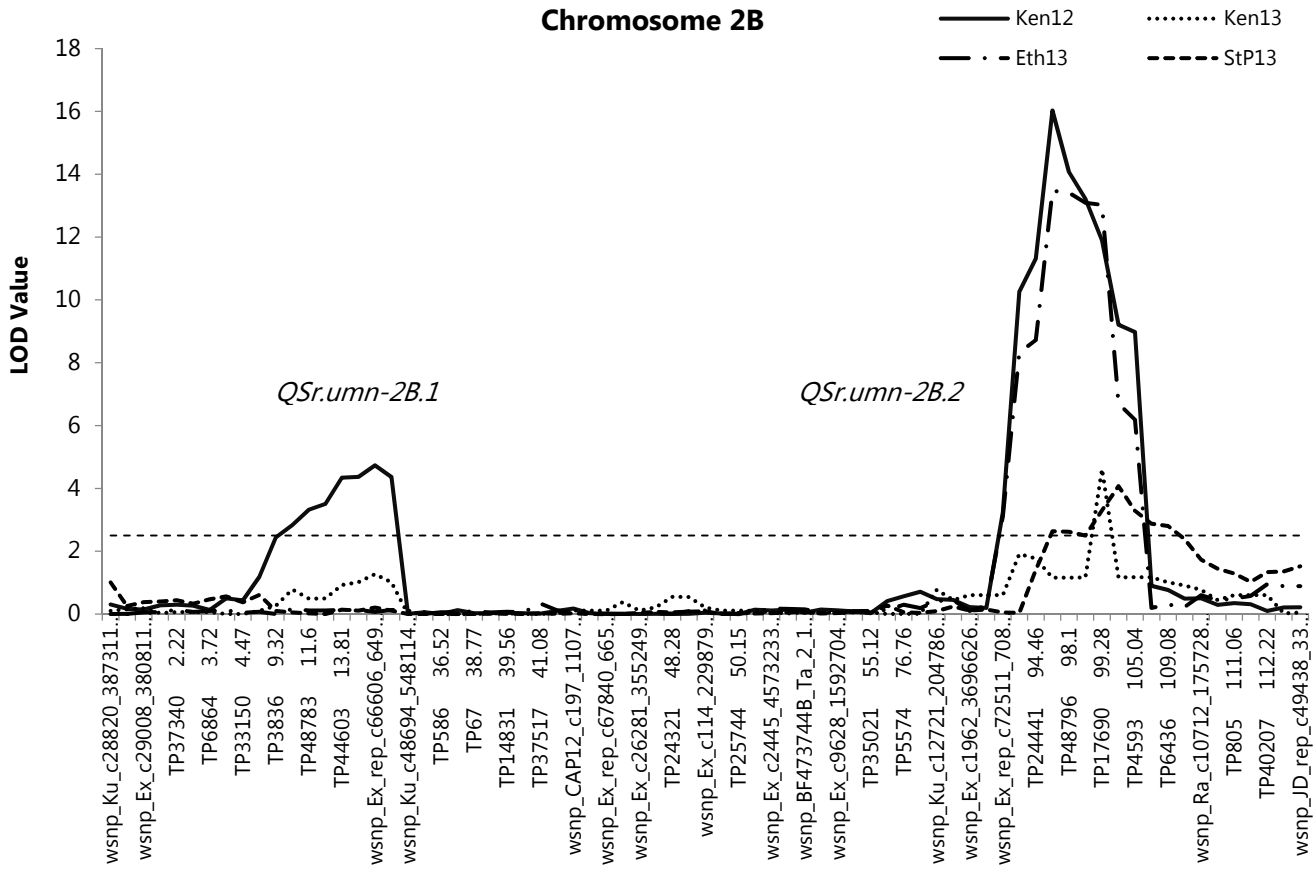


Figure 3: Quantitative trait loci (QTL) interval map for *QSr.umn-2B.2* detected on wheat chromosome 2B. The QTL is associated with stem rust resistance in the RB07/MN06113-8 population of recombinant inbred lines, and provides resistance to the Ug99 race group as well as North American *Pgt* races. The Y-axis indicates the logarithm of the odds (LOD) values with the dotted line representing the threshold LOD score of 2.5. The X-axis is labeled with SNP markers and the genetic distances in centimorgans (cM) between the markers. Markers beginning with 'TP' are *de novo* SNPs discovered from GBS approach. Ken12 = Kenya 2012; Ken13 = Kenya 2013; Eth13 = Ethiopia 2013; StP13 = St. Paul 2013.

Chapter 3

Comparison Between SNP-chip based Genotyping and Genotyping-by-Sequencing Approaches in Linkage Mapping of Quantitative Trait Loci in Wheat

In the past decade, high-throughput genomic technologies have facilitated genomic studies to be more efficient and economical. Generation of abundant, high quality polymorphic markers by such platforms have allowed for in-depth studies of organisms with both simple and complex genomes. Single nucleotide polymorphism (SNP) markers are preferred for genotyping in high throughput platforms because of their abundance in the genome, and lower cost per data point compared to older marker technologies. Genotyping by sequencing (GBS), a relatively newer approach of genotyping, is even more appealing because of its even lower cost per data point, and also due to the avoidance of ascertainment bias during genotyping. In this study, we compared the SNP-chip based genotyping method with GBS approach for several attributes, including cost per polymorphic marker, distribution of markers on a linkage map, and the genome size and recombination sites covered by the markers, as well as an overall comparison of the methodology and workflow. Additionally, imputation accuracy of the GBS dataset was examined by introducing missing values randomly and imputing the missing SNPs using a probabilistic principal components algorithm. Imputation results suggest recovery of the missing alleles with reasonable accuracy in datasets with low (up to 40%) amount of missing data is possible and can provide acceptable accuracy

in gene mapping. Overall, the comparative results indicate that the GBS approach provides equal or more genome coverage and similar mapping results at a much lower cost per data point compared to the SNP-chip genotyping method.

INTRODUCTION

The era of molecular markers, which began in earnest with restriction fragment length polymorphism (RFLP) markers in the late 1980s, has evolved remarkably. RFLP markers were followed by PCR-based marker technologies such as randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers in the late 1990s (Gupta et al. 2010). The discovery and implementation of these marker systems facilitated the mapping of genes controlling many traits, including several stem rust resistance genes in wheat. Recently, SNP markers are increasingly being used in gene and QTL mapping approaches, primarily because of the lower cost per data point and the relative ease in assay-design as well as in scoring and interpretation of the results. SNP-chip based genotyping is adaptable to high-throughput systems.

The SNP markers on SNP-chips are discovered in diversity panels, and selected to be included on the genotyping chip based on allele frequency, polymorphism information content, and marker segregation, among other factors. However, SNP markers selected via this approach are known to introduce ascertainment bias to downstream data analysis and mapping procedures (Albrechtsen et al. 2010; Li and Kimmel 2013). Ascertainment bias is a type of sampling bias that gets introduced during the SNP discovery process in which either 1) a small number of non-random individuals are used for SNP discovery, or 2) SNPs with allele frequencies significantly different than that in the population under study are selected for genotyping. Ascertainment bias can impact the study of genetic relationships among individuals that may not be easily corrected (Moragues et al. 2010;

Frascaroli et al. 2013). Therefore, to avoid false discoveries and erroneous interpretation of data, the use of genotyping methods that avoid this bias are preferred. Genotyping by sequencing (GBS), driven by the dramatically lower sequencing cost and freedom from ascertainment bias (Poland et al. 2012a), is one such method as SNPs are discovered directly on the population of interest.

The GBS procedure, also known as restriction site associated DNA (RAD) sequencing, allows the discovery of population-specific SNPs from sequencing of DNA libraries obtained after restriction-digest of samples. The usability of this approach in genotyping a crop species was first demonstrated in maize (Elshire et al. 2011), and has been successfully used in other crop species including wheat (Saintenac et al. 2013a), and barley (Liu et al. 2014). In addition to the high-quality polymorphism data free of ascertainment bias produced by GBS, the lower cost per sample (and also cost per polymorphic marker) gives this approach greater appeal for genomic studies in crops and non-crop species. Because the GBS approach yields sequence data as opposed to only the allele calls in other existing genotyping approaches, the likelihood of using the sequences for *in silico* annotation for a given functional polymorphism adds to the functionality of this genotyping approach.

One potential problem of the GBS approach is the generation of significant amounts of missing allele calls when a large number of lines are multiplexed during sequencing (Williams et al. 2010; Poland et al. 2012a; Fu et al. 2014). As more samples are multiplexed for sequencing, the number of reads per sample decreases, leading to a higher frequency of missing allele calls during SNP detection in the population.

Construction of accurate linkage maps may not be possible with a genotype matrix with a large proportion of missing allele calls. There are two ways of solving this problem: 1) increasing the sequencing depth either by multiplexing fewer individuals to generate DNA libraries or by sequencing the libraries multiple times; and 2) by imputation of missing genotyped calls using the relationship among lines in a population. Of the two, the former might be less appealing, as it requires additional time and resources to sequence and analyze the DNA libraries multiple times. The latter is increasingly preferred, because of the richness in data mining offered by the GBS procedure expands to the ability to impute missing haplotypes in related individuals with low read coverage during sequencing. Recent studies in several crop species, including wheat, have shown that imputing markers to construct missing haplotypes can be done with relatively high accuracy (Rutkoski et al. 2013; Fu 2014). Thus, in mapping populations, where the individuals are either linked in families (F_2, recombinant inbred line (RIL) populations) or via population structure (association mapping panels), the utilization of GBS SNPs appears promising. Imputation of missing genotypes might be most beneficial in populations where the allelic origin based on the parents is known or if co-ancestry information can be utilized. However, caution must be applied when doing marker imputations, especially in unrelated individuals such as in an association mapping panel.

The main objective of this study was to compare the GBS approach with the 9000 Infinium iSelect SNP assay (9K) (Cavanagh et al. 2013) based genotyping method in mapping QTL associated with resistance to stem rust of wheat. We also compared attributes such as recombination frequency, and genome coverage between these two

high-throughput genotyping methods and explore the practicality of using one approach over another.

MATERIALS AND METHODS

Molecular Marker Assay

Genomic DNA was extracted from ground seeds of the parents 'MN06113-8' and 'RB07' and 141 $F_{6:7}$ RILs using a modified cetyltrimethylammonium bromide (CTAB) protocol (Kidwell and Osborn 1992). The extracted DNA was quantified using an ND 1000 Spectrophotometer (NanoDrop Technologies, Montchanin, Delaware, USA). The population was genotyped using SNP markers obtained from two approaches: 1) the 9000 Infinium iSelect SNP assay (Cavanagh et al. 2013) and 2) genotyping by sequencing (GBS) (Elshire et al. 2011).

For genotyping using the 9K Infinium iSelect assay, DNA suspended in ddH2O at approximately 80 ng/µl was submitted to the USDA-ARS Small Grain Genotyping Center, Fargo, ND, USA. The data generated was manually called using Illumina's GenomeStudio 2011.1 (Illumina Inc., Hayward, CA). Briefly, each SNP call across the RIL population was manually analyzed, curated, and exported in diploid format (AA/AB/BB). Monomorphic markers, markers with the same calls for the entire population, markers with more than 10% missing data, and markers that deviated from 1:1 segregation ratio were discarded. Markers with 5% or less heterozygous calls were retained to avoid false purging of heterozygous loci. This resulted in 1,050 high quality markers that were retained for linkage mapping.

In the GBS approach, a double-digested library was created using the restriction enzymes *Pst*I and *Msp*I on 200 ng of DNA/sample, following Poland et al. (2012a). Each library was 76-plexed, with the parents repeated six times each, and the libraries were

sequenced in two lanes of Illumina HiSeq 2000, generating 100 bp paired-end sequences. The sequences were processed via the UNEAK pipeline (Lu et al. 2013) using the parameters -c 10 –e 0.025 to obtain *de novo* SNPs. Reads containing SNPs were used as query sequences and blastn-searched against the wheat chromosome survey sequences (WCSS) to assign SNPs to unique chromosomes. The wheat CSS sequences were obtained by assembling reads obtained from sequencing flow-sorted wheat chromosomes from the 'Chinese Spring' variety (International Wheat Genome Sequencing Consortium, http://wheaturgi.versailles.inra.fr/Seq-Repository/). To ensure that correct SNPs were obtained, only the full length alignment of a query sequence with the survey sequences allowing either one base mismatch or one gap was permitted. To circumvent retaining redundant SNPs on paralog sequences and duplicated regions among the A, B, and D subgenomes, SNPs thus obtained were filtered to remove those that mapped more than once to multiple chromosomes. SNPs that were monomorphic, had no allele calls for >10 individuals (>7% missing data), or were heterozygous in >10 individuals (7% heterozygosity) were also discarded. SNPs obtained after these steps were converted to diploid format (AA/AB/BB) from allelic phases (A/C/G/T). This process resulted in 932 high quality SNP loci that were retained for linkage mapping.

Linkage Map Construction and QTL Mapping

SNPs obtained from both genotyping approaches (9K, GBS) were combined to assign markers to linkage groups. Linkage groups were constructed using Mapdisto version 1.7.7.0.1 (Lorieux 2012) using a minimum logarithm of odds (LOD) value of 3.0. Genetic distances between the markers were calculated based on the Kosambi mapping

function (Kosambi 1943). Phenotypic data (stem rust severity) were collected on the RB07/MN06113-8 recombinant inbred line (RIL) population (Chapter 2 of this thesis). The program Windows QTL Cartographer 2.5_011, which implements the composite interval mapping (CIM) method to identify QTL, was used to analyze marker-trait associations (Wang et al. 2012b). A walk speed of 1 cM was used for QTL detection on linkage groups, and a QTL was declared to be present if the LOD threshold was calculated by 1,000 permutations at $\alpha = 0.05$. QTL effects were estimated as the proportion of phenotypic variance (R^2) explained by the QTL.

Imputation of GBS SNPs

One problematic property of the GBS approach is the generation of substantial amount of missing genotype data (Poland et al. 2012a; Fu et al. 2014). Construction of linkage maps for genome mapping can be a difficult task with a dataset that is missing a significant portion of allele calls. In this dataset, the samples were 76-plexed, and therefore the issue of missing data was not egregious (< 7% missing data; see Chapter 2 of this thesis). However, to simulate scenarios where missing data could be a problem, the genotype matrix comprising of 932 GBS SNPs for the RIL population was modified to introduce missing allele calls. Missing values ('NA') were introduced randomly in the GBS dataset using R 3.0.2 (R Development Core Team, 2013) to simulate the genotype matrix with 20%, 30%, 40%, 50%, 60%, 75%, and 90% missing data. These datasets are hereafter referred as GBS20, GBS30, GBS40, GBS50, GBS60, GBS75, and GBS90, respectively.

Imputation of missing SNPs on the simulated data was done using principal component analysis (PCA) based imputation using the probabilistic PCA (ppca) algorithm in the R package 'pcaMethods' (Stacklies et al. 2007). The ppca algorithm first assigns row average values to the missing values, and then uses the singular value decomposition of the SNP matrix to create orthogonal principal components. In turn, the principal component values corresponding to the largest eigenvalues are used to reconstruct the missing SNP genotypes in the genotype matrix. The algorithm 'ppca' was chosen for its high imputation accuracy and efficiency in regards to the use of computational resources compared to other imputation algorithms of similar caliber (Moser et al. 2009; Fu 2014). In the algorithm, 25 PCA values were used to reconstruct all genotype matrices with missing data. Prediction values < 2 were assigned the homozygous genotype '1' to represent alleles originating from the first parent; values equal to 2 were assigned the heterozygous genotype '2'; and values greater than 2 were assigned the homozygous genotype '3' to represent alleles originating from the second parent. These values were assigned after careful and replicated manual scans of the imputed data to predict the true genotypes as much as possible.

Summary Statistics of Genotype Matrices

Both non-imputed and imputed datasets were analyzed using PowerMarker V3.25 (Liu and Muse 2005) to estimate population statistics. The method of moments estimator was used to estimate the within-population inbreeding coefficient and relatedness between the individuals (Ritland 1996). An unbiased estimator that uses the inbreeding coefficient calculated using the method of moments in the previous step was used to

calculate gene diversity. Polymorphism information content (PIC), a diversity measure among the individuals in a population, was calculated according to Botstein et al. (1980). Estimation of these population statistics was done with 10,000 nonparametric bootstraps across different loci at a confidence interval of 95% ($\alpha = 0.05$). The number of recombinations in each RIL on all datasets (9K, GBS-unimputed, and GBS-imputed) was estimated using the R package 'hsphase' (Ferdosi et al. 2014).

We were also interested in knowing the distribution of recombination sites across the genome, and specifically the distribution within chromosomal arms. Although recombination in the wheat genome differs by chromosome, gene-rich regions (GRRs) are known to be the recombination hotspots (Sandhu and Gill 2002; Sidhu and Gill 2004). In general, studies report that most recombinations occur on distal 20-50% of both long and short arms of the chromosomes (Lukaszewski and Curtis 1993; Faris et al. 2000). Therefore, recombination events were calculated on the distal 40% of each arm versus the remaining region on each chromosome. The distal 40% of each arm was determined as 40% length of the genetic distance of the chromosomal arm.

Methodology and Workflow Comparison

The cost, time, and resources required to genotype the RIL population using both methods were compared so that some characteristics between the two methods could be understood to determine their usability in programs and/or projects that are similar to ours. The discussed cost estimate strictly pertains to the genotyping cost, and does not include the cost of the manual labor involved. The amount of time required to genotype the population is the 'active' time used in preparation of the DNA samples, sequencing,

and data analysis; and does not include the latent time between procedures. Computational resources and skills needed to analyze the data obtained from both procedures are also briefly discussed.

RESULTS AND DISCUSSION

Genotype Properties

The genotype calls obtained from both the 9K SNPs and the GBS SNPs were first analyzed for deviation from an expected segregation ratio of 1:1. However, as described in the 'Materials and Methods' section, the 9K dataset (1,050 SNPs) was devoid of any markers deviating from the 1:1 segregation ratio as SNPs that deviated from the ratio were discarded. Of the 932 GBS SNPs, 164 SNPs were found to be skewed towards either parental genotype. Of these 164 SNPs, 48 SNPs were over-representative of the MN06113-8 genotype, whereas 116 SNPs were over-representative of the RB07 genotype. Overall, 49.5% of marker genotypes were inherited from MN06113-8 in the 9K data, with 50% inherited from RB07, and 0.5% heterogyzous genotypes. In the GBS data, 46% of the marker genotypes originated from MN06113-8, 49% from RB07, 3% were heterozygotes, and 2% had missing data. The RILs used for genotyping were inbred to $F_{6:7}$ generation, and as such, only 1.6% of genotype calls in the population were expected to be heterozygous; and these numbers observed in both genotyping methods are expected of a highly inbred population. These results also indicated a higher proportion of heterozygous and missing allele calls in the GBS dataset, as permitted during the filtering of genotype calls.

Linkage Groups Construction

Linkage groups for both 9K and GBS SNPs datasets were constructed using the same parameters in Mapdisto version 1.7.7.0.1 (Lorieux 2012). Of 1,050 9K SNPs, 964 SNPs were placed in 30 linkage groups representing 18 wheat chromosomes.

Chromosomes 2D, 4D, and 7D were not represented by any SNPs. The number of SNPs per linkage group ranged from 2 (chromosomes 2A, 3D, 6B) to 145 (chromosome 5B) with an average of 32 SNPs. The 964 SNPs distributed over the 30 linkage groups covered a total of 1,294 cM of the wheat genome. The sizes of the smallest and the largest linkage groups were 0.4 cM and 168 cM, respectively, with an average size of 43 cM. Similarly, 925 of the 932 GBS SNPs were assigned to 31 linkage groups that represented all 21 wheat chromosomes. The number of SNPs in these linkage groups ranged from 2 (chromosomes 4A, 6B, 6D, 7D) to 131 (chromosome 2B) with an average of 30 SNPs per linkage group. The smallest linkage group was 0.8 cM, the largest group was 147 cM long, and the average size of all linkage groups was 42 cM. The size of the wheat genome covered by the 31 linkage groups constructed using 925 GBS SNPs was 1,305 cM. Properties of the linkage groups constructed using the two genotyping methods are presented in Table 1.

One particular advantage GBS has over the SNP chip based genotyping is the slightly better coverage of the wheat D genome. The wheat D genome is often the least represented in genotyping platforms, owing to its lower frequency of polymorphic sequences (Chao et al. 2009; Allen et al. 2011). While the number of markers mapped to the D genome in our GBS dataset is not large compared to the A and B genomes, the retrieval of all seven D chromosomes during linkage mapping is a good indication that the GBS approach can be manipulated to obtain more SNP markers from the D genome. One possible way to achieve this is by using less stringent filtering parameters than the ones used in our study. However, a more reliable approach may be to map the reads

obtained from sequencing of GBS libraries to the respective sub-genomes of wheat in order to obtain sub-genome specific polymorphic markers. With two of three wheat's subgenomes already sequenced the A genome from wheat's diploid ancestor *Triticum urartu* (Ling et al. 2013) and the D genome from *Aegilops tauschii* (Jia et al. 2013) – sub-genome specific mapping of sequence reads would help to retain reads that would otherwise be discarded during SNP-calling due to mismatches among the reads originating from the A, B, and D sub-genomes.

Recombination and Genome Coverage

To better understand the distribution of parental genotypes in the population, the number of recombination events in each RIL was estimated. The average number of recombinations per RIL in both 9K and GBS genotype matrices was 27. Seventy-eight (55.3%) individuals had more recombinations than the average in GBS dataset, compared to 65 (46.1%) in the 9K dataset. Only 2 individuals had the same number of recombinations in both matrices. The total number of recombinations in the 9K matrix was 3,746 which was slightly lower than that in the GBS matrix, at 3,790.

To visualize these results, recombination blocks per line observed in both genotype matrices were plotted against the SNPs in each chromosome (Figure 1). As seen in the figure, both congruous and incongruous patterns of recombination blocks exist between the two genotype matrices. The difference in reported recombination breakpoints likely arises from the difference in assay design between the two genotyping methods as different parts of the genome might have been sampled, which alter the genome coverage in these two genotyping methods. This can potentially lead to over- or

under-representation of different haplotype matrices, which results in the detection of different recombination sites and number of recombination events. The difference in genome coverage is also corroborated by the difference in properties of the linkage groups constructed using SNPs obtained from the two methods. This is illustrated in Figure 1C where the portions of the genome represented by the two methods are compared. The figure shows the differences in genome sampling between the two genotyping methods in one RIL ('MN06_01') across the chromosomes, especially visible among the chromosomes 1A, 1B, and 5A. To visualize distribution of 9K SNPs and GBS SNPs along the linkage maps, linkage maps constructed separately using 9K and GBS datasets were compared with the linkage map constructed using a combined '9K+GBS' SNPs (see Chapter 2 for more details). The results support the patterns of haplotypic distribution as observed in Figure 1C. The 9K SNPs are located away from the GBS SNPs on chromosomes 1A, 1B, and 5A in combined linkage map, owing to the differences in genome sampling between the two SNP types (Supplemental File 1). On chromosomes that show similar sampling of the genomic regions (such as 2B, 3A, and 4B), both 9K and GBS SNPs are intercalated in proximity to each other in the combined '9K+GBS' linkage maps (Supplemental File 1). This difference in SNP distribution between the two datasets illustrates that dissimilar portions of the genome are represented by the two genotyping approaches. Such differences are also observed in all RILs in the population in general (data not shown). We believe that the accumulation of such differences over the whole genome across the population is the reason behind the observation of incongruous recombination patterns.

As expected, the distribution of recombination events favored the distal ends of the chromosomes compared to the centromeric regions in both datasets. In the GBS dataset, 1,523 recombination events were observed on the distal 40% of both arms (623 on short arms, 900 on long arms), and 1,257 in the centromeric region. Similarly, 843 recombination events were observed on the distal 40% of both arms in the 9K dataset (396 on short arms, and 447 on long arms), with the centromeric region recording 706 recombinations. In our population, we observe that the distal ends have 10% more recombinations in the GBS dataset, and 9% more recombinations in the 9K dataset, relative to the centromeric regions in each respective dataset. The highest number of recombinations among all three sub-genomes in the GBS dataset was observed in the B-genome (1,608 recombinations) whereas the A-genome recorded the most recombinations in the 9K dataset (1,791 recombinations). The D-genome in both datasets had the fewest recombination events with 691 recombinations in the GBS dataset, and 369 in the 9K dataset.

To uncover the genetic architecture controlling the traits of interest, the use of molecular markers representative of the genome is important in gene mapping studies. Markers that are significantly linked to the trait can provide remarkable improvements in breeding for allele enrichment and trait improvement. Since the choice of genotyping platform can impact the quantity of molecular markers and their distribution in different genomic regions, understanding the differences in marker-related genome properties can assist in such choice among the different available genotyping approaches. In our investigation of such properties between the two genotyping methods discussed here, the

difference in number of polymorphic markers identified and used in creating linkage groups was not strikingly different; whereas the distribution of those markers in different genomic regions as shown by the difference in sites of recombination was noteworthy. The difference in genome coverage is also corroborated by the difference in properties of the linkage groups constructed using SNPs obtained from the two methods, and the differences in QTL detection between the two methods, as described in upcoming sections.

QTL Mapping

Mapping of QTL associated with resistance to stem rust of wheat was carried out in the RIL population using stem rust severity data collected at three locations over four seasons. This was done to assess the impact on QTL mapping using linkage groups constructed from markers obtained from the two genotyping approaches. For detailed information on disease phenotyping and data statistics, see Materials and Methods, Chapter 2 of this thesis.

The CIM method of QTL mapping detected nine QTL in the 9K dataset on linkage groups representing the chromosomes 2B, 3A, 4A, 4B, 5B, and 6D (Appendix I). Similarly, eight QTL distributed on chromosomes 1A, 2A, 2B, 2D, 4A, 4B, and 7A were detected in the GBS dataset (Appendix I). Four QTL detected in the 9K dataset, namely the QTL on chromosomes 2B (*QSr.umn-2B.1*, *QSr.umn-2B.2*), 4A (*QSr.umn-4A*), and 4B (*QSr.umn-4B.2*) were also detected in the GBS dataset. The similarity of the QTL was determined based on their location, parent contributing the QTL allele, and SNPs associated with the QTL (Table 3, Chapter 3 of this thesis). QTL not common between

the 9K and GBS datasets suggests that the portions of the genome represented by the SNPs obtained from the two genotyping approaches are different, as we discussed earlier. However, uniform genome coverage, inclusion of more environments, or highly correlative phenotype data between the environments might absolve this issue. On the whole, both methods performed reasonably well, and were able to predict the location and effect of QTL involved in stem rust resistance.

Imputation of GBS Markers

One advantage of the GBS approach over the chip-based genotyping approach is the flexibility of the system that allows some control over the number of polymorphic markers that can be generated among lines in a population. An imputation study has not been conducted in an RIL population yet, therefore we studied the effect and accuracy of genotype imputation on the GBS dataset after simulating datasets with several successive missing proportions.

The highest imputation accuracy was observed when the proportion of missing data was the lowest (Figure 2). The missing genotypes were predicted with accuracy of 96% when the dataset was missing 20% of the genotypes. The accuracy of genotype imputation was reduced as datasets had a higher proportion of missing data, with a major drop-off in datasets > 60% missing data, which had imputation accuracies of < 84%. The effect of missing data did not appear to impact population parameters such as gene diversity, polymorphism information content (PIC), and the amount of heterozygosity. An overall decreasing trend for each of these population characteristics can, however, be observed with increasing missing values in the dataset. The inbreeding coefficient had an

overall increasing trend as datasets had more missing proportion of genotype calls. There was no significant change in the allele type from imputation in the imputed datasets, except in the GBS90 dataset where the proportion of RB07 alleles increased by 3%.

Our dataset contained only 932 high-quality GBS markers, which is sufficient for gene mapping studies, given the high LD of hexaploid bread wheat (Chao et al. 2007; Chao et al. 2009). However, the marker imputation results presented here show that more *de novo* markers obtained from the GBS approach with moderate levels of missing allele calls (up to 40%) may be used in order to obtain higher resolution in genome mapping studies. With as much as 40% missing data, the genotype matrix can be predicted with imputation accuracy of 90% or higher. While this may potentially introduce some biases in the study, the level of gene diversity, PIC, and heterozygosity were not significantly altered, implying that the population does not deviate significantly from the expected levels of inbreeding. This is illustrated in Figure 2, where the inbreeding coefficient increases negligibly from non-imputed GBS dataset to GBS90 dataset. The observed slight increase is most likely due to the introduction of false genotype calls in imputed datasets with a higher proportion of missing data. In most crop breeding programs, higher level of inbreeding is desired so that the allelic composition of genes that control the desired traits in the breeding germplasm may be preserved. Yet, the consistency in inbreeding levels among the imputed datasets observed here should not be the only determinant behind the use of a dataset with large amounts of missing allele calls. Datasets with large amounts of missing data introduce severe problems such as inaccurate and inflated linkage groups, and erroneous QTL detection.

QTL Mapping Using Imputed Datasets

The effect of marker imputation in construction of linkage groups and QTL mapping in a biparental population has not been investigated yet. Thus, we used imputed GBS datasets, as described in the Materials and Methods section, to investigate QTL mapping. Construction of linkage groups and QTL mapping, however, were carried out using the imputed dataset with 40% and 75% missing data only, which had imputation accuracies of 90% and 71%, respectively.

Twenty-nine linkage groups were formed in the GBS40 dataset, and covered 4,785 cM of the genome (Table 2). Similarly, 30 linkage groups were formed in the GBS75 dataset, covering 14,879 cM of the genome. Genome sizes represented by these linkage groups are approximately 3.5 times and 11 times larger than the size covered by the original, non-imputed GBS dataset. This increase in genome size is most likely due to the introduction of inaccurate genotype calls in the dataset with the higher proportion of missing values. Inflation of linkage groups by introduction of false genotype calls is also supported by the average marker interval distance of 5 cM in the GBS40 dataset and 19 cM in the GBS75 dataset. Therefore, marker imputation in datasets with a large proportion of missing data appears to introduce errors, and as such, avoiding the use of such datasets is pragmatic. As the imputation accuracy dropped, more markers were also unlinked to any linkage group (Table 2). Two linkage groups (chromosomes 3D and 7D) were not detected in the GBS75 dataset.

The number of QTL discovered in both GBS40 and GBS75 datasets using the CIM approach was 8 and 7, respectively (Table 2). However, most of the QTL detected

in these datasets were different relative to the non-imputed dataset, with only one consistent QTL among all datasets (Appendix I). The large-effect QTL observed on chromosome 2B in the non-imputed GBS datasets was the only consistent QTL in the GBS40 and GBS75 datasets, with lower accuracy of the QTL positions in the imputed datasets (Figure 3, A-C). Only the GBS40 dataset correctly predicted the same SNPs (TP24441 and TP17690) linked to the large-effect QTL as predicted in the non-imputed GBS dataset. Although the large-effect QTL is still detected in the GBS75 dataset, the QTL was incorrectly placed near the end of the linkage map; and the significant markers TP24441 and TP17690 were incorrectly assigned wrong map positions. Both imputed datasets predicted the percentage of phenotypic variation and allelic effect similarly to the non-imputed dataset. The small-effect QTL observed in the Kenya 2012 environment was detected in both imputed datasets, although its position and linked markers in GBS40 and GBS75 were different relative to the non-imputed dataset. It is likely, given the inflation in size of linkage groups experienced with imputing, that most of the QTL detected in the imputed datasets are inaccurate. While validation of the detected QTL would provide a definitive answer, the results indicate that large-effect QTL can be detected if the dataset comprises a large proportion of imputed genotypes. The small-effect QTL may also be detected, but such prediction might not necessarily be accurate.

Comparison of Methodology and Required Resources

With regard to the methodologies and workflow, chip based genotyping is relatively easier than the GBS approach, though the latter had a faster turnaround time in our case. The 9K genotype calls were obtained from the USDA Genotyping Facility at

Fargo, ND that needed manual inspection using Illumina's GenomeStudio program version 2011.1 (Illumina Inc., Hayward, CA) before use. In the GBS method, DNA libraries for sequencing were generated in-house in less than two days. The sequences were filtered based on barcodes and trimmed before calling SNPs in the population. In addition, sequence alignment to the WCSS and further data parsing was required. Thus, the need for high-end computational resources and bioinformatic expertise (ability to work in unix environment, as well as programming skills) is essential in the GBS approach to manage and work with the large amount of sequence data generated from parallel sequencing. On average, 256 gigabytes (gb) of memory was requested on any available node with the Minnesota Supercomputing Institute (MSI, https://www.msi.umn.edu/) while working with the GBS procedures such as quality control of the sequences, SNP calling, and sequence alignment. Similarly, several hundred gb of hard disk space was needed to store the sequence files and any output files created during the procedures mentioned above. The SNP chip based method, however, required less computational resources but a proprietary program (GenomeStudio) was needed to visualize and analyze the data generated from the genotyping assay. The program was run on a Windows machine with 64 gb of available memory. The hard disk space requirement to store the data files was less than 50 megabytes.

The biggest advantage of the GBS approach however, is perhaps the economical aspect of this method. In our study, we obtained a comparable number of usable SNPs for QTL mapping from both genotyping approaches (964 from 9K, 925 from GBS). The cost per 9K SNP used for mapping was approximately $8.20, whereas the cost per GBS SNP

used in QTL mapping was approximately $2.10. These figures are exclusive of the labor cost, in which the GBS method is also advantageous over the chip-based genotyping method. Although several filters can be applied within the program GenomeStudio to parse the genotype calls obtained using the 9K chip, the genotypes still need to be manually inspected for each SNP between the two parents in order to re-cluster the individuals to distinct genotype groups. In our dataset, 2,524 polymorphic SNPs were obtained after applying the filter to remove SNPs monomorphic between the two parents. Each of these SNPs had to be inspected to assign the correct genotype calls to the RILs. At the inspection rate of approximately three SNPs every two minutes, the total time required to tag the population with correct genotype calls was approximately 28 hours. As the program did not allow correction of the incorrect genotype calls, the exported data had to be edited to assign final genotypes to each individual. On the other hand, the GBS procedure required less than 3 hours to obtain the final genotype calls. Creating the input file with each individual labeled with the barcode used during library preparation was essential before running the SNP-calling program UNEAK. This task was completed in about 30 minutes, followed by approximately one hour to obtain the SNPs from UNEAK. As the allele calls were reported in base format (AA/CC/GG/TT), they were converted to a biallelic format (AA/BB), which was accomplished in about 30 minutes. While the time needed to troubleshoot errors that appeared during both procedures has not been discussed here, the GBS approach was more efficient because of its lower economical burden and advantages in automated data processing. Yet, as we indicated in the results, the genome areas targeted by these approaches are slightly different, which may lead to

detection of some different QTL between the two methods. If marker coverage were better, and perhaps uniform, between the two methods, these differences should disappear. Based on the comparative analysis, we are confident that the GBS approach can be used as an independent, standalone approach for genotyping and conducting genome studies.

CONCLUSION

At present day, the cost of genotyping is lower than ever. Highly economical genotyping approaches provide a user the options of using different types of high-throughput genotyping methods, ranging from simple sequence repeat based genotyping to sequence-based genotyping. In our study, we compared two high throughput genotyping methods used in genomic studies of wheat: SNP-chip assay, and genotyping by sequencing (GBS). The results showed that both methods are powerful means of studying the genome and provide enough resolution to carry out marker-trait association studies. The key attributes of high interest to a researcher might be the cost and data turnaround time, in which the GBS approach bests the SNP-genotyping method. The GBS approach was also able to provide a better coverage of the wheat genome, including that of the often poorly represented D-genome. The SNP-chip based genotyping however requires less computational knowledge and resources to process the data. The choice of the genotyping platform for gene mapping and other genome studies may come down to the question of cost and available resources.

Table 1: Results of linkage groups formation using the 9K and GBS SNPs.

A.

Method	9K																				Total
Chromosome	1A	2A	3A	4A	5A	6A	7A	1B	2B	3B	4B	5B	6B	7B	1D	2D	3D	4D	5D	6D	7D	
Linkage groups[a]	2	4	1	1	3	2	1	2	1	1	1	1	4	1	1	0	1	0	1	2	0	30
SNPs[b]	76	32	138	26	23	121	43	17	110	7	82	145	12	66	23	0	2	0	21	18	0	
Size (cM)[c]	75	38	156	61	31	76	105	57	141	47	94	168	34	95	73	0	1	0	24	18	0	
SNPs per sub-genome	459							439							64							962
Size per sub-genome	542							636							116							1,294

B.

Method	GBS																				Total
Chromosome	1A	2A	3A	4A	5A	6A	7A	1B	2B	3B	4B	5B	6B	7B	1D	2D	3D	4D	5D	6D	7D	
Linkage groups[a]	1	2	1	2	1	3	1	2	1	1	1	1	2	1	1	1	1	2	2	2	2	31
SNPs[b]	61	29	56	63	15	87	80	62	131	40	61	117	11	46	10	16	3	11	13	9	4	
Size (cM)[c]	95	45	146	58	23	36	98	80	111	43	90	147	25	83	29	44	23	35	64	12	18	
SNPs per sub-genome	391							468							66							925
Size per sub-genome	501							579							225							1,305

[a] Number of linkage groups formed for each wheat chromosome

[b] Number of SNPs that mapped to all linkage groups representing each chromosome

[c] Size of the linkage groups combined if >2 linkage groups were observed for a chromosome

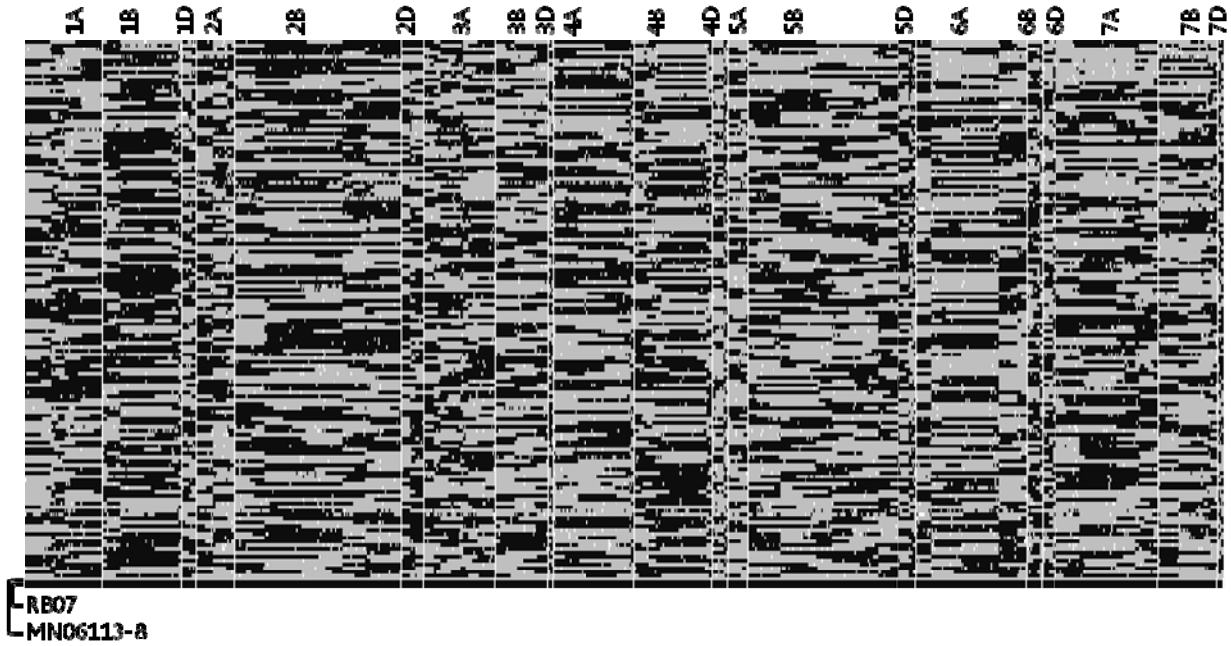
Table 2: Comparison of linkage mapping results among the 9K dataset, non-imputed GBS dataset, and datasets with 40% and 75% missing allele calls.

Method[a]	SNPs[b]	Linkage Groups	SNPs (%) in Linkage Groups	Unlinked SNPs (%)	Genome Size (cM)	QTL
9K	1,050	30	91.8	8.2	1,295	4
GBS	932	31	99.2	0.8	1,306	8
GBS40	932	29	98.8	1.2	4,785	8
GBS75	932	30	89.2	10.8	14,879	7

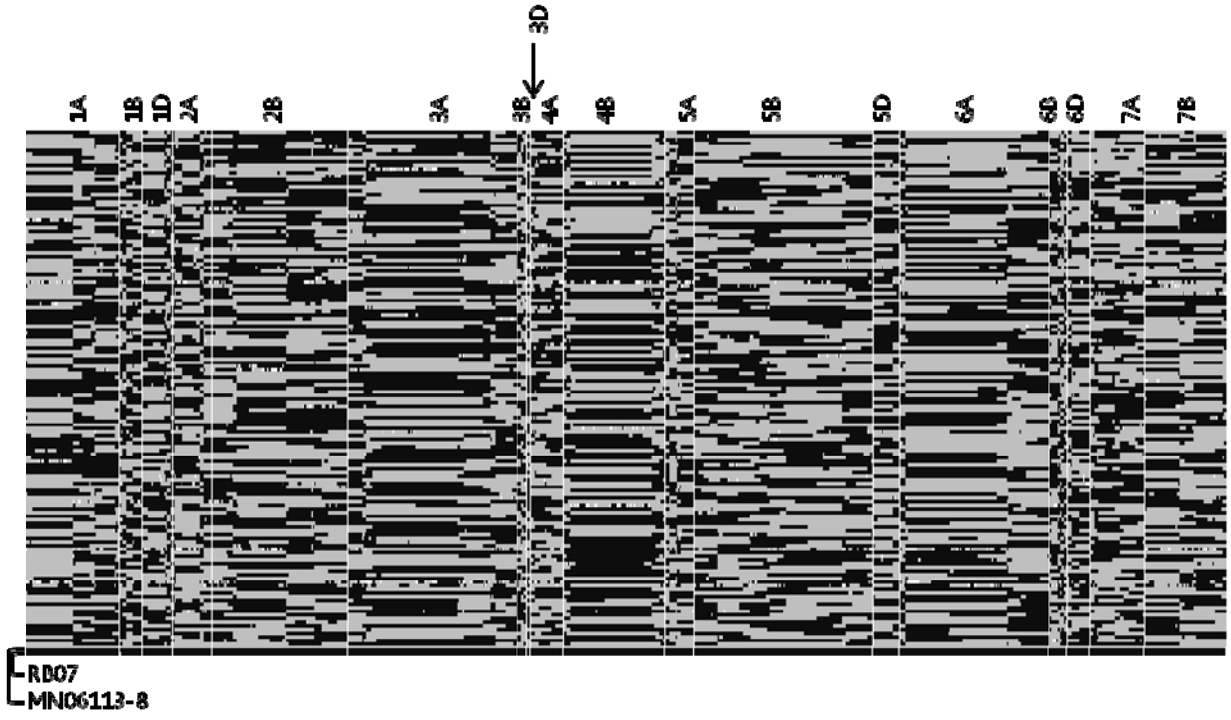
[a] GBS40 = imputed GBS dataset with 40% missing allele calls; GBS75 = imputed GBS dataset with 75% missing allele calls

[b] Number of polymorphic markers used in linkage mapping and imputation of missing genotype data

A.


1A
1B
1D
2A
2B
2D
3A
3B
3D
4A
4B
4D
5A
5B
5D
6A
6B
6D
7A
7B
7D
RB07
MN06113-8

B.


3D
1A
1B
1D
2A
2B
3A
3B
4A
4B
5A
5B
5D
6A
6B
6D
7A
7B
RB07
MN06113-8

C.

Figure 1: Recombination blocks observed in 9K (Panel A) and GBS (Panel B) genotype datasets; and a representative example of the difference in genome sampling between the 9K and GBS methods (Panel C). In Panels A and B, the population is arranged in descending order on the Y-axis and SNPs are arranged by the chromosomes they belong to on the X-axis separated by the white vertical bar. Panel C represents an example of the difference in genome sampling between the 9K and GBS methods in the recombinant inbred line (RIL) 'MN06_1' across all 21 chromosomes. The size difference within each chromosome (for example, within 1A_9K and 1A_GBS) is due to the differences in number of SNP markers between the two methods that are distributed along the chromosome. No linkage groups were obtained for chromosomes 2D, 4D and 7D using the 9K SNPs. In all panels, the colors gray and black represent MN06113-8 and RB07 haplotype blocks, respectively whereas the white dots (white vertical lines on panel C) indicate missing data.

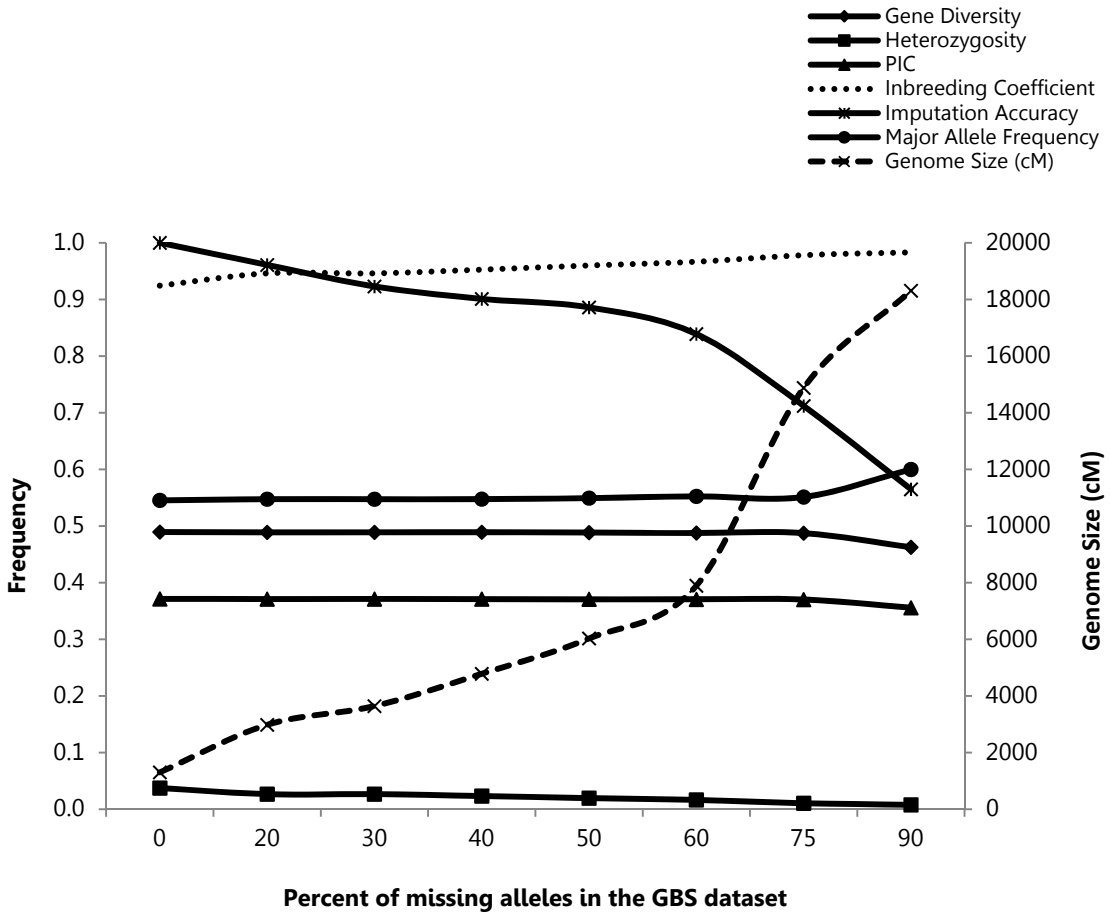
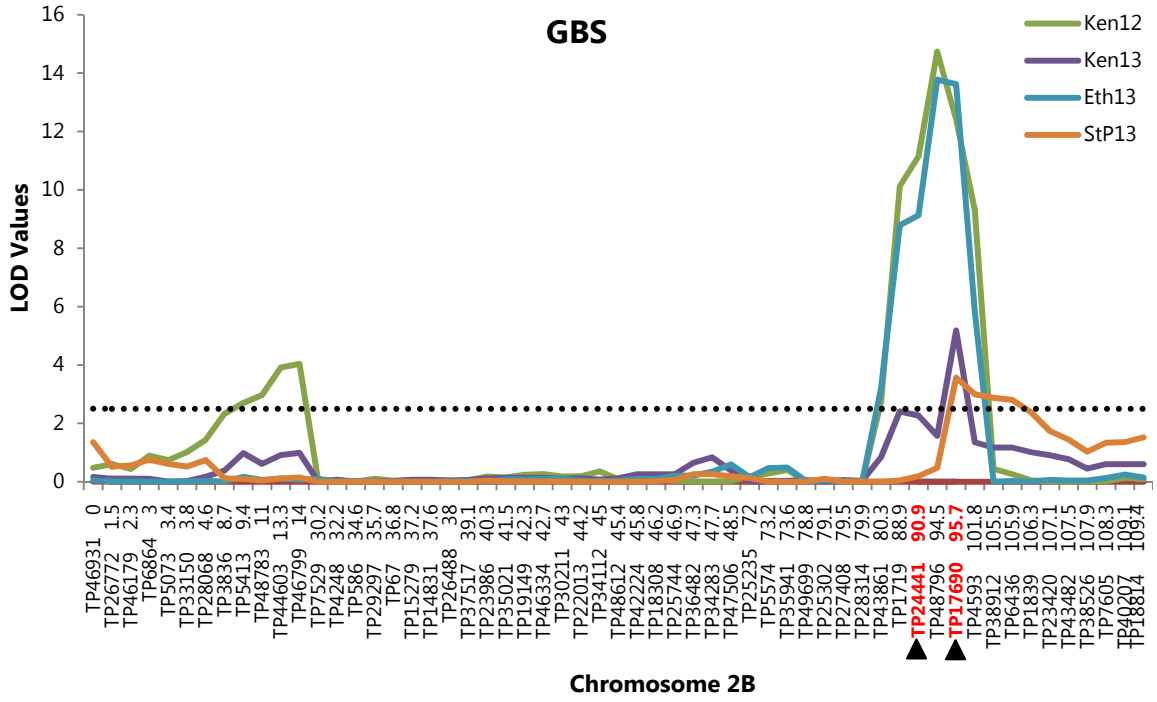
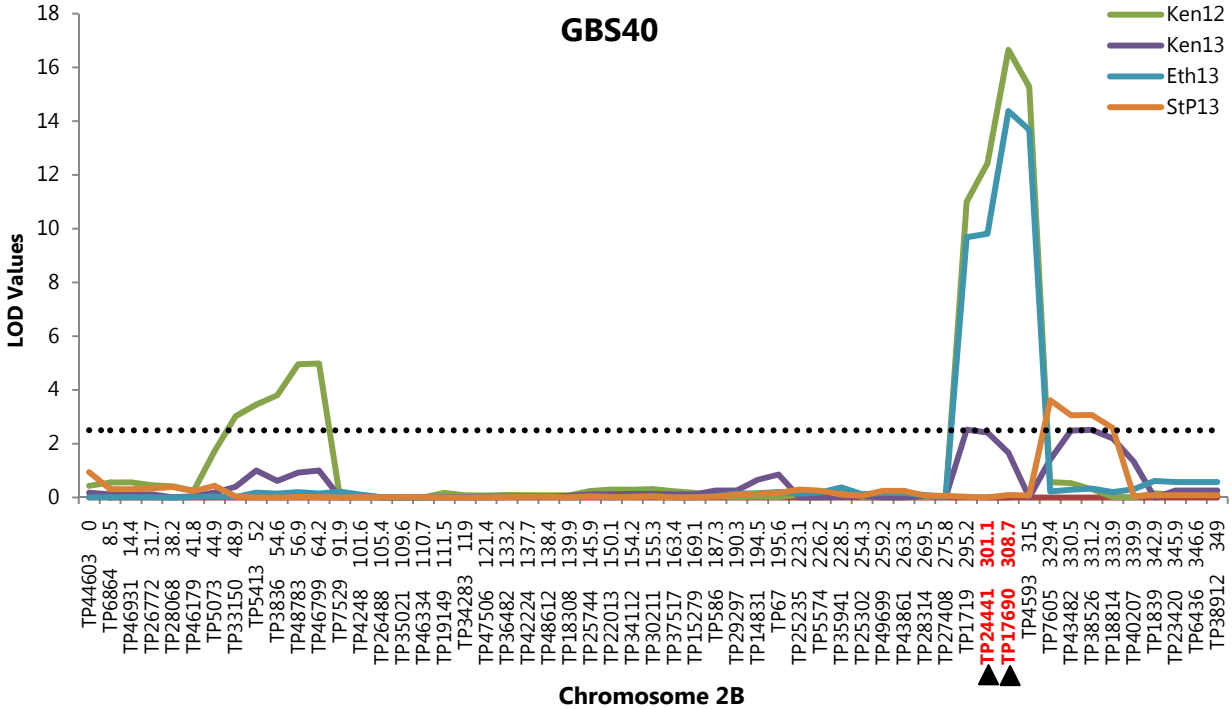


Figure 2: Characteristics of genotype matrices in imputed and non-imputed datasets. '0' represents the original GBS dataset with no missing allele calls introduced. The genome sizes (sizes of linkage groups summed together) of imputed and non-imputed GBS datasets are shown on the secondary Y-axis to the right.

A.


GBS
Ken12
Ken13
Eth13
StP13
LOD Values
Chromosome 2B

B.


GBS40
Ken12
Ken13
Eth13
StP13
LOD Values
Chromosome 2B

C.

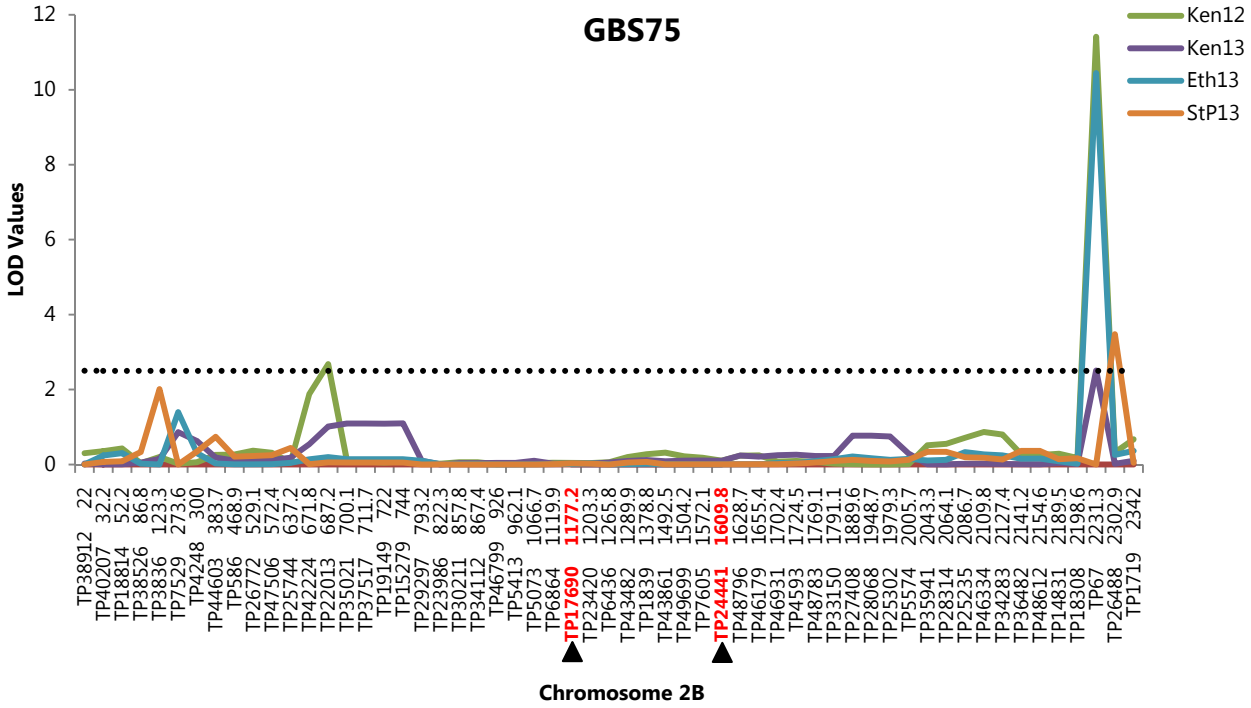


Figure 3: A comparison of the QTL detected on chromosome 2B among the three GBS datasets: non-imputed GBS dataset (Panel A), GBS40 (Panel B), and GBS75 (Panel C). The Y-axis indicates the logarithm of the odds (LOD) values with the dotted line representing the threshold LOD score of 2.5. The X-axis is labeled with SNP markers and the genetic distances in centimorgans (cM) between the markers. The significant SNP markers (TP24441 and TP17690) in non-imputed GBS dataset are labeled in red color (and black triangles) in all three panels to indicate their positions relative to the QTL peaks. Note the increase in sizes of the GBS40 and GBS75 linkage groups compared to that of non-imputed GBS dataset. In comparison with the non-imputed GBS dataset, the order of the SNP markers also changes in linkage groups constructed using genotype information in imputed datasets with higher proportion of missing alleles.

Chapter 4

Nested Association Mapping of Stem Rust Resistance in Wheat Using Genotyping by Sequencing

Nested association mapping (NAM) is an approach to map trait loci in which families within populations are interconnected by a common parent. By implementing joint-linkage association analysis, this approach is able to map causative loci with higher power and resolution compared to biparental linkage mapping. The recently developed genotyping by sequencing (GBS) approach is a relatively fast, efficient and cost-effective method of genotyping a large number of individuals. Additionally, the GBS method allows for discovery of *de novo* markers that are specific to the individuals or populations under study. We combined the NAM approach with GBS to dissect and understand the genetic architecture controlling stem rust resistance in wheat. Ten stem rust resistant wheat varieties were crossed to the susceptible line LMPG-6 to generate F_6 recombinant inbred lines (RILs). The RIL populations were phenotyped at four environments in Kenya, South Africa, and St. Paul, Minnesota, USA. We identified several minor-effect QTL contributing towards adult plant resistance (APR) to North American *Pgt* races as well as the highly virulent Ug99 race group. Validation of markers that are significantly associated with each QTL is necessary to generate diagnostic markers for marker assisted resistance breeding. The usefulness of GBS-derived *de novo* SNPs in mapping APR to

stem rust shown in this study could be used as a model to conduct similar marker-trait association studies in other plant species.

INTRODUCTION

Since the earliest days of wheat cultivation, the disease stem rust of wheat caused by the fungus *Puccinia graminis* f. sp. *tritici* (*Pgt*) has been a major constraint to wheat production in many areas of the world. Historical records show that the disease is highly damaging during epidemics and is capable of posing a serious threat to world food security. The pathogen is well known for its ability to travel long distances and capacity to evolve at the site of deposition (Hodson et al. 2011). Genetic resistance has been the main means of fighting this disease primarily because of its effectiveness. However, the pathogen has historically proven that it is able to overcome the deployed genes, giving rise to disease epidemics in regions with susceptible cultivars (Roelfs 1985a; McIntosh et al. 1995). The recent evolution and spread of the virulent African stem rust race TTKSK and its derivative races are the perfect example of this phenomenon. Race TTKSK and six other lineal races of the same group have already defeated resistance genes that had been effective for several decades (Mago et al. 2011a). As predicted by wind trajectories, the arrival of these stem rust races in the breadbaskets of the world is likely. Also, the *Pgt* races in North America are changing, as documented by detection of a highly virulent race, TTTTF, only a few years ago (Jin 2005). Therefore, discovery of novel sources of resistance and their deployment is an essential process in order to fight off this constantly evolving pathogen.

The nature of genetic resistance to stem rust of wheat is mainly qualitative, in the form of major genes derived from hexaploid bread wheat and related species. All-stage resistance (also known as seedling resistance) has been a significant part of rust

resistance breeding. One of the first resistance genes to be deployed, *Sr12* in the University of Minnesota cultivar 'Thatcher' protected the majority of the wheat acreage from rust epidemics in the northern US wheat growing regions in the 1930's and 1940's (McIntosh et al. 1995). Varieties with the gene *Sr31*, bred and distributed by CIMMYT from the mid 1960s, were popular globally until they were defeated by the highly virulent *Pgt* race TTKSK (aka isolate Ug99) in 1998 (Pretorius et al. 2000). TTKSK, along with six other races in the Ug99 lineage, has defeated almost the half of the discovered stem rust resistance genes (Jin et al. 2008; Jin et al. 2009b; Mago et al. 2011a; Pretorius et al. 2012), and the majority of deployed genes. This is one example among several cases where *Pgt* races have evolved to defeat several important stem rust resistance genes that had been protecting the wheat crop from this disease. As the selection pressure on the pathogen leads to its evolution and renders genes ineffective, the threat of crop loss from possible rust epidemics increases.

In addition to seedling resistance, adult plant resistance (APR) is considered to be effective against a wider array of *Pgt* races, and are assumed to be durable, mainly because of their nonspecific effectiveness (Lindhout 2002; Stuthman et al. 2007). APR genes are expressed mainly during the adult plant stages, and provide APR during important plant growth and development phases. While QTL mapping studies have reported several QTL providing APR against *Pgt* races, only five APR genes (*Sr2*, *Sr55*, *Sr56*, *Sr57*, and *Sr58*) has been discovered to date. As APR genes contribute less additive effectiveness compared to seedling genes, the level of resistance conferred by these genes in the field is usually inadequate. Thus, pyramiding both APR and seedling genes to

obtain durable APR is essential for effective disease control (Singh et al. 2000; Evanega et al. 2014). In addition, several of the discovered genes, mainly the ones discovered in wild species, are linked with undesired traits. Discovery of genes in breeding lines and adapted germplasm, on the other hand, would more readily facilitate the use of such genes in breeding programs.

Of various strategies implemented to map causative loci for segregating traits, nested association mapping (NAM) is used to map loci in a multi-cross mating design where one common parent is shared among all other 'founder' parents. Also known as joint mapping or joint linkage mapping, this strategy uses the benefits of both linkage mapping and linkage disequilibrium mapping to provide higher mapping power and resolution. Briefly, the strategy involves crossing several founder lines to a single common parent to generate segregating progenies in each population. The genetic background is normalized by virtue of having a common parent, which allows mapping of segregating alleles in different populations with reference to common-parent specific alleles (Blanc et al. 2006; Yu et al. 2008). Therefore, joint mapping helps to minimize problems that may arise due to genetic heterogeneity, different environmental effects, or simply experimental and sampling differences as all populations are connected by a common parent. The NAM design is also able to detect QTL with various effect sizes, including rare alleles, because of its higher statistical power. The efficacy of a NAM design in mapping important QTL have been demonstrated in recent studies in maize (Buckler et al. 2009; McMullen et al. 2009), Arabidopsis (Buckler and Gore 2007), and in a few other crop species (see Guo et al. 2010).

One key requirement of a sound mapping study is the abundance of high quality markers for genotyping the population. High marker coverage of the genome allows high resolution and accurate mapping of causative loci. Single nucleotide polymorphism (SNP) markers are preferred over other marker systems for genotyping because of their abundance, low cost per data point, and utility to high-throughput technologies. However, genotyping of populations using SNPs from pre-designed assays (or chips) is known to introduce founder ascertainment bias and are known to result in less accurate and biased data analysis (Albrechtsen et al. 2010; Heslot et al. 2013). One alternative to obtain a high number of high-throughput SNPs without ascertainment bias is the complexity-reduced genome sequencing approach called genotyping by sequencing (GBS). GBS uses restriction enzymes for targeted complexity reduction of genomes followed by next-gen sequencing of multiplexed samples (Elshire et al. 2011; Poland et al. 2012b). The millions of reads thus obtained are used to discover SNPs, thus allowing the discovery of high quality population-specific SNPs for genomic studies. This approach is also appealing because of the low cost per sample, relatively faster turnaround time, and malleability in terms of sequence manipulation and data mining. This technique has been successfully used in wheat studies to obtain *de novo* genetic maps (Saintenac et al. 2013a), and in barley to map alleles influencing plant height (Liu et al. 2014).

In this study, we use ten biparental spring wheat RIL populations to map stem rust resistance QTL. The objectives of this study were to conduct a genome-wide scan for QTL in a joint analysis of all ten populations, linked by a common parent. We use GBS markers for the first time to conduct a NAM analysis to dissect the relationship between

the genetics of a large mapping population and their resistance to African and North American *Pgt* races. All male parents used in the crosses have been released as commercial wheat varieties in their target areas. As such, we expect that the significant markers identified in this study will provide higher value and incentive towards introgression of the detected QTL for resistance breeding in areas that are hotspots and/or are vulnerable to the stem rust disease of wheat.

MATERIALS AND METHODS

Plant Material

Nine Kenyan spring wheat lines ('Fahari', 'Gem', 'Kudu', 'Kulungu', 'Ngiri', 'Paka', 'Pasa', 'Popo', and 'Romany') that were released as varieties in the 1960s, 70s, and the 80s were selected for crossing based on the high level of APR exhibited by these lines during screening in the Njoro stem rust nursery in Kenya (Table 1). 'Ada', a recent hard red spring wheat variety released by the University of Minnesota (Anderson et al. 2007), also exhibits moderate resistance to African *Pgt* races in the Ug99 lineage and a high level of resistance against North American *Pgt* races (Table 1). These ten lines were crossed to the universal susceptible line 'LMPG-6' to develop recombinant inbred line populations (F_6 or more inbred) via the single seed descent method at the University of Minnesota. The number of inbred lines in the ten populations ranged from 55 to 110 (Table 1). The pedigree information, known genes based on marker screening, and stem rust reactions of each parent are provided in Table 1.

Disease Phenotyping

The ten RIL populations, comprised of 852 lines, were evaluated for adult plant reaction to stem rust in four environments: St. Paul, MN, USA during May-August 2012 & 2013 (referred as StP12 and StP13 in the text); South Africa during October 2012 – January 2013 (referred as SA12 in the text); and Njoro, Kenya during May-October 2013 (referred as Ken13 in the text).

In the Njoro nursery, lines were planted in an augmented design with 1 check, 'Red Bobs'. Each line was sown in double 70 cm long rows, 20 cm apart. On each side of the plot, and in the middle of the plots, a twin-row of susceptible spreader wheat cultivar 'Cacuke' was sown. The field was also surrounded by a border of several spreader rows comprised of susceptible wheat varieties which were artificially inoculated using a bulk inoculum of *Pgt* urediniospores collected at the Njoro field site; however wheat stem rust differential lines with known stem rust resistance genes indicated that the predominant, if not only, race present in the nursery since 2008 was race TTKST (avirulence/virulence formula on the wheat stem rust differential panel: *Sr36*, *SrTmp*/*Sr5*, *Sr6*, *Sr7b*, *Sr8a*, *Sr9a*, *Sr9b*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr10*, *Sr11*, *Sr17*, *Sr21*, *Sr24*, *Sr30*, *Sr31*, *Sr38*, *SrMcN*) (Njau et al. 2010).

In the St. Paul 2012 environment, lines were planted in 2 m long single rows with 20 cm between the rows. The populations were planted in an augmented design with 4 check varieties 'Oklee' (Anderson et al. 2005), 'Thatcher' (Hayes et al. 1936), 'Tom' (Anderson et al. 2012), and 'Verde' (Busch et al. 1996) planted after every 30 entries. In the St. Paul 2013 environment, lines were planted in hill-plots with 20 cm distance between the hills. The same checks were planted in the same manner as in the 2012 season. In both environments, the lines were surrounded by a mixture of susceptible lines 'Morocco', 'Thatcher', 'Max', and 'Little Club' planted perpendicular to the lines on all sides. To initiate disease, spreader rows were syringe-injected with a mixture of North American stem rust races MCCFC (isolate 59KS19), QFCSC (isolate 03ND76C), QTHJC (isolate 75ND717C), RCRSC (isolate 77ND82A), RKQQC (isolate 99KS76A),

and TPMKC (isolate 74MN1409) at the jointing stage. The spreader rows were sprayed with a bulked mixture of *Pgt* races suspended in a light mineral oil suspension using an Ulva+ sprayer (Micron Sprayers Ltd., Bromyard, UK) after heading stage.

In South Africa, populations were planted in Cedara of KwaZulu-Natal Province as an augmented design with the check lines 'Morocco' and 'Kariega' planted after every 50 entries. Seeds for each line were planted in hill plots with 30 cm between plots in a row and 60 cm between rows. To initiate the disease, spreader rows containing 'Morocco' and 'McNair' were inoculated with a mixture of the *Pgt* races TTKSF and PTKST of the Ug99 race group using an ultra low volume sprayer twice in the season – once during the booting stage, and again at flowering stage.

Additionally, all 11 parent lines were inoculated with race TTKSK at seedling stages to postulate the presence/absence of major genes providing resistance to this widely virulent *Pgt* race. Disease inoculation and phenotyping procedures were carried out as described by Rouse et al. (2012).

Statistical Analysis

Mixed linear models were fitted to each environment to assess family and line within family genotypic effects, and also the effects of heading date and growth stages of the lines. Factors explaining significant amounts of variation were retained in the model based on the likelihood ratio tests calculated in 'lme4' package 3.0.3 in R (R Development Core Team 2013, http://www.r-project.org/). Phenotypic data for each line was adjusted based on either the number of days to heading (St. Paul data) or growth stage (Kenya data) by estimating the effect of days to heading or growth stages on stem

rust severity for each line within each RIL population, and subtracting the estimate from rust severity for each line. The number of days to heading was measured as the day after planting when half the spikes in the plot fully emerged above the flag leaf. Growth stages of each line in Kenya were determined mainly by assessing grain development stages such as watery, milky, soft dough, and hard dough; and also for stages of booting and flowering, as explained by Zadoks et al. (1974). The phenotypic scale used in South Africa was different than the other locations as the disease was evaluated on a quantitative scale of 0 to 10 representing highly resistant (scores 0-2), resistant (score 3), moderate resistance (scores 4-5), moderate susceptible (scores 6-7), susceptible (score 8), and highly susceptible (scores 9-10) (Tsilo et al. 2014). Therefore, all data sets were fitted into a mixed model with environments as fixed effects and lines as random effects to correct for data distortion due to trial effects. Using this model, best linear unbiased predictors (BLUPs) for each line were predicted from the combined analysis model using SAS 9.1, from which final adjusted trait values were obtained.

Genotyping, SNP Discovery & Imputation of Missing Alleles

All lines used in the study were genotyped using the genotyping by sequencing (GBS) approach modified after Elshire et al. (2011). Genomic DNA was extracted from ground leaf tissue of $F_{6:7}$ RILs and parent seedlings using the BioSprint 96 DNA Plant Kit (QIAGEN, Valencia, CA). Extracted DNA was quantified using the picogreen assay and diluted to 20ng/µl. The restriction enzymes *Pst*I and *Msp*I were used to generate double-digested complexity-reduced DNA libraries following Poland et al. (2012b). The *Pst*I restriction overhang was preceded by barcoded forward adapters, whereas the *Msp*I

overhang was fitted with a common Y-adapter. This design prevented either *Pst*I-*Pst*I or *Msp*I-*Msp*I fragments from amplifying, thereby yielding sequences of only the *Pst*I-*Msp*I fragments, producing a uniform library. Few modifications were introduced into the protocol, mainly: 1) each sample was ligated to two unique barcodes to minimize sequencing bias of reads with certain barcodes; and 2) the concentrations of the barcode adapter and the common adapter were increased to 0.1 μM and 50 μM, respectively, in an attempt to capture most of the digested DNA fragments. Ten 96-plex libraries were generated, with each parent repeated at least six times to obtain higher read coverage of parental alleles. Each library was sequenced in one lane of Illumina HiSeq 2000, generating 100 bp single-end sequences.

The Universal Network Enabled Analysis Kit (UNEAK) was used to call *de novo* SNPs across the populations using the parameters -c 10 –e 0.025 (Lu et al. 2013). Chromosome locations of the SNPs were obtained by aligning the reads containing SNPs to the wheat chromosome survey sequences (CSS), using the reads as query sequences in a local 'blastn' search. The wheat CSS sequences were obtained by assembling reads obtained from sequencing flow-sorted wheat chromosomes from 'Chinese Spring' (International Wheat Genome Sequencing Consortium, http://wheaturgi.versailles.inra.fr/Seq-Repository/). To ensure that correct SNPs were obtained, the following filters were applied to the blast results: 1) each read was allowed only one alignment to a unique chromosomal location; 2) full length alignment of each read was required; and 3) either one base mismatch or one gap between the query sequence and the wheat CSS was permitted, assuming some level of sequence

polymorphism existed between Chinese Spring and lines in our study. Our stringency parameters are modeled after Wang et al. (2014) who used a similar approach to assign their 91,829 SNPs on the Infinium iSelect array to wheat chromosomes. The workflow of genotyping, map construction, data analysis, and genetic mapping in this study is presented in Figure 1.

One unique property of the GBS approach is the generation of significant proportions of missing genotype data (Williams et al. 2010; Poland et al. 2012b; Fu et al. 2014). As such, construction of linkage maps with missing data is an arduous task. However, the NAM design of mapping populations allows for imputation of parental SNPs to the segregating progeny with high accuracy (Guo and Beavis 2011). As both parents were sequenced multiple times, the high-confidence allele calls between the parents can be used to impute missing haplotypes in the progeny. Hence, imputation of missing SNP data based on family relationship was carried out using principal component analysis (PCA) based imputation using the probabilistic PCA (ppca) algorithm in the freely available R package 'pcaMethods' (Stacklies et al. 2007). The ppca algorithm first assigns row average values to the missing values, and then uses the singular value decomposition of the SNP matrix to create orthogonal principal components. In turn, the PC values corresponding to the largest eigenvalues are used to reconstruct the missing SNP genotypes in the data matrix. The algorithm 'ppca' was chosen for its high imputation accuracy and efficiency in regards to the use of computational resources compared to other imputation algorithms of similar caliber (Moser et al. 2009; Fu 2014).

Map Construction and Joint QTL Mapping

The genetic map was constructed using the program IciMapping 3.3 (Wang et al. 2012a) using a minimum logarithm of odds (LOD) value of 5.0. Genetic distances between the markers were calculated based on the Kosambi mapping function (Kosambi 1943). A search for QTL across all populations, and in each population was done using MCQTL 5.2.6 (Jourjon et al. 2005). Within MCQTL, the iterative QTL mapping (iQTLm) procedure (Charcosset et al. 2000) was implemented with walk distance of 5 cM to detect significant QTL. The iQTL method works in two steps where significant QTL are first scanned along chromosomes for their precise locations, followed by adding new QTL to or dropping QTL from the model based on Fisher's test. Permutation analysis of the trait data was implemented using a resampling method as described by Churchill and Doerge (1994). Trait threshold was computed as described by Jourjon et al. (2005) to adapt multiple cross design using 1,000 permutation steps. Significant QTL were declared if the LOD threshold level was greater than the calculated value at $\alpha = 0.05$. Epistasis effects among the detected QTL were also estimated, using the joint connected model in MCQTL 5.2.6, with LOD threshold calculated as described above. QTL mapping of each population was also carried out separately using the composite interval mapping (CIM) approach in MCQTL 5.2.6. For each detected QTL, the percent of phenotypic variance explained (R^2) and allelic effects were also estimated.

RESULTS AND DISCUSSION

Genotyping

We generated ten 96-plex GBS libraries representing 852 RILs from ten biparental populations. Each parent line was sequenced at least six times to obtain sequences with higher read depth for confidence in SNP-calling and imputation. As a result, the sequencing of each library on one lane of Illumina HiSeq 2000 generated a total of 1.5 billion 100 bp reads, with 154 million reads on average per lane. On average, 90% of the generated bases passed the Q30 filter with a median Q-score of 34.98. The reads were then filtered for having intact barcode sequences and a complete *Pst*I overhang, which led to 73% of total reads assigned to each individual. This figure is comparable to recent GBS studies in wheat (Saintenac et al. 2013a) and barley (Liu et al. 2014) that follow the same library construction and sequencing protocols. This resulted in a read distribution per individual from 175,443 to 31,381,071 with a median of 1,225,681 reads per RIL. A total of 158,182,462 reads were obtained for the 11 parents used in the study, with a median value of 13,348,322 reads, and ranging from 7,951,649 (Kudu) to 31,381,071 of the common parent LMPG-6 which was replicated 10 times in the library (Appendix IV).

The GBS approach has been used in species where reference sequences are available (Elshire et al. 2011; Liu et al. 2014), and in species where reference sequences are not available (Baxter et al. 2011; Saintenac et al. 2013a). Owing to the power in data analysis from the generated sequences, this approach has been successfully used in several genomic studies in non-model species as well as in species with high genome

complexity, such as that of hexaploid bread wheat. The large number of reads generated, as evident by the numbers reported here, enable the discovery of abundant polymorphic markers to be used in studies ranging from genetic mapping to population genomics. Regardless of the approach used to analyze the sequences and research questions asked, this method is efficient, generates better value to cost ratio than most other available genotyping systems, and therefore, is of high usability in investigating marker-trait associations (Chapter 3 of this thesis).

SNP Discovery and Linkage Mapping

SNPs for which the SNP-containing reads matched uniquely to the WCSS were retained and passed through imputation steps. The imputed SNPs were compared against all 10 populations to remove inter-population common SNPs and retain only unique polymorphic SNPs. This led to identification of 992 SNPs that were used for construction of linkage maps and QTL mapping. Of the 992 high quality SNPs, 930 SNPs were assigned to 21 linkage groups with all 21 wheat chromosome represented. The number of markers per linkage group ranged from 1 (Chromosomes 3D, 7D) to 288 (Chromosome 6A). The total genetic distance covered by these groups was 4,977 cM, with one SNP marker placed at every 5 cM on average. The distribution of SNPs by linkage groups is shown in Figure 2. Linkage groups representing chromosomes 3B and 5B, were larger compared to other linkage groups, at 604 cM, and 696 cM, respectively. We suspect that the large amount of missing data in the genotype matrices representing these chromosomes led to this size inflation. The amounts of missing data on 3B and 5B genotype matrices were 18% and 39%, respectively. The remaining SNPs mapped to the

other 19 chromosomes together had less than 6% missing data. We decided to retain these SNPs as their removal would have resulted in insufficient number of SNPs for linkage group construction.

Figure 2 also shows the read coverage per SNP, arranged by their positions on each linkage group. The median number of read coverage per SNP was 346, and ranged from 105 to 7,943. This range in read coverage per SNP is likely an attribute of uneven sampling of the genome during sequencing, either from non uniform representation of the samples, or from a large number of reads being discarded during our stringent sequence filtering and SNP calling procedures.

Population Characteristics and Stem Rust Reaction

Population structure among the populations was determined by singular value decomposition of the genotype data using the 'princomp' package in R 3.0.3 (R Development Core Team 2013, http://www.r-project.org/). No significant structuring among the lines was observed (Appendix V), which is expected of NAM population design (Yu et al. 2008). The lack of population structure among the populations despite having diverse founder lines is an attractive aspect of the NAM design. As different founder lines are crossed to a common parent, shuffling of the parental genomes in the progeny normalizes allelic differences by virtue of the common parent. Therefore, population stratification is effectively controlled in the NAM design, which minimizes spurious associations that could arise from population structure. Because of the lack of structure, relationship matrices were not used as cofactors during the data analysis.

We phenotyped all ten RIL populations and the parent lines over two years at four environments where disease epidemics were artificially established by inoculating spreader lines with local *Pgt* races. Screening for APR to North American and African *Pgt* races showed that all male parents used in the cross were resistant to moderately resistant (Table 1). Some susceptible pustules were observed on the stem of 'Ada' in Kenya 2013 season, albeit at low severity (average of 11%). The susceptible parent 'LMPG-6' succumbed to high disease severity in all four environments and exhibited high susceptibility. Phenotyping of all populations was initiated once disease severity on 'LMPG-6' was at its maximum at each environment. All populations exhibited continuous disease distribution in most trials, allowing mapping of quantitative loci associated with disease resistance (Appendix V). The frequency of resistant lines was higher in the Kenya 2013 environment than any other, most likely due to the unusually cooler day and night temperatures at this site during the 2013 main season, which led to slower onset of disease than usual. Adjustment of phenotypic data was able to resolve this, as trait means among all environments were similar post-adjustment.

Seedling screening of the parent lines with the race TTKSK showed that all parents were susceptible to this race (Table 1). The parent lines were screened for the presence of resistance genes effective against races of the Ug99 lineage for which diagnostic DNA markers have been developed (MAS Wheat 2014). While the results indicated that the lines 'Gem', and 'Romany' contain the gene *Sr22* which is effective to all seven races of the Ug99 lineage (Singh et al. 2011), all lines were found to be susceptible to the Ug99 lineal race TTKSK during seedling screening, suggesting that

these lines lack any seedling gene (Table 1). The discrepancy between detection of *Sr22* and the susceptibility of these two lines to race TTKSK has been discussed in detail in the 'Comparison with Previously Reported Genes and QTL Conferring Resistance to *Pgt*' section.

Stem Rust QTL Mapping

We implemented the iterative QTL mapping (iQTLm) method to map QTL in all RIL populations providing APR to stem rust in four environments. Results of the joint QTL analysis of the NAM panel are presented in Figure 3 and Appendix IX99. We also mapped QTL in individual populations using the composite interval mapping (CIM) method, and the results from this analysis are summarized in Figure 4 and Appendix IV.

The iQTLm approach discovered 27 QTL on 11 of 21 wheat chromosomes (Figure 3). This number of QTL detected in joint mapping is larger than that detected in any single-population QTL detection using the CIM approach (Appendix IV), although the QTL effects and R^2 values are comparable between the two approaches. Our findings are in agreement with published NAM studies where additional QTL have been discovered in joint mapping compared to single-populations (Buckler et al. 2009; Negeri et al. 2011). Six of the 27 QTL were represented multiple times in more than one environment. The QTL *QSr.umn-2A.1* and *QSr.umn-4B.1* were common between Ken13 and Stp13 environments; *QSr.umn-4A.1* between Ken13 and StP12; and *QSr.umn-2B.2* (no relationship with *QSr.umn-2B.2* in Chapter 2), *QSr.umn-3B.4*, and *QSr.umn-3B.7* between StP12 and StP13 environments (Table 3, Appendix VI).

The 27 QTL detected by the NAM approach explained a wide range of phenotypic variation (R^2) for adult plant resistance to stem rust of wheat, with the lowest R^2 of 0.6% and the highest of 5.0%. These values indicate the absence of any large-effect QTL/gene contributing towards disease resistance. Rather, several QTL seem to contribute towards resistance in a cumulative fashion, as confirmed by the relatively small additive values in Figure 3. However, small population size can inflate phenotypic estimation variation (Eucharia 1990). As some populations in our study are small, the R^2 values and additive effect estimates will have greater error vs. populations of larger size. Further phenotyping with larger populations is required to more accurately assess the R^2 value contributed by the detected QTL. Interestingly, the additive effects of the QTL were similar to those that have been reported in stem rust QTL mapping studies carried out in biparental populations (Bhavani et al. 2011; Macharia 2013; Singh et al. 2013d). For each QTL, the parental alleles contributed different effects towards stem rust resistance, as exhibited by the distribution of total additive phenotypic effects of each parent, as shown in Figure 3. Overall, each parent contributed towards resistance in each environment. The positive values associated with the parent lines for specific SNPs in certain environments do not imply that the parents are not important sources of resistance. As the effects of all parental alleles are considered simultaneously during joint mapping, variations in disease resistance in each family, contributed by each parent, are expected to occur. Moreover, the choice of a parent for introgression of alleles for APR is also the function of several important factors defining the target environment such as frequency, and types of *Pgt* races, and disease pressure. While the contribution to disease

resistance from alleles with low additive values might be difficult to visually observe in the field, APR genes acting in an additive manner should elevate the resistance, thereby assisting in phenotypic selection.

We did not detect epistasis within the populations, or in a joint setting with all populations combined. Recent studies have mixed results for detection of epistatic interactions between the mapped QTL as some studies report significant levels of interaction (Yu et al. 2012; Singh et al. 2013a; Singh et al. 2013b; Rouse et al. 2014) whereas other studies do not (Bhavani et al. 2011; Macharia 2013; Singh et al. 2013d). Taken together the number of QTL detected during our analysis and magnitude of the marker effects, it could be said that the nature of resistance to stem rust in these ten populations is polygenic, with several loci of minor to major effect acting in additive fashion. A point to consider is that the lack of epistasis should not deter a breeder from adoption of a QTL as long as it contributes to elevated APR, recognized mainly from its additive value.

It should also be noted that the absence of a QTL effect in some environments does not necessarily mean that the locus had no effect on disease resistance. It is possible that the stringent threshold settings of the analysis filtered the QTL out from being detected in certain chromosomal regions, especially if they were slightly under the threshold value. This is likely to happen especially due to inadequate marker coverage in such regions. Another possible explanation is because environmental differences alter the expression of the genes, causing the loci to go undetected in some environments during analysis. The breeding history of the founder parents, in part, could also be responsible

for the observed differences in QTL x environment effects. As the varieties are selected for optimal performance in different environments, it is possible that certain loci are fixed while others are purged, producing a wide range of results during genome mapping.

In our study, the distance of a SNP nearest to the reported QTL peak ranged from 0 cM to 20 cM, with an average value of 4 cM. Only four QTL were detected with SNPs further than 5 cM from the QTL peak position. The average extent of intra-chromosomal LD in spring wheat population is reported to be 20.8 cM (median value of 11.5 cM) and the rate of LD decay is about 6 cM (Chao et al. 2010). Hence, as most of the discovered significant markers close to the QTL peak, they can be expected to remain in LD with the causative loci and can provide high value in marker assisted selection for resistance breeding.

Comparison of Joint Mapping to Single-population QTL Mapping

We were interested in knowing if the QTL detected during joint mapping could also be detected by conducting QTL mapping of each population separately. The CIM approach of QTL detection conducted separately on each RIL population found a total of 56 QTL on 17 chromosomes. No QTL were detected on chromosomes 1D, 3D and 4D, 5A. The population LMPG-6/Fahari had the largest number of QTL (15 QTL), and the populations LMPG-6/Kulungu and LMPG-6/Paka had the least number of QTL (1 QTL each) (Figure 4). No QTL were detected in LMPG-6/Ngiri population. No QTL were detected in all four environments in any individual RIL population, although seven populations had at least one QTL effective to *Pgt* races in three of the four environments.

The joint mapping approach detected QTL in all environments and on all chromosomes where QTL were detected by single-population QTL mapping. The single-population mapping (upon summation of number QTL detected in all populations) and joint mapping methods detected the same number of QTL (11) in the Kenya 2013 environment. In all other environments, this number was higher in single-population mapping with 9 QTL detected in SA12, 14 QTL in StP12, and 21 QTL in StP13 relative to 2, 11, and 9 QTL detected in SA12, StP12, and StP13 environments, respectively, in joint mapping (Figure 4, Appendix IV). The percent of phenotypic variation (R^2) explained by the QTL as well as the additive effects contributed by the parents between the two methods were different. The R^2 value of a QTL detected in single-population mapping ranged from 0.1% to 24.6% whereas it ranged from 0.6% to 5% in joint mapping. The additive values of the parents calculated for each QTL in the joint mapping approach ranged from -33 to +19 whereas it ranged from -4 to +8 for QTL detected in individual populations. As QTL in joint mapping are declared based on their overall contribution to resistance (or susceptibility) across all 10 families, a 'normalized' value is calculated with the effect of each parent taken into account. In other words, a joint analysis estimates the allelic effects of a QTL by assuming that the allele from a founder parent is uniform across all populations (Rebai and Goffinet 1993; Blanc et al. 2006). This might not necessarily be true for each QTL detected, especially in case of linked genes, and among populations that differ in the number and effect of loci associated with the trait. In our study, via CIM mapping of single populations, we have found that not all populations share the same locus for stem rust resistance (Appendix IV). Hence, such an

assumption in the joint mapping model could be one of the main reasons why the joint mapping differs from individual mapping in estimation of the amount of total phenotypic variance and the allelic effect accounted by the significant SNPs.

The two methods also detected a few QTL that were consistent in several disease environments. Eleven QTL distributed on eight chromosomes that were detected by the joint mapping approach were also detected by CIM in single populations (Table 3). The Pearson correlation value for the percent of phenotypic variance explained (R^2) by the QTL detected between the two methods was moderately high at 0.55, implying the similarity between the QTL detected by the two methods. The biggest differences in the R^2 value were observed in four QTL detected between the SA12 and StP12 environments where the values differed from 11.2% to 20.5%. Differences due to environmental conditions, inoculum load, and *Pgt* races present in each environment are probable to contribute to the differences in QTL detected in different environments. As shown in Table 3, three of the 11 QTL were detected in different environments between the two methods: the QTL detected on chromosome 3B in Ken13 and StP13 environments by joint mapping were detected in environments StP13 and Kenya 2013, respectively, in single-population mapping; and the QTL detected on 5D in the StP12 environment by joint mapping was detected in SA12 environment in single-population mapping. The difference of the environments where the QTL were detected suggests that these loci might be involved in providing resistance that is almost specific to the disease environments, and perhaps to the *Pgt* races. Validation of the QTL detected in both

approaches using the markers significantly associated with the QTL will help to further understand the relationship between the QTL detected in these two mapping approaches.

Comparison of Joint Mapping Results to Previously Reported Genes and QTL Conferring Resistance to *Pgt*

Previous mapping studies have reported QTL conferring resistance to the *Pgt* races on the same chromosomes as detected in the present study. However, a joint mapping study with such objectives has not been carried out before. In our study, we detected one QTL on chromosome 2D, two QTL on 2A, and three QTL on 2B, of which the QTL *QSr.umn-2A.1* was common between Ken13 and StP13 environments and the QTL *QSr.umn-2B.2* was observed in both St. Paul environments (Figure 3). Several QTL conferring APR to *Pgt* races, including the Ug99 lineage races, on chromosomes 2A, 2B, and 2D have been detected in several spring wheat RIL populations and genome-wide association study (GWAS) panels as well as in durum wheat populations and GWAS panels (see Haile and Röder 2013; Yu et al. 2014). Several Sr genes are also located in the chromosomal regions where QTL for stem rust resistance were detected in this study. The genes *Sr32* (located on 2A and/or 2B), and *Sr28*, *Sr36*, *Sr39*, *Sr40*, *Sr47* (located on 2B) are effective to the races in the Ug99 lineage, but are unlikely to exist in our population based on screening for seedling resistance to TTKSK and screening of the parents using molecular markers. Several other genes such as, *Sr6* (located on 2D), *Sr9a*, *9b*, *9d*, *9e* (located on 2BL); *Sr21* (located on 2AL), and *Sr38* (located on 2AS) are ineffective against African races but are resistant to one or more North American *Pgt*

races (Rouse and Jin 2011; Zhang et al. 2014). Also located on the long arm of 2B, *QSr.umn-2B.2* was detected in both StP environments, and QSr.umn-2B.3 was detected in StP12 environment. Kenyan lines are known to be the sources of many Sr genes discovered to date (McIntosh et al. 1995). Examples include *Sr6*, which is quite common in Kenyan lines (Knott 1962), and *Sr9b*, which was first observed in Kenyan lines (Knott and Anderson 1956). The line 'Frontana', present in the pedigree of the founder parent 'Gem', is known as one of the sources of *Sr9b* (McIntosh et al. 1995). Hence, our NAM population may possess these Sr genes, in addition to previously unidentified genes, and could be involved in providing resistance to North American *Pgt* races.

The QTL *QSr.umn-3B.4* detected on distal end of chromosome 3B may indicate the presence of *Sr2*, which is segregating in four populations: LMPG-6/Gem, LMPG-6/Kudu, LMPG-6/Paka, and LMPG-6/Romany. The presence of *Sr2* in 'Gem', 'Kudu', 'Paka', and 'Romany' was confirmed by marker screening (Table 1) as well as expression of the trait pseudo-black chaff (PBC) in St. Paul 2012 and Kenya 2013 environments on plant internodes (not observed in Romany). PBC is conditioned by the expression of the partially dominant gene *Pbc* causing dark coloring of the glumes and inter-nodal regions in an adult wheat plant, and is considered to be associated with the *Sr2* gene (McFadden 1939; Sharp et al. 2001). The expression of PBC is highly dependent on the environment for its expression, and is also affected by the genetic background. The additive effect of the resistance allele at this QTL was observed to contribute towards resistance in populations (Figure 3) with the founder parents 'Kudu', 'Paka', and 'Romany', whereas the parent 'Gem' was found to contain an allele that

contributed towards susceptibility. *Sr2* is an APR gene that provides broad-spectrum resistance to all known isolates of *Pgt* worldwide (McIntosh et al. 1995; Spielmeyer et al. 2003). Therefore, it is expected to be effective in all environments. As *QSr.umn-3B.4* was observed in only two environments (StP12 and StP13) in our study, further field screening of the population is needed to confirm its identity relative to *Sr2*.

We detected three QTL on the proximal end of 4B, which explained 0.6% to 4.0% of phenotypic variance observed in the four environments. Chromosome 4A contains the genes *Sr7a*, and *7b*, and chromosome 4B hosts the genes *Sr37*. *Sr37* is resistant to Ug99, and therefore is unlikely to be present in our population as the parents were seedling-susceptible to Ug99. Both *Sr7a* and *Sr7b* are resistant to multiple N. American *Pgt* races (Roelfs and McVey 1979; Rouse and Jin 2011). *Sr7a* was first identified in several Kenyan wheat lines, (Knott and Anderson 1956; Knott 1962); and 'Ngiri' could have acquired this gene from the line 'Manitou' which is present in its pedigree (McIntosh et al. 1995). Thus, this gene may have been bred into one or more of the Kenyan varieties used to create our NAM population. Further screening of the population, preferably in single rust race nurseries, is needed to confirm this proposition.

We detected three different QTL on 5B in Ken13, StP12 and StP13 environments that explained 1.6%, 1.8%, and 2.1%, respectively, of the variance observed for rust resistance in these environments. Two of these three QTL were also detected in single-population mapping (Table 3). The QTL *QSr.umn-5B.2* detected in StP12 environment was also detected in the same environment in LMPG-6/Popo population; and the QTL *QSr.umn-5B.3* detected in StP13 environment was detected in the same environment in

the LMPG-6/Fahari population. No known Sr gene has been mapped to 5B, and the previously reported QTL effective to Pgt differ from those identified in our study in terms of their position and allelic effects.

The QTL on 5D (*QSr.umn-5D.2*) detected in this study may be *Sr30*, which is known to exhibit resistance to North American *Pgt* races. *Sr30* is ineffective against African races such as TTKSK at both seedling and adult plant stages (Mago et al. 2011a; Yu et al. 2014), yet exhibits seedling resistance to a few North American *Pgt* races (Rouse and Jin 2011) and all-stage resistance to Australian *Pgt* pathotypes (Bariana et al. 2001; Kaur et al. 2009). The total phenotypic variance explained by *Sr30* was high (≥ 20%) in the study by Kaur et al. (2009) relative to that explained by *QSr.umn-5D.2* (2%). Differences among the *Pgt* races used to initiate the disease, environments used for APR screening, and the genetic background of the parental lines could be the probable factors causing this variation.

The QTL *QSr.umn-6A*, detected only in the Ken13 environment, explained 2.6% of phenotypic variance in this environment. QTL effective against races in the Ug99 lineage have been detected on 6A previously, including *Sr26* in common wheat (Prins et al. 2011a; Yu et al. 2011); and *Sr13* in durum wheat (*Triticum durum* Desf.) (Pozniak et al. 2008; Letta et al. 2013). *Sr26* is effective against all races of the Ug99 lineage (Mago et al. 2011a), whereas all the NAM parents were susceptible to race TTKSK (Table 1). Also, as *Sr13* is not prevalent in common wheat (McIntosh et al. 1995; Yu et al. 2014), it is unlikely that either of these genes is present in our population. Chromosome 6A also contains the gene *Sr52*, which was recently introgressed into hexaploid wheat from its

diploid relative *Dasypyrum villosum*, and is resistant to TTKSK (Qi et al. 2011). Being a novel gene introduced to wheat via translocation, *Sr52* is not currently used in breeding programs (Yu et al. 2014), and therefore is not expected to be present in old Kenyan varieties used in this study. It is possible that *QSr.umn-6A* is a novel source of resistance to the Ug99 group of races. Further tests are needed to confirm this hypothesis.

Pozniak et al. (2008), in their association mapping study comprised of durum wheat lines, report a QTL providing APR to the Ug99 lineage races located at 83 cM on 7A. The gene *Sr22* is also located on 7A. This gene was detected in lines 'Gem' and 'Romany' during our marker screening (Table 1). While this may suggest the presence of *Sr22* in our population, it should be noted that *Sr22* is effective to TTKSK at the seedling stage (Mago et al. 2011a), which we did not observe in either 'Gem' or 'Romany' during our tests for seedling resistance. We suspect that either the markers we used amplified a null allele, or that the markers are not completely diagnostic, as confirmed by Olson et al. (2010a), whose recommended markers were used for screening of the gene.

The QTL *QSr.umn-7B* detected in the SA12 environment explained 4.1% of the observed phenotypic variance, and is likely novel. The only gene mapped to 7B is *Sr17*, which is ineffective to races of the Ug99 lineage as well as to the North American *Pgt* races (Mago et al. 2011a; Rouse and Jin 2011). Previously reported QTL on chromosome 7B that confer resistance to the Ug99 group of races include the QTL linked to the microsatellite markers *wPt-0318* (Yu et al. 2012) and *cfa-2040* (Pozniak et al. 2008). Additional research, either jointly or in individual populations, is necessary to further

understand the relationship between discussed All-stage genes with the QTL detected in our study.

CONCLUSION

In this study, we present results of a nested association mapping approach, an effective mapping strategy that utilizes the mapping power of several RIL populations. In addition, by using a statistical procedure designed for NAM, we estimate allele effects by using the common parent as a common genomic reference. Joint QTL mapping for stem rust resistance resulted in detection of a large number of QTL with small to large effects. Epistatic effects among the detected loci did not have significant contributions to stem rust resistance, suggesting that the differences between the populations in different environments are largely due to additive effects of several QTL. Also, by sequencing complexity-reduced genomes of the whole NAM panel, we obtained population-specific *de novo* SNPs, and show their usability in mapping QTL effective against stem rust of wheat. While validation of the detected loci is required to confirm the significant markers, GBS offers an efficient and economical approach for genome mapping and genomewide studies. However, multiple rounds of sequencing is recommended to obtain enough reads to minimize the complications that could arise from having substantial amounts of missing data. Sequencing the parent lines at higher read depth helps towards more accurate imputation of missing haplotypes. The genomic regions associated with stem rust resistance identified in this study will be useful to breeders in introgression of resistance to stem rust of wheat, including the widely virulent African stem rust races. Validation of our *de novo* markers by fine mapping of the regions and/or in marker assisted breeding is needed to confirm the locations of the reported QTL.

Table 1: Origin, pedigree, and stem rust reaction of parent lines used to develop the NAM population.

Parent	Origin[a]	Pedigree[a]	Sr markers[b]	TTKSK[c]	Field reaction[d]		RILs[e]
					USA	Africa	
Ada	USA (2007)	SBY189H/'2375'		3+	22.3	11.3	71
Fahari	Kenya (1977)	TOBARI-66/3/SRPC-527-67//CI-8154/2*FROCOR		33+	11.7	3.7	90
Gem	Kenya (1964)	BT908/FRONTANA//CAJEME 54	*Sr2*, *Sr22*	3-	10.2	1.0	97
Kudu	Kenya (1966)	KENYA-131/KENYA-184-P	*Sr2*	3+	19.8	3.7	80
Kulungu	Kenya (1982)	ON/TR207/3/CNO//SN64/4/KTM		33+	22.5	0.3	59
Ngiri	Kenya (1979)	SANTACATALINA/3/MANITOU/4/2*TOBARI-66		33+	10.2	2.3	52
Paka	Kenya (1975)	WISCONSIN-245/II-50-17//CI-8154/2*TOBARI-66	*Sr2*	3+	16.5	3.7	104
Pasa	Kenya (1989)	BUCK BUCK/CHAT		3+	8.6	5.3	93
Popo	Kenya (1982)	KLEIN-ATLAS/TOBARI-66//CENTRIFEN/3/BLUEBIRD/4/KENYA-FAHARI		3+	5.0	2.3	97
Romany	Kenya (1966)	COLOTANA 261-51/YAKTANA 54A	*Sr2*, *Sr22*	3+	11.1	5.3	109
LMPG-6	Canada (1990)	LITTLE-CLUB//PRELUDE*8/MARQUIS/3/GABO		4	64.9	25.0	-

[a] Information obtained from Njau et al. (2009), Macharia (2013), Anderson et al (2007), and Knott (1990). Year (in parenthesis) indicates the year the line was released or published

[b] Putative genes present in the NAM parents based on marker screening. Marker screening was conducted using diagnostic markers for genes resistant to Ug99 races, as catalogued in MAS Wheat (2014) http://maswheat.ucdavis.edu/. The lines 'Gem' and 'Romany' appear positive for the presence of *Sr22* but are susceptible to Ug99, implying that the markers used for screening are not diagnostic of the gene (see 'Results & Discussion' section for detailed explanation)

[c] Seedling screening of the parent lines with race TTKSK (isolate '04KEN156/04')

[d] For USA environment, the rust response of parent lines were averaged from StP12 and StP13 environments. For Africa environment, only Ken13 data is shown. Mean severity reactions (%) are shown for both sites

[e] The number of progenies in the population obtained by crossing the parent line to LMPG-6

Table 2: Common quantitative trait loci (QTL) detected between the iterative QTL mapping (iQTLm) in nested association mapping of 10 RIL populations, and composite interval mapping (CIM) methods in individual populations for stem rust adult plant resistance in four environments.

Env	Chr[a]	Left SNP	Right SNP	Pos[b]	LOD[c]	PVE[d] (%)
Ken13	3B	SNP4	SNP813	723.8	2.9	1.9
	4A	SNP371	SNP375	164.3	2.5	1.5
	6A	SNP935	SNP425	210.9	4.2	2.6
SA12	4B	SNP177	SNP203	358.8	7.3	4.0
	7B	SNP268	SNP337	43.0	7.6	4.1
StP12	2B	SNP294	SNP970	439.9	6.1	3.8
	3B	SNP320	SNP359	811.0	8.3	5.0
	4A	SNP371	SNP375	164.9	2.5	1.5
	5B	SNP142	SNP784	624.2	2.8	1.8
	5D	SNP714	SNP481	138.0	3.0	1.9
StP13	2B	SNP294	SNP970	434.9	3.2	2.3
	3B	SNP885	SNP790	684.5	3.6	2.6
	3B	SNP320	SNP359	811.0	3.5	2.6
	5B	SNP141	SNP765	575.9	2.9	2.1

Env	Chr[a]	Left SNP	Right SNP	Pos[b]	LOD[c]	PVE[d] (%)	Population
Ken13	3B	SNP885	SNP790	684.5	3.5	2.7	LMPG-6/Kudu
	4A	SNP371	SNP375	164.3	3.3	1.2	LMPG-6/Kudu
	6A	SNP935	SNP425	210.3	2.6	10.2	LMPG-6/Romany
SA12	4B	SNP177	SNP203	358.8	4.7	17.4	LMPG-6/Pasa
	5D	SNP714	SNP481	143.0	3.7	15.6	LMPG-6/Gem
	7B	SNP268	SNP337	43.0	7.6	24.6	LMPG-6/Romany
StP12	2B	SNP294	SNP970	439.9	3.0	5.7	LMPG-6/Gem
	5B	SNP142	SNP784	629.2	3.1	12.9	LMPG-6/Popo
StP13	2B	SNP294	SNP970	429.9	2.7	12.0	LMPG-6/Pasa
	2B	SNP294	SNP970	434.9	3.0	6.5	LMPG-6/Gem
	3B	SNP4	SNP813	713.8	3.7	3.4	LMPG-6/Ada
	3B	SNP320	SNP359	806.0	2.5	1.7	LMPG-6/Gem
	5B	SNP141	SNP765	575.9	2.6	1.2	LMPG-6/Gem

[a] Chromosome location of the QTL

[b] Position (centiMorgan) of the detected QTL peak in Chromosome 'Chrom'

[c] Logarithm of odds scores for the QTL detected at position 'Pos', based on joint mapping

[d] Percentage of phenotypic variation explained by the observed QTL, based on joint mapping

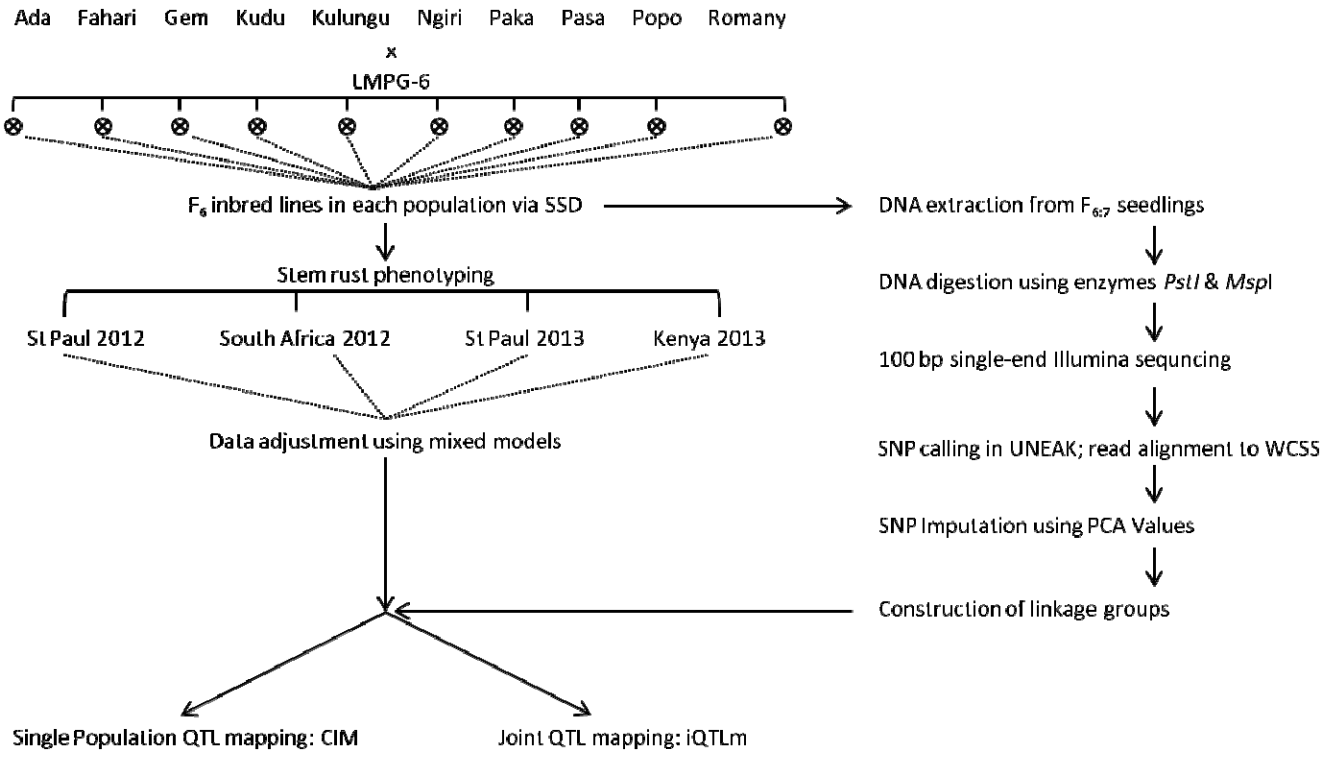


Figure 1: Schematic workflow of the study. See Materials & Methods for detailed explanation of each step.

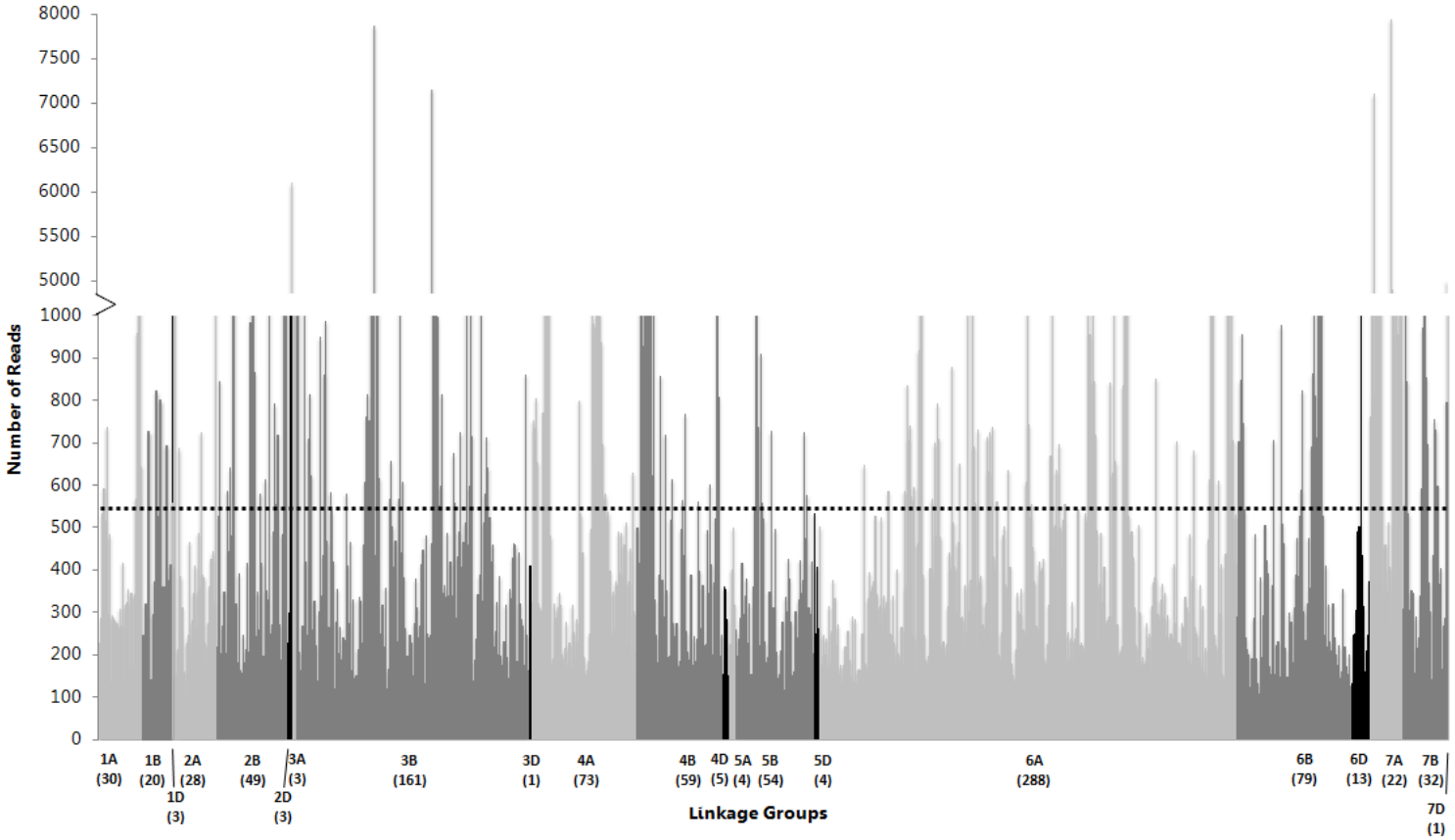


Figure 2: Distribution of SNPs by linkage groups (X-axis) and read depth (Y-axis) obtained from GBS approach. In each linkage group, SNPs are sorted by position. The numbers of SNPs in each linkage group are shown in parenthesis next to the linkage group. The dotted line represents the average SNP read coverage (546 reads).

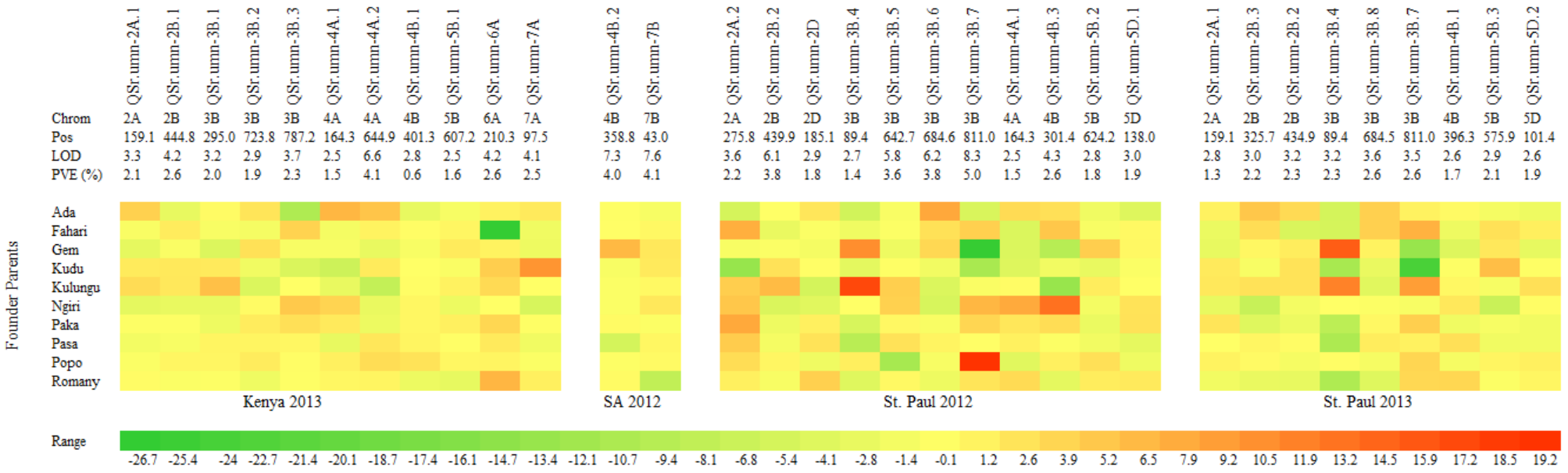


Figure 3: Heat map of additive effect estimates of alleles contributed by the 10 founder lines at the QTL for resistance to *Pgt* races. QTL (columns) are named according to McIntosh et al. (2003); the allelic effect estimates for each founder allele (rows) are color coded by increments in the allelic effect estimate (legend); each block represents the environments where the QTL were detected, as labeled. Chrom = Chromosome location of the QTL; Pos = Position (centiMorgan) of the detected QTL peak in Chromosome 'Chrom'; LOD = Logarithm of odds scores for the QTL detected at position 'Pos', based on joint mapping; PVE = Percentage of phenotypic variation explained by the observed QTL, based on joint mapping.

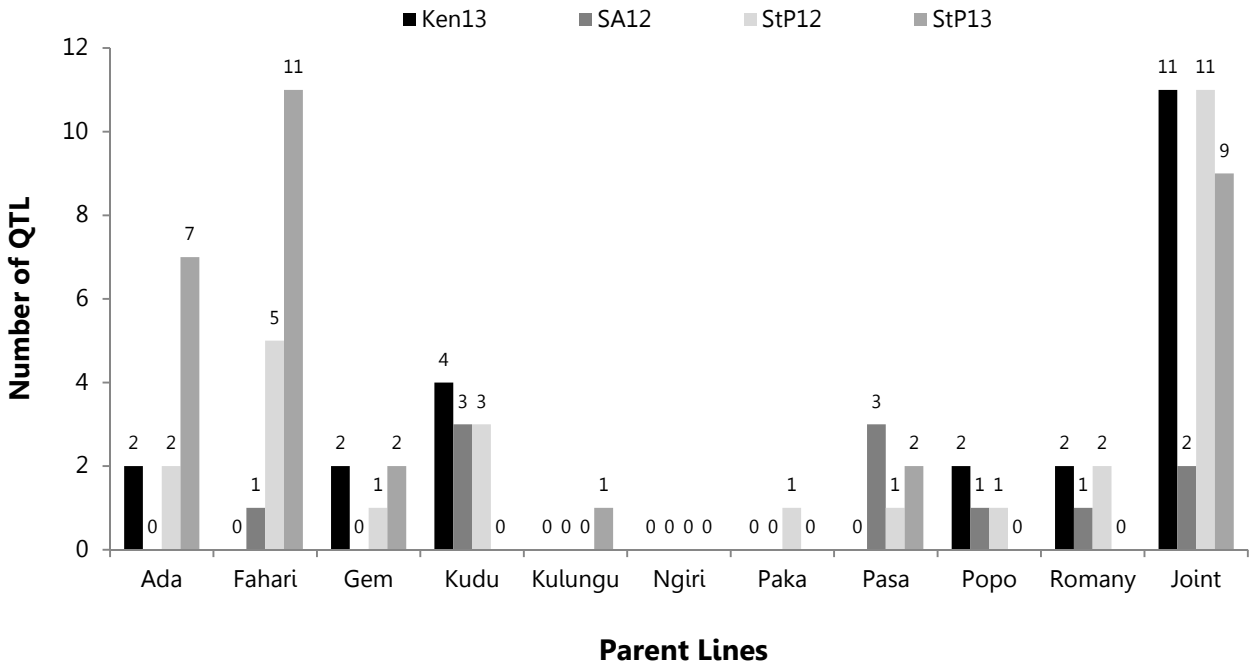


Figure 4: Frequency distribution of number of QTL discovered in different environments from composite interval mapping (CIM) of each RIL population and iterative QTL mapping (iQTLm) in joint mapping of all populations combined. Numbers next to the bars represent the number of QTL detected in all environments respective to the parent line. The environments Saint Paul 2012, South Africa 2012, Saint Paul 2013, and Kenya 2013 are abbreviated as StP12, SA12, StP13, and Ken13, respectively.

REFERENCES

Albrechtsen A, Nielsen FC, Nielsen R (2010) Ascertainment biases in SNP chips affect measures of population divergence. Molecular Biology and Evolution 27:2534-2547

Allen AM, Barker GLA, Berry ST, Coghill JA, Gwilliam R, Kirby S, Robinson P, Brenchley RC, D'Amore R, McKenzie N, Waite D, Hall A, Bevan M, Hall N, Edwards KJ (2011) Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). Plant Biotechnology Journal 9:1086-1099

Anderson JA, Busch RH, McVey DV, Kolmer JA, Jin Y, Linkert GL, Wieserma JV, Dill-Macky R, Wieserma JJ, Hareland GA (2007) Registration of 'Ada' wheat. Crop Science 47:434-435

Anderson JA, Busch RH, Mcvey DV, Kolmer JA, Linkert GL, Wiersma JV, Dill-Macky R, Wiersma JJ, Hareland GA (2005) Registration of 'Oklee' wheat. Crop Science 45:784-785

Anderson JA, Linkert GL, Busch RH, Wiersma JJ, Kolmer JA, Jin Y, Dill-Macky R, Wiersma JV, Hareland GA, McVey DV (2009) Registration of 'RB07' wheat. J Plant Reg 3:175-180

Anderson JA, Ogihara Y, Sorrells ME, Tanksley SD (1992) Development of a chromosomal arm map for wheat based on RFLP markers. Theoretical and Applied Genetics 83:1035-1043

Anderson JA, Wiersma JJ, Linkert GL, Kolmer JA, Jin Y, Dill-Macky R, Wiersma JV, Hareland GA, Busch RH (2012) Registration of 'Tom' wheat. Journal of Plant Registrations 6:180-185

Arthur JC, Kern FD, Orton CR, Fromme FD, Jackson HS, Mains EB, Bisby GR (1929) The Plant Rusts (Uredinales). J. Wiley & sons, Inc, New York

Ayliffe M, Singh R, Lagudah E (2008) Durable resistance to wheat stem rust needed. Current Opinion in Plant Biology 11:187-192

Bansal U, Bariana H, Wong D, Randhawa M, Wicker T, Hayden M, Keller B (2014) Molecular mapping of an adult plant stem rust resistance gene *Sr56* in winter wheat cultivar Arina. Theor Appl Genet 127:1441-1448

Bansal UK, Bossolini E, Miah H, Keller B, Park RF, Bariana HS (2008) Genetic mapping of seedling and adult plant stem rust resistance in two European winter wheat cultivars. Euphytica 164:821-828

Bariana HS, Cupitt CF, Warburton T (2001) Diversity of resistance to rust diseases in Australian wheats in 1999 and 2000 crop seasons. 10th Assembly of Wheat Breeding Society of Australia, Mildura, Australia, pp 233-236

Bariana HS, Kailasapillai S, Brown GN, Sharp PJ (1998) Marker assisted identification of *Sr2* in The National Cereal Rust Control Program in Australia. In: Slinkard AE (ed) Proceedings of the 8th International Wheat Genetics Symposium. University Extension Press, University of Saskatchewan, Saskatoon, Canada, pp 78-80

Bariana HS, McIntosh RA (1993) Cytogenetic studies in wheat. XV. Location of rust resistance genes in VPM1 and their genetic linkage with other disease resistance genes in chromosome 2A. Genome 36:476-482

Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD, Blaxter ML (2011) Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. PLoS ONE 6:e19315

Bhavani S, Singh RP, Argillier O, Huerta-Espino J, Singh S, Njau P, Brun S, Lacam S, Desmouceaux N (2011) Mapping durable adult plant stem rust resistance to the race Ug99 group in six CIMMYT wheats. In: McIntosh RA (ed) BGRI 2011 Technical Workshop, Borlaug Global Rust Initiative, St Paul, MN, USA, pp 43-53

Bidwell PW, Falconer JI (1925) History of agriculture in the northern United States, 1620-1880. Carnegie Institute, Washington DC, USA

Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor Appl Genet 113:206-224

Bonnett DG, Rebetzke GJ, Spielmeyer W (2005) Strategies for efficient implementation of molecular markers in wheat breeding. Mol Breeding 15:75-85

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American journal of human genetics 32:314-331

Browning JA (1979) Genetic protective mechanisms of plant-pathogen populations: their coevolution and use in breeding for resistance. In: Harris MK (ed) Biology and Breeding for Resistance. Texas A&M University, College Station, TX, USA, pp 52-75

Buckler E, Gore M (2007) An Arabidopsis haplotype map takes root. Nat Genet 39:1056-1057

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE,

Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Romay MC, Romero S, Salvo S, Villeda HS, Sofia da Silva H, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD (2009) The genetic architecture of maize flowering time. Science 325:714-718

Burdon JJ, Barrett LG, Rebetzke G, Thrall PH (2014) Guiding deployment of resistance in cereals using evolutionary principles. Evolutionary Applications 7:609–624

Burton GH (1928) Report of plant breeder, Department of Agriculture, Kenya

Busch RH, McVey DV, Linkert GL, Wiersma JV, Warnes DO, Wilcoxson RD, Hareland GA, Edwards I, Schmidt H (1996) Registration of 'Verde' wheat. Crop Science 36:1418

Butler FC (1948) Stem rust of wheat: the value of resistant varieties. Agricultural Gazette of New South Wales 59:511–514

Caldwell RM (1968) Breeding for general and/or specific plant disease resistance. In: Findlay KW, Shepherd KW (eds) Proceedings of the 3rd Int Wheat Genet Symposium. Australian Academy of Science, Canberra, Australia, pp 263-272

Carleton MA (1905) Lessons from the grain-rust epidemic of 1904. Farmers' Bulletin 219:1–24

Carter AH, Chen XM, Garland-Campbell K, Kidwell KK (2009) Identifying QTL for high-temperature adult-plant resistance to stripe rust (*Puccinia striiformis* f. sp. *tritici*) in the spring wheat (*Triticum aestivum* L.) cultivar 'Louise'. Theoretical and Applied Genetics 119:1119-1128

Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A, See D, Bai G, Pumphrey M, Tomar L, Wong D, Kong S, Reynolds M, da Silva ML, Bockelman H, Talbert L, Anderson JA, Dreisigacker S, Baenziger S, Carter A, Korzun V, Morrell PL, Dubcovsky J, Morell MK, Sorrells ME, Hayden MJ, 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 doi: 101073/pnas1217133110

Chao S, Dubcovsky J, Dvorak J, Luo M-C, Baenziger S, Matnyazov R, Clark D, Talbert L, Anderson J, Dreisigacker S, Glover K, Chen J, Campbell K, Bruckner P, Rudd J, Haley S, Carver B, Perry S, Sorrells M, Akhunov E (2010) Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). BMC Genomics 11:727

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theoretical and Applied Genetics 78:495-504

Chao S, Zhang W, Akhunov E, Sherman J, Ma Y, Luo M-C, Dubcovsky J (2009) Analysis of gene-derived SNP marker polymorphism in US wheat (*Triticum aestivum* L.) cultivars. Mol Breeding 23:23-33

Chao S, Zhang W, Dubcovsky J, Sorrells M (2007) Evaluation of genetic diversity and genome-wide linkage disequilibrium among U.S. wheat (*Triticum aestivum* L.) germplasm representing different market classes. Crop Science 47:1018-1030

Charcosset A, Mangin B, Moreau L, Combes L, Jourjon M-F (2000) Heterosis in maize investigated using connected RIL populations. Quantitative genetics and breeding methods: The way ahead. INRA Editions, Paris, France, pp 89–98

Chester KS, Gilbert FA, Hay RE, Newton N (1951) Cereal rusts: epidemiology, losses, and control. Battellc Memorial Institute, Columbus, Ohio

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

Cook, Hims, Vaughan (1999) Effects of fungicide spray timing on winter wheat disease control. Plant Pathology 48:33-50

de Bary A (1853) Investigations of the brand fungi and the diseases of plants caused by them with reference to grain and other useful plants [translated from German by Heffner RMS, Arny DC, Moore JD Phytopathological Classics 11; American Phytopathological Society, St. Paul, Minnesota, USA (1969)]

Devey ME, Hart GE (1993) Chromosomal localization of intergenomic RFLP loci in hexaploid wheat. Genome 36:913-918

Deynze AEV, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, Dvořák J, Gill BS, Lagudah ES, McCouch SR, Appels R (1995) Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. Genome 38:45-59

Dubin HJ, Brennan JP (2009) Combating stem and leaf rust of wheat: historical perspective, impacts, and lessons learned. International Food Policy Research Institute Discussion Paper:1-64

Dyck PL (1992) Transfer of a gene for stem rust resistance from *Triticum araraticum* to hexaploid wheat. Genome 35:788-792

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379

Eucharia EN-O (1990) Health research design and methodology. CRC Press, FL, USA

Evanega SD, Singh RP, Coffman R, Pumphrey MO (2014) The Borlaug Global Rust Initiative: reducing the genetic vulnerability of wheat to rust. In: Tuberosa R, Graner A, Frison E (eds) Genomics of Plant Genetic Resources. Springer Netherlands, pp 317-331

Fantham E (1998) Ovid: fasti book IV. Cambridge University Press, UK

Faris JD, Haen KM, Gill BS (2000) Saturation mapping of a gene-rich recombination hot spot region in wheat. Genetics 154:823-835

Ferdosi M, Kinghorn B, van der Werf J, Lee S, Gondro C (2014) hsphase: an R package for pedigree reconstruction, detection of recombination events, phasing and imputation of half-sib family groups. BMC Bioinformatics 15:172

Fontana F (1767) Observations on the rust of grain. American Phytopathological Society, St Paul, Minnesota, USA

Frascaroli E, Schrag TA, Melchinger AE (2013) Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. Theor Appl Genet 126:133-141

Fu Y-B (2014) Genetic diversity analysis of highly incomplete SNP genotype data with imputations: an empirical assessment. G3: Genes|Genomes|Genetics doi: 10.1534/g3.114.010942

Fu Y-B, Cheng B, Peterson G (2014) Genetic diversity analysis of yellow mustard (*Sinapis alba* L.) germplasm based on genotyping by sequencing. Genet Resour Crop Evol 61:579-594

Fulling EH (1943) Plant life and the law of man. IV. Barberry, currant and gooseberry, and cedar control. Botanical Review 9:483-592

Ghazvini H, Hiebert C, Zegeye T, Liu S, Dilawari M, Tsilo T, Anderson J, Rouse M, Jin Y, Fetch T (2012) Inheritance of resistance to Ug99 stem rust in wheat cultivar Norin 40 and genetic mapping of *Sr42*. Theoretical and Applied Genetics 125:817-824

Gill BS, C. SH, Raupp WJ, Browder LE, Hatchett JH, Harvey TL, Moseman GJ, Waines JG (1985) Evaluation of *Aegilops species* for resistance to wheat powdery mildew, wheat leaf rust, hessian fly and greenbug. Plant Disease 69:314–316

Gold J, Harder D, Townley-Smith F, Aung T, Procunier J (1999) Development of a molecular marker for rust resistance genes *Sr39* and *Lr35* in wheat breeding lines. Electronic Journal of Biotechnology 2:1-6

Green GJ, Martens JW, Ribeiro O (1969) Epidemiology and specialization of wheat and oat stem rusts in Kenya in 1968. Phytopathology 60:309-314

Guo B, Beavis W (2011) In silico genotyping of the maize nested association mapping population. Mol Breeding 27:107-113

Guo B, Sleper DA, Beavis WD (2010) Nested association mapping for identification of functional markers. Genetics 186:373-383

Gupta PK, Kumar J, Mir RR, Kumar A (2010) Marker-assisted selection as a component of conventional plant breeding. Plant Breeding Reviews. John Wiley & Sons, Inc., pp 145-217

Haile JK, Röder MS (2013) Status of genetic research for resistance to Ug99 race of *Puccinia graminis* f. sp. *tritici*: A review of current research and implications. African Journal of Agricultural Research 8:6670-6680

Hamilton LM (1939) Stem rust in the spring wheat area in 1878. Minnesota History 20:156–164

Harder DE, Mathenge GR, Mwaura LK (1971) Physiologic specialization and epidemiology of wheat stem rust in East Africa. Phytopathology 62:166-171

Hare RA, McIntosh RA (1979) Genetic and cytogenetic studies of durable adult-plant resistances in Hope and related cultivars to wheat rusts. Zeitschrift für Pflanzenzüchtung 83:350-367

Hayden MJ, Kuchel H, Chalmers KJ (2004) Sequence tagged microsatellites for the Xgwm533 locus provide new diagnostic markers to select for the presence of stem rust resistance gene *Sr2* in bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 109:1641-1647

Hayes HK, Ausemus ER, Stakman EC, Bailey CH, Wilson HK, Bamberg RH, Morkley MC, Crim RF, Levine MN (1936) Thatcher wheat. Station Bulletin - Minnesota Agriculture Experimental Station 325:1–36

Hayes HK, Stakman EC, Aamodt OS (1925) Inheritance in wheat of resistance to black stem rust. Phytopathology 15:371-387

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

Heslot N, Rutkoski J, Poland J, Jannink J-L, Sorrells ME (2013) Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. PLoS ONE 8:e74612

Hiebert C, Fetch T, Zegeye T, Thomas J, Somers D, Humphreys DG, McCallum B, Cloutier S, Singh D, Knott D (2011) Genetics and mapping of seedling resistance to Ug99 stem rust in Canadian wheat cultivars 'Peace' and 'AC Cadillac'. Theor Appl Genet 122:143-149

Hodson DP, Nazari K, Park RF, Hansen J, Lassen P, Arista J, Fetch T, Hovmøller M, Jin Y, Pretorius ZA, Sonder K (2011) Putting Ug99 on the map: an update on current and future monitoring. http://www.globalrust.org/db/attachments/about/309/1/Hodson.pdf [Accessed on 01-21-2014]

Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, Tao Y, Zhang X, Jing R, Zhang C, Ma Y, Gao L, Gao C, Spannagl M, Mayer KF, Li D, Pan S, Zheng F, Hu Q, Xia X, Li J, Liang Q, Chen J, Wicker T, Gou C, Kuang H, He G, Luo Y, Keller B, Xia Q, Lu P, Wang J, Zou H, Zhang R, Xu J, Gao J, Middleton C, Quan Z, Liu G, Wang J, Yang H, Liu X, He Z, Mao L, Wang J (2013) *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature 496:91-95

Jin Y (2005) Races of *Puccinia graminis* identified in the United States during 2003. Plant Disease 89:1125-1127

Jin Y, Szabo LJ, Pretorius ZA, Singh RP, Ward R, Fetch T (2008) Detection of virulence to resistance gene *Sr24* within race TTKS of *Puccinia graminis* f. sp. *tritici*. Plant Disease 92:923-926

Jin Y, Szabo LJ, Rouse MN, Fetch T, Pretorius ZA, Wanyera R, Njau P (2009a) Detection of virulence to resistance gene *Sr36* within the TTKS race lineage of *Puccinia graminis* f. sp. *tritici*. Plant Disease 93:367–370

Jin Y, Szabo LJ, Rouse MN, G. FT, Pretorius ZA, Wanyera R, Njau P (2009b) Detection of virulence to resistance gene *Sr36* within the TTKS race lineage of *Puccinia graminis* f. sp. *tritici*. Plant Disease 93:367-370

Johnston SJ, Sharp PJ, McIntosh RA (1998) Molecular markers for the *Sr2* stem rust resistance gene. In: Slinkard AE (ed) Proceedings of the 9th International Wheat Genetics Symposium. University Extension Press, University of Saskatchewan, Saskatoon, Canada, pp 78-80

Jones SS, Dvořák J, Knott DR, Qualset CO (1991) Use of double-ditelosomic and normal chromosome 1D recombinant substitution lines to map *Sr33* on chromosome arm 1DS in wheat. Genome 34:505-508

Joshi LM, Palmer LT (1973) Epidemiology of stem, leaf and stripe rusts of wheat in northern India. The Plant Disease Reporter 57:8–12

Jourjon M-F, Jasson S, Marcel J, Ngom B, Mangin B (2005) MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics 21:128-130

Kaur J, Bansal U, Khanna R, Saini R, Bariana H (2009) Molecular mapping of stem rust resistance in HD2009/WL711 recombinant inbred line population. International Journal of Plant Breeding 3:77-33

Kerber ER, Dyck PL (1973) Inheritance of stem rust resistance transferred from diploid wheat (*Triticum monococum*) to tetraploid and hexaploid wheat and chromosome location of the gene involved. Canadian Journal of Genetics and Cytology 15:397-409

Kerber ER, Dyck PL (1990) Transfer to hexaploid wheat of linked genes for adult-plant leaf rust and seedling stem rust resistance from an amphiploid of *Aegilops speltoides* × *Triticum monococcum*. Genome 33:530-537

Kidwell K, Osborn T (1992) Simple plant DNA isolation procedures. In: Beckmann JS, Osborn TC (eds) Plant Genomes: Methods for Genetic and Physical Mapping. Springer Netherlands, pp 1-13

Kislev ME (1982) Stem rust of wheat 3300 years old found in Israel. Science 216:993-994

Klindworth DL, Miller JD, Williams ND, Xu SS (2011) Resistance to recombinant stem rust race TPPKC in hard red spring wheat. Theoretical and Applied Genetics 123:603-613

Klindworth DL, Niu Z, Chao S, Friesen TL, Jin Y, Faris JD, Cai X, Xu SS (2012) Introgression and characterization of a goatgrass gene for a high level of resistance to Ug99 stem rust in tetraploid wheat. G3: Genes|Genomes|Genetics 2:665-673

Knott DR (1962) Inheritance of rust resistance. VIII. Additional studies on Kenya varieties of wheat. Crop Science 2:130-132

Knott DR (1968) The inheritance of resistance to stem rust races 56 and 15B-IL (Can.) in the wheat varieties Hope and H-44. Canadian Journal of Genetics and Cytology 10:311-320

Knott DR (1982) Multigenic inheritance of stem rust resistance in wheat. Crop Sci 22:393-399

Knott DR (1989) The Wheat Rusts — Breeding for Resistance. Theoretical and Applied Genetics 12:201

Knott DR (1990) Near-isogenic lines of wheat carrying genes for stem rust resistance. Crop Sci 30:901-905

Knott DR (2001) The relationship between seedling and field resistance to two races of stem rust in Thatcher wheat. Canadian Journal of Plant Science 81:415-418

Knott DR, Anderson RG (1956) The inheritance of rust resistance. I. The inheritance of stem rust resistance in ten varieties of common wheat. Canadian Journal of Agricultural Science 36:174-195

Kosambi DD (1943) The estimation of map distance from recombination values. Annals of Eugenics 12:172-175

Lagudah ES, McFadden H, Singh RP, Huerta-Espino J, Bariana HS, Spielmeyer W (2006) Molecular genetic characterization of the *Lr34*/*Yr18* slow rusting resistance gene region in wheat. Theor Appl Genet 114:21-30

Letta T, Maccaferri M, Badebo A, Ammar K, Ricci A, Crossa J, Tuberosa R (2013) Searching for novel sources of field resistance to Ug99 and Ethiopian stem rust races in durum wheat via association mapping. Theoretical and Applied Genetics 126:1237-1256

Letta T, Olivera P, Maccaferri M, Jin Y, Ammar K, Badebo A, Salvi S, Noli E, Crossa J, Tuberosa R (2014) Association mapping reveals novel stem rust resistance loci in durum wheat at the seedling stage. Plant Gen 0:-

Li B, Kimmel M (2013) Factors influencing ascertainment bias of microsatellite allele sizes: impact on estimates of mutation rates. Genetics 195:563-572

Lindhout P (2002) The perspectives of polygenic resistance in breeding for durable disease resistance. Euphytica 124:217-226

Ling H-Q, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, Gao C, Wu H, Li Y, Cui Y, Guo X, Zheng S, Wang B, Yu K, Liang Q, Yang W, Lou X, Chen J, Feng M, Jian J, Zhang X, Luo G, Jiang Y, Liu J, Wang Z, Sha Y, Zhang B, Wu H, Tang D, Shen Q, Xue P, Zou S, Wang X, Liu X, Wang F, Yang Y, An X, Dong Z, Zhang K, Zhang X, Luo M-C, Dvorak J, Tong Y, Wang J, Yang H, Li Z, Wang D, Zhang A, Wang J (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496:87-90

Liu H, Bayer M, Druka A, Russell J, Hackett C, Poland J, Ramsay L, Hedley P, Waugh R (2014) An evaluation of genotyping by sequencing (GBS) to map the *Breviaristatum-e* (*ari-e*) locus in cultivated barley. BMC Genomics 15:104

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128-2129

Liu S, Yu L-X, Singh R, Jin Y, Sorrells M, Anderson J (2010) Diagnostic and co-dominant PCR markers for wheat stem rust resistance genes *Sr25* and *Sr26*. Theoretical and Applied Genetics 120:691-697

Liu W, Rouse M, Friebe B, Jin Y, Gill B, Pumphrey M (2011) Discovery and molecular mapping of a new gene conferring resistance to stem rust, *Sr53*, derived from *Aegilops geniculata* and characterization of spontaneous translocation stocks with reduced alien chromatin. Chromosome Res 19:669-682

Lorieux M (2012) MapDisto: fast and efficient computation of genetic linkage maps. Mol Breeding 30:1231-1235

Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genet 9:e1003215

Lukaszewski AJ, Curtis CA (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theoretical and Applied Genetics 86:121-127

Macharia GK (2013) Molecular diversity, linkage disequilibrium and genetic mapping in East Africa wheat. Department of Agronomy. University of Minnesota, St Paul, MN, USA

Mago R, Bariana HS, Dundas IS, Spielmeyer W, Lawrence GJ, Pryor AJ, Ellis JG (2005) 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, Brown-Guedira G, Dreisigacker S, Breen J, Jin Y, Singh R, Appels R, Lagudah ES, Ellis J, Spielmeyer W (2011a) An accurate DNA marker assay for stem rust resistance gene *Sr2* in wheat. Theoretical and Applied Genetics 122:735-744

Mago R, Lawrence G, Ellis J (2011b) 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, Zhang P, Bariana HS, Verlin DC, Bansal UK, Ellis JG, Dundas IS (2009) Development of wheat lines carrying stem rust resistance gene *Sr39* with reduced *Aegilops speltoides* chromatin and simple PCR markers for marker-assisted selection. Theoretical and Applied Genetics 119:1441-1450

MAS Wheat (2014) http://maswheat.ucdavis.edu/protocols/StemRust/index.htm. Accessed on 05-24-2014. UC-Davis

Mayfield AH (1985) Efficacies of fungicides for control of stem rust of wheat. Australian Journal of Experimental Agriculture 25:440–443

McAlpine D (1906) The rusts of Australia. Their structure, nature, and classification. Department of Agriculture, Melbourne, Australia

McDonald J (1931) The existence of physiologic forms of wheat stem rust in Africa. Transactions of the British Mycological Society 15:235

McFadden ES (1930) A successful transfer of emmer characters to vulgare wheat. Journal of the American Society of Agronomy 22:1020-1034

McFadden ES (1939) Brown necrosis, a discoloration associated with rust infection in certain rust resistant wheats. Journal of Agricultural Research 58:805–819

McIntosh RA, Dubcovsky J, Rojers WJ, Morris C, Sommers DJ, Appels R, Xia XC (2011) Catalogue of gene symbols for wheat: 2011 supplement. Annual Wheat Newsletter 57:303–321

McIntosh RA, Gyarfas J (1971) *Triticum timopheevi* as a source of resistance to wheat stem rust. Zeitschrift für Pflanzenzüchtung 66:240-248

McIntosh RA, Wellings CR, Park RF (1995) Wheat rusts: an atlas of resistance genes. CSIRO Publications, Victoria, Australia

McIntosh RA, Yamazaki Y, Dubcovsky J, Rogers J, Morris C, Somers DJ, Appels R, Devos KM (2003) Catalogue of gene symbols for wheat. 11th International Wheat Genetics Symposium, Brisbane, Australia

McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, Brown P, Browne C, Eller M, Guill K, Harjes C, Kroon D, Lepak N, Mitchell SE, Peterson B, Pressoir G, Romero S, Rosas MO, Salvo S, Yates H, Hanson M, Jones E, Smith S, Glaubitz JC, Goodman M, Ware D, Holland JB, Buckler ES (2009) Genetic properties of the maize nested association mapping population. Science 325:737-740

Moragues M, Comadran J, Waugh R, Milne I, Flavell AJ, Russell JR (2010) Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. Theor Appl Genet 120:1525-1534

Moser G, Khatkar MS, Raadsma HW (2009) Imputation of missing genotypes in high density SNP data. Proceedings of 18th Conference Of The Association For The Advancement Of Animal Breeding And Genetics, Barrosa Valley, Australia, pp 612-615

Mundt CC (2014) Durable resistance: A key to sustainable management of pathogens and pests. Infection, Genetics and Evolution DOI: 10.1016/j.meegid.2014.01.011

Nazareno NRX, Roelfs AP (1981) Adult plant resistance of Thatcher wheat to stem rust. Phytopathology 71:181-185

Negeri A, Coles N, Holland J, Balint-Kurti P (2011) Mapping QTL controlling southern leaf blight resistance by joint analysis of three related recombinant inbred line populations. Crop Science 51:1571-1579

Nelson JC, Deynze AEV, Sorrells ME, Autrique E, Lu YH, Merlino M, Atkinson M, Leroy P (1995) Molecular mapping of wheat homoeologous group 2. Genome 38:516-524

Njau PN, Jin Y, Huerta-Espino J, Keller B, Singh RP (2010) Identification and evaluation of sources of resistance to stem rust race Ug99 in wheat. Plant Disease 94:413-419

Njau PN, Wanyera R, Macharia GK, Macharia J, Singh R, Keller B (2009) Resistance in Kenyan bread wheat to recent eastern African isolate of stem rust, *Puccinia graminis* f. sp. *tritici*, Ug99. Journal of Plant Breeding and Crop Science 1:22-27

Olson EL, Brown-Guedira G, Marshall D, Stack E, Bowden RL, Jin Y, Rouse M, Pumphrey MO (2010a) Development of wheat lines having a small introgressed segment carrying stem rust resistance gene *Sr22*. Crop Science 50:1823-1830

Olson EL, Brown-Guedira G, Marshall DS, Jin Y, Mergoum M, Lowe I, Dubcovsky J (2010b) Genotyping of U.S. wheat germplasm for presence of stem rust resistance genes *Sr24*, *Sr36* and $Sr1RS^{Amigo}$. Crop Science 50:668-675

Parlevliet JE (1976) Partial resistance of barley to leaf rust, *Puccinia hordei*. III. The inheritance of the host plant effect on latent period in four cultivars. Euphytica 25:241-248

Paull JG, Pallotta MA, Langridge P, The TT (1994) RFLP markers associated with *Sr22* and recombination between chromosome 7A of bread wheat and the diploid species *Triticum boeoticum*. Theoretical and Applied Genetics 89:1039-1045

Periyannan S, Bansal U, Bariana H, Deal K, Luo M-C, Dvorak J, Lagudah E (2014) Identification of a robust molecular marker for the detection of the stem rust resistance gene *Sr45* in common wheat. Theoretical and Applied Genetics 127:947-955

Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, Deal K, Luo M, Kong X, Bariana H, Mago R, McIntosh R, Dodds P, Dvorak J, Lagudah E (2013) The gene *Sr33*, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. Science 341:786-788

Persoon CH (1794) Neuer versuch einer systematischen eintheilung der schwämme. Roemer's Neues Magazin für die Botanik 1:63–128

Peterson RF, Campbell AB, Hannah AE (1948) A diagramatic scale for estimating rust intensity of leaves and stem of cereals. Canadian Journal of Research 26c:496-500

Plessers AG (1954) The genetics of stem and leaf rust reactions and other characters in Lee wheat with Chinese monosomic testers. Dissertation Abstracts 15

Poland JA, Brown PJ, Sorrells ME, Jannink J-L (2012a) Development of high-density genetic maps for barley and wheat using a aovel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7:e32253

Poland JA, Brown PJ, Sorrells ME, Jannink J-L (2012b) Development of high-density genetic maps for barley and wheat using a novel two-enzyme gnotyping-by-sequencing approach. PLoS ONE 7:e32253

Pozniak CJ, Reimer S, Fetch T, Clarke JM, Clarke FR, Somers DJ, Knox RE, Singh AK (2008) Association mapping of Ug99 resistance in diverse durum wheat population. In: Rudi A, Russell E, Peter L, Michael M, Lynne M, Peter S (eds) 11th International Wheat Genetics Symposium, Brisbane, Australia, pp 809–811

Pretorius ZA, Singh RP, Wagoire WW, Payne TS (2000) Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis* f. sp. *tritici* in Uganda. Plant Disease 84

Pretorius ZA, Szabo LJ, Boshoff WHP, Herselman L, Visser B (2012) First report of a new TTKSF race of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) in South Africa and Zimbabwe. Plant Disease 96:590-590

Prins R, Pretorius ZA, Bender CM, Lehmensiek A (2011a) QTL mapping of stripe, leaf and stem rust resistance genes in a 'Kariega' × 'Avocet S' doubled haploid wheat population. Mol Breeding 27:259–270

Prins R, Pretorius ZA, Bender CM, Lehmensiek A (2011b) QTL mapping of stripe, leaf and stem rust resistance genes in a Kariega × Avocet S doubled haploid wheat population. Mol Breeding 27:259-270

Qi LL, Pumphrey MO, Friebe B, Zhang P, Qian C, Bowden RL, Rouse MN, Jin Y, Gill BS (2011) A novel Robertsonian translocation event leads to transfer of a stem rust resistance gene (*Sr52*) effective against race Ug99 from *Dasypyrum villosum* into bread wheat. Theor Appl Genet 123:159-167

Rebai A, Goffinet B (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. Theoretical and Applied Genetics 86:1014-1022

Remington DL, Ungerer MC, Purugganan MD (2001) Map-based cloning of quantitative trait loci: progress and prospects. Genetics Research 78:213-218

Risk JM, Selter LL, Chauhan H, Krattinger SG, Kumlehn J, Hensel G, Viccars LA, Richardson TM, Buesing G, Troller A, Lagudah ES, Keller B (2013) The wheat *Lr34* gene provides resistance against multiple fungal pathogens in barley. Plant Biotechnol J 11:847-854

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. Genetical Research 67:175-185

Roelfs AP (1977) Foliar fungal diseases of wheat in the People's Republic of China. The Plant Disease Reporter 61:836–841

Roelfs AP (1985a) Epidemiology in North America. In 'The Cereal Rusts'. Vol. II. Academic Press, Orlando, Florida, USA

Roelfs AP (1985b) Wheat and rye stem rust. In: Roelfs AP, Bushnell WR (eds) The Cereal Rusts. Academic Press, Orlando, FL, USA, pp 3–37

Roelfs AP, Martens JW (1988) An international system of nomenclature for *Puccinia graminis* f. sp. *tritici*. Phytopathology 78:526-533

Roelfs AP, McVey DV (1979) Low infection types produced by *Puccinia graminis* f.sp. *tritici* and wheat lines with designated genes for resistance. Phytopathology 69:722-730

Roelfs AP, Singh RP, Saari EE (1992) Rust diseases of wheat: Concepts and methods of disease management. CIMMYT, Mexico

Rouse M, Nava I, Chao S, Anderson J, Jin Y (2012) Identification of markers linked to the race Ug99 effective stem rust resistance gene *Sr28* in wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 125:877-885

Rouse MN, Jin Y (2011) Stem rust resistance in A-genome diploid relatives of wheat. Plant Disease 95:941-944

Rouse MN, Talbert LE, Singh D, Sherman JD (2014) Complementary epistasis involving *Sr12* explains adult plant resistance to stem rust in Thatcher wheat (*Triticum aestivum* L.). Theor Appl Genet 127:1549-1559

Rutkoski JE, Poland J, Jannink JL, Sorrells ME (2013) Imputation of unordered markers and the impact on genomic selection accuracy. G3 (Bethesda, Md) 3:427-439

Saintenac C, Jiang D, Wang S, Akhunov E (2013a) Sequence-based mapping of the polyploid wheat genome. G3: Genes|Genomes|Genetics 3:1105-1114

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

Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends in Plant Science 10:297-304

Sandhu D, Gill KS (2002) Gene-containing regions of wheat and the other grass genomes. Plant Physiology 128:803–811

Sawhney RN, Singh D, Joshi BC (1981) Monosomic analysis of genes for resistance against stem rust races in bread wheat. Theoretical and Applied Genetics 59:313-316

Seah S, Spielmeyer W, Jahier J, Sivasithamparam K, Lagudah ES (2000) Resistance gene analogs within an introgressed chromosomal segment derived from *Triticum ventricosum* that confers resistance to nematode and rust pathogens in wheat. Molecular Plant-Microbe Interactions 13:334-341

Sears ER (1954) The aneuploids of common wheat. University of Missouri Research Bulletin 572

Sears ER, Loegering WQ, Rodenhiser HA (1957) Identification of chromosomes carrying genes for stem rust resistance in four varieties of wheat. Agronomy Journal 49:208-212

Sharp PJ, Johnston S, Brown G, McIntosh RA, Pallotta M, Carter M, Bariana HS, Khatkar S, Lagudah ES, Singh RP, Khairallah M, Potter R, Jones MGK (2001) Validation of molecular markers for wheat breeding. Australian Journal of Agricultural Research 52:1357-1366

Sidhu D, Gill KS (2004) Distribution of genes and recombination in wheat and other eukaryotes. Plant Cell, Tissue and Organ Culture 4653PB:1-14

Simons K, Abate Z, Chao S, Zhang W, Rouse M, Jin Y, Elias E, Dubcovsky J (2011) Genetic mapping of stem rust resistance gene *Sr13* in tetraploid wheat (*Triticum turgidum* ssp. *durum* L.). Theoretical and Applied Genetics 122:649-658

Singh A, Knox RE, DePauw RM, Singh AK, Cuthbert RD, Campbell HL, Singh D, Bhavani S, Fetch T, Clarke F (2013a) Identification and mapping in spring wheat of genetic factors controlling stem rust resistance and the study of their epistatic interactions across multiple environments. Theor Appl Genet 126:1951-1964

Singh A, Pandey MP, Singh AK, Knox RE, Ammar K, Clarke JM, Clarke FR, Singh RP, Pozniak CJ, DePauw RM, McCallum BD, Cuthbert RD, Randhawa HS, Fetch TG, Jr. (2013b) Identification and mapping of leaf, stem and stripe rust resistance

quantitative trait loci and their interactions in durum wheat. Mol Breeding 31:405-418

Singh RP (2012) Pros and cons of utilizing major, race-specific resistance genes versus partial resistance in breeding rust resistant wheat. International Maize and Wheat Improvement Center (CIMMYT), Mexico

Singh RP, Herrera-Foessel SA, Huerta-Espino J, Lan CX, Basnet BR, Bhavani S, Lagudah ES (2013c) Pleiotropic gene *Lr46*/*Yr29*/*Pm39*/*Ltn2* confers slow rusting, adult plant resistance to wheat stem rust fungus. Borlaug Global Rust Initiative 2013 Technical Workshop, New Delhi, India, p 17

Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Bhavani S, Njau P, Herrera-Foessel S, Singh PK, Singh S, Govindan V (2011) The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annual Review of Phytopathology 49:465-481

Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Njau P, Wanyera R, Herrera-Foessel SA, Ward RW (2008a) Will stem rust destroy the world's wheat crop? In: Donald LS (ed) Advances in Agronomy. Academic Press, pp 271-309

Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Njau P, Wanyera R, Herrera-Foessel SA, Ward RW (2008b) Will stem rust destroy the world's wheat crop? Advances in Agronomy 98:272-309

Singh RP, Hodson DP, Jin Y, Huerta-Espino J, Kinyua MG, Wanyera R, Njau P, Ward RW (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 RP, Huerta-Espino J, Rajaram S (2000) Achieving near-immunity to leaf and stripe rusts in wheat by combining slow rusting resistance genes. Acta Phytopathologica et Entomologica Hungarica 35:133-139

Singh RP, Huerta EJ, Manilal HM (2005) Genetics and breeding for durable resistance to leaf and stripe rusts in wheat. Turkish Journal of Agriculture and Forestry 29:121-127

Singh S, Singh R, Bhavani S, Huerta-Espino J, Eugenio L-V (2013d) QTL mapping of slow-rusting, adult plant resistance to race Ug99 of stem rust fungus in PBW343/Muu RIL population. Theor Appl Genet 126:1367-1375

Spielmeyer W, Sharp PJ, Lagudah ES (2003) Identification and validation of markers linked to broad-spectrum stem rust resistance gene *Sr2* in wheat (*Triticum aestivum* L.). Crop Sci 43:333-336

Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007) pcaMethods—a bioconductor package providing PCA methods for incomplete data. Bioinformatics 23:1164-1167

Stakman EC, Harrar JG (1957) Principles of plant pathology. Ronald Press, New York

Stakman EC, Stewart DM, Loegering WQ (1962) Identification of physiologic races of *Puccinia graminis* var. *tritici*. US Dep Agric Agric Res Serv E 6/7

Steffenson BJ, Olivera P, Roy JK, Jin Y, Smith KP, Muehlbauer GJ (2007) A walk on the wild side: mining wild wheat and barley collections for rust resistance genes. Australian Journal of Agricultural Research 58:532-544

Stubbs RW, Prescott JM, Saari EE, Dubin HJ (1986) Cereal disease methodology manual, CIMMYT, Mexico

Stuthman DD, Leonard KJ, Miller-Garvin J (2007) Breeding crops for durable resistance to disease. In: Donald LS (ed) Advances in Agronomy. Academic Press, pp 319-367

Suenaga K, Singh RP, Huerta-Espino J, William HM (2003) Microsatellite markers for genes *Lr34*/*Yr18* and other quantitative trait Loci for leaf rust and stripe rust resistance in bread wheat. Phytopathology 93:881-890

Targioni TG (1767) True nature, causes and sad effects of the rust, the bunt, the smut, and other maladies of wheat and of oats in the field. American Phytopathological Society, St Paul, Minnesota, USA

Theophrastus Enquiry into plants II. Harvard University Press Cambridge, Massachussetts

Tsilo TJ, Jin Y, Anderson JA (2007) Microsatellite markers linked to stem rust resistance allele *Sr9a* in wheat. Crop Science 47:2013-2020

Tsilo TJ, Kolmer JA, Anderson JA (2014) Molecular mapping and improvement of leaf rust resistance in wheat breeding lines. Phytopathology 104:865-870

Tulasne LR (1853) Note sur la germination des spores des uredinees. comptes rendus hebdomadaires des séances de l'Académie des Sciences 36:1093-1095 [Translation in Annals and Magazine of Natural History 2, 221-223 (1853)].

Tulasne LR, Tulasne C (1847) Mémoire sur les ustilaginees comparees aux

Wang J, Li H, Zhang L, Meng L (2012a) Users' manual of QTL IciMapping Version 3.3. The Quantitative Genetics Group, Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS), China, and Genetic Resources Program, International Maize and Wheat Improvement Center (CIMMYT), Mexico

Wang S, Basten CJ, Zeng Z-B (2012b) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC

Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J, International Wheat Genome Sequencing C, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, 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 KJ, Hayden M, Akhunov E (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnology Journal:1-10

Wanyera R, Kinyua MG, Jin Y, Singh RP (2006) The spread of stem rust caused by *Puccinia graminis* f. sp. *tritici*, with virulence on *Sr31* in wheat in Eastern Africa. Plant Disease 90:113-113

Wanyera R, Macharia JK, Kilonzo SM, Kamundia JW (2009) Foliar fungicides to control wheat stem rust, race TTKS (Ug99), in Kenya. Plant Disease 93:929-932

Waterhouse WL (1929) Australian rust studies. I. Proceedings of the Linnean Society of New South Wales 54:615–680

Watson IA (1981) Wheat and its rust parasites in Australia. In: Evans LT, Peacock WJ (eds) Wheat Science – Today and Tomorrow. Cambridge University Press, London, UK, pp 129–147

Werner JE, Endo TR, Gill BS (1992) Toward a cytogenetically based physical map of the wheat genome. Proceedings of the National Academy of Sciences of the United States of America 89:11307-11311

Wheat MAS (2014) http://maswheat.ucdavis.edu/protocols/StemRust/index.htm. Accessed on 05-24-2014. UC-Davis

William M, Singh RP, Huerta-Espino J, Islas SO, Hoisington D (2003) Molecular marker mapping of leaf rust resistance gene *Lr46* and its association with stripe rust resistance gene *Yr29* in wheat. Phytopathology 93:153-159

Williams L, Ma X, Boyko A, Bustamante C, Oleksiak M (2010) SNP identification, verification, and utility for population genetics in a non-model genus. BMC Genetics 11:32

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539-551

Yu L-X, Barbier H, Rouse M, Singh S, Singh R, Bhavani S, Huerta-Espino J, Sorrells M (2014) A consensus map for Ug99 stem rust resistance loci in wheat. Theor Appl Genet 127:1561-1581

Yu L-X, Lorenz A, Rutkoski J, Singh R, Bhavani S, Huerta-Espino J, Sorrells M (2011) Association mapping and gene–gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. Theor Appl Genet 123:1257-1268

Yu L-X, Morgounov A, Wanyera R, Keser M, Singh S, Sorrells M (2012) Identification of Ug99 stem rust resistance loci in winter wheat germplasm using genome-wide association analysis. Theor Appl Genet 125:749-758

Zadoks JC (1963) Epidemiology of wheat rusts in Europe. FAO Plant Protection Bulletin 13:97–108

Zadoks JC, Chang TT, Konzak CF (1974) A decimal code for the growth stages of cereals. Weed Research 14:415-421

Zeller FJ (1973) 1B/1R wheat-rye chromosome substitutions and translocations. In: Sears ER, Sears LMS (eds) Proceedings of the Fourth International Wheat Genetics Symposium. Agricultural Experiment Station, University of Missouri, Columbia, Missouri, USA, pp 209-221

Zhang D, Bowden RL, Yu J, Carver BF, Bai G (2014) Association analysis of stem rust resistance in U.S. winter wheat. PLoS ONE 9:e103747

Zhensheng K, Jie Z, Dejun H, Hongchang Z, Xiaojie W, Chenfang W, Qingmei H, Jun G, Lili H (2010) Status of wheat rust research and control in China. BGRI 2010 Technical Workshop 30-31 May, St. Petersburg, Russia

APPENDIX

Appendix I: QTL detected via CIM in the 9K (Table A), non-imputed GBS (Table B), GBS40 (Table C), and GBS75 (Table D) datasets.

A.

Environment	Chromosome	Marker	Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken12	2B	26	24.7	15.7	32.6	-6.2
	2B	93	122.5	3.1	6.0	2.6
	4A	9	9.2	3.4	6.6	-2.8
	6D	2	2.4	2.9	5.1	2.4
Ken13	2B	17	14.2	4.8	23.7	-5.0
	4B	9	25.8	2.6	11.0	-3.3
	5B	3	0.4	4.0	18.5	5.3
Eth13	2B	26	24.7	17.0	57.6	-11.2
	3A	22	60.3	2.8	6.1	-3.7
StP13	2B	20	21.4	3.8	9.8	-2.6
	4B	9	24.8	2.6	5.7	2.1
	4B	34	49.0	4.7	10.4	-3.0
	4B	77	62.1	3.0	12.1	-3.0

B.

Environment	Chromosome	Marker	Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken12	2B	23	15.0	4.1	7.1	-2.9
	2B	104	93.9	15.4	32.1	6.2
	7A	6	0.8	4.1	6.7	-2.8
Ken13	1A	57	92.9	3.4	14.3	3.8
	2B	107	95.7	5.2	29.6	5.6
	2D	7	32.2	3.3	14.7	-3.8
Eth13	2A	11	30.6	2.8	7.4	3.9
	2B	105	94.5	13.8	49.6	10.4

Environment	Chromosome	Marker	Pos (cM)	LOD[a]	R^{2b}	Add[c]
Eth13	4B	17	33.6	2.7	5.7	3.6
StP13	2B	115	107.1	3.8	8.8	2.5
	4A	40	21.8	2.6	6.0	2.1

C.

Environment	Chromosome	Marker	Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken12	2B	17	58.9	5.0	8.3	-3.2
	2B	104	307.1	17.2	36.1	6.6
	7A	5	24.4	4.4	7.2	-2.9
Ken13	1A	57	312.1	3.4	14.4	3.9
	2B	103	295.2	2.5	11.2	3.7
	2B	113	331.6	2.5	9.9	4.0
	2D	9	97.8	3.3	13.6	-3.7
Eth13	2B	105	308.7	14.4	52.6	10.7
StP13	2B	107	321.0	3.6	9.2	2.5
	4A	37	60.4	2.6	5.9	2.1
	4B	50	208.1	2.9	6.6	2.2

D.

Environment	Chromosome	Marker	Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken12	2B	35	696.2	2.8	4.6	-2.3
	2B	122	2237.3	16.3	34.1	6.4
	7B	9	167.2	3.6	6.1	-2.7
Ken13	1A	56	1004.8	3.3	14.4	3.9
	2B	121	2228.4	2.5	11.8	3.8
	3A	18	377.1	3.2	13.5	-3.6
Eth13	2B	123	2238.2	15.2	53.5	10.8
StP13	2B	128	2328.3	3.7	9.4	2.6
	4B	16	172.3	2.8	6.2	2.1
	5B	24	567.3	3.0	6.9	2.3

[a] LOD values are the peak logarithm of odds score for the given QTL

[b] Value indicates the phenotypic variation explained by the QTL

[c] Value indicates the estimated additive effect of the QTL; negative value means that the allele was contributed by RB07

Appendix II: Quantitative trait loci (QTL) detected for stem rust adult plant resistance by joint mapping of 10 RIL populations in four environments.

Env Chr[a]	QTL[b]	Left SNP	Right SNP	Pos[c]	LOD[d]	PVE[e] (%)	Ada[f]	Fahari[f]	Gem[f]	Kudu[f]	Kulungu[f]	Ngiri[f]	Paka[f]	Pasa[f]	Popo[f]	Romany[f]
Ken13																
2A	QSr.umn-2A.1	SNP126	SNP652	159.1	3.3	2.1	4.4	-0.7	-3.2	1.9	3.3	-3.4	-0.2	-1.5	-0.4	0.3
2B	QSr.umn-2B.1	SNP970	SNP358	444.8	4.2	2.6	-3.0	1.8	-0.6	2.1	2.2	-2.8	-0.2	-0.9	0.9	-0.4
3B	QSr.umn-3B.1	SNP666	SNP243	295.0	3.2	2.0	0.4	-0.9	-4.5	2.0	6.0	-2.8	-2.1	1.0	0.9	-0.2
3B	QSr.umn-3B.2	SNP4	SNP813	723.8	2.9	1.9	2.5	-1.3	2.9	-1.8	-4.7	0.4	1.8	1.0	1.8	-2.0
3B	QSr.umn-3B.3	SNP350	SNP353	787.2	3.7	2.3	-10.4	4.2	-0.8	-4.8	0.0	5.1	3.0	0.8	0.2	-0.1
4A	QSr.umn-4A.1	SNP371	SNP375	164.3	2.5	1.5	6.8	-1.2	-0.9	-6.5	-2.8	4.1	2.0	-3.0	1.5	0.8
4A	QSr.umn-4A.2	SNP973	SNP522	644.9	6.6	4.1	5.5	1.2	-2.8	1.9	-7.8	-2.7	-2.3	2.3	3.2	1.0
4B	QSr.umn-4B.1	SNP159	SNP435	401.3	2.8	0.6	-2.9	0.2	-0.8	0.0	0.2	0.8	0.8	1.0	2.7	-2.3
5B	QSr.umn-5B.1	SNP455	SNP150	607.2	2.5	1.6	-1.0	0.9	1.9	-0.6	1.3	-2.3	1.5	0.1	0.8	-3.1
6A	QSr.umn-6A	SNP935	SNP425	210.3	4.2	2.6	1.3	-26.7	1.2	4.7	3.5	0.0	3.8	2.0	1.0	6.9
7A	QSr.umn-7A	SNP191	SNP181	97.5	4.1	2.5	2.0	-1.9	-2.1	10.0	0.0	-5.2	-0.1	-1.8	-0.5	1.5
SA12																
4B	QSr.umn-4B.2	SNP177	SNP203	358.8	7.3	4.0	-0.1	0.1	6.6	-0.8	-0.4	-0.4	0.4	-5.5	-0.2	0.4
7B	QSr.umn-7B	SNP268	SNP337	43.0	7.6	4.1	-1.0	0.7	2.2	2.0	0.6	2.2	-0.3	0.8	0.5	-8.1
StP12																
2A	QSr.umn-2A.2	SNP537	SNP494	275.8	3.6	2.2	-5.4	7.6	-0.9	-13.5	4.7	5.3	8.1	4.8	3.2	1.5
2B	QSr.umn-2B.2	SNP294	SNP970	439.9	6.1	3.8	0.0	-2.7	-0.4	2.8	6.4	-4.8	-2.2	-1.2	0.9	-0.9
2D	QSr.umn-2D	SNP116	SNP687	185.1	2.9	1.8	2.4	-0.4	-1.3	0.1	-5.5	-3.9	1.5	2.7	-1.9	4.5

Env																
Chr[a]	QTL[b]	Left SNP	Right SNP	Pos[c]	LOD[d]	PVE[e] (%)	Ada[f]	Fahari[f]	Gem[f]	Kudu[f]	Kulungu[f]	Ngiri[f]	Paka[f]	Pasa[f]	Popo[f]	Romany[f]
StP12																
3B	QSr.umn-3B.4	SNP89	SNP395	89.4	2.7	1.4	-6	-1.9	10.7	-3.5	17.3	0.0	-5.2	-9.1	1.6	-4.1
3B	QSr.umn-3B.5	SNP154	SNP93	642.7	5.8	3.6	-0.7	-0.3	-2.2	-0.1	4.6	4.4	0.7	2.8	-11.5	0.6
3B	QSr.umn-3B.6	SNP790	SNP924	684.6	6.2	3.8	8.2	2.9	3.7	-1.6	-4.7	-5.2	-0.7	1.1	-0.8	-2.1
3B	QSr.umn-3B.7	SNP320	SNP359	811.0	8.3	5.0	-4.9	4.5	-26.2	-11.2	-0.6	6.8	3.9	0.1	19.2	2.5
4A	QSr.umn-4A.1	SNP371	SNP375	164.3	2.5	1.5	3.4	-4.7	-4.6	-4.3	0.3	8.0	2.3	1.4	-3.8	3.4
4B	QSr.umn-4B.3	SNP205	SNP210	301.4	4.3	2.6	3.0	5.3	-9.3	-1.6	-13.4	13.4	3.2	-0.3	1.5	-3.3
5B	QSr.umn-5B.2	SNP142	SNP784	624.2	2.8	1.8	-1.9	-0.8	4.7	-2.8	1.7	-1.4	-2.6	-1.7	3.0	1.7
5D	QSr.umn-5D.1	SNP714	SNP481	138.0	3.0	1.9	-4.3	0.6	0.7	-0.4	0.0	2.7	2.8	-3.2	-2.0	2.0
StP13																
2A	QSr.umn-2A.1	SNP126	SNP652	159.1	2.8	1.3	1.3	-2.6	-3.0	2.0	2.2	-3.0	2.5	0.0	1.2	-1.7
2B	QSr.umn-2B.3	SNP840	SNP618	325.7	3.0	2.2	5.3	3.2	0.8	-0.8	2.8	-7.2	-4.1	0.9	0.5	-2.7
2B	QSr.umn-2B.2	SNP294	SNP970	434.9	3.2	2.3	3.5	-4.9	1.7	2.7	2.7	-1.4	-2.3	0.3	-1.1	-3.0
3B	QSr.umn-3B.4	SNP89	SNP395	89.4	3.2	2.3	-5.5	-5.4	15.3	-11.3	11.8	0.0	-8.7	-10.7	0.2	-10.4
3B	QSr.umn-3B.8	SNP885	SNP790	684.5	3.6	2.6	4.5	4.5	1.2	-2.7	-3.8	-0.3	0.7	1.7	-1.1	-4.1
3B	QSr.umn-3B.7	SNP320	SNP359	811.0	3.5	2.6	1.3	7.3	-13.4	-23.7	9.2	-0.9	4.6	1.3	3.9	3.7
4B	QSr.umn-4B.1	SNP159	SNP435	396.3	2.6	1.7	0.5	-2.1	-4.2	-0.2	0.7	2.2	-1.9	3.0	-1.1	3.8
5B	QSr.umn-5B.3	SNP141	SNP765	575.9	2.9	2.1	-1.2	2.9	2.2	6.2	-1.0	-7.0	-1.0	-1.8	1.0	-0.3
5D	QSr.umn-5D.2	SNP474	SNP977	101.4	2.6	1.9	-2.1	1.6	-3.1	-0.1	3.0	0.2	-1.7	-1.7	1.5	0.9

[a] Chromosome location of the QTL

[b] QTL were named according to McIntosh et al. (2003)

[c] Position (centiMorgan) of the detected QTL peak in Chromosome ‘Chrom’

[d] Logarithm of odds scores for the QTL detected at position ‘Pos’, based on joint mapping

[e] Percentage of phenotypic variation explained by the observed QTL, based on joint mapping

[f] Estimated additive effects of resistance allele for each parent

Appendix III: QTL detected via CIM in the RIL populations LMPG-6/Ada (Table A), LMPG-6/Fahari (Table B), LMPG-6/Gem (Table C), LMPG-6/Kudu (Table D), LMPG-6/Kulungu (Table E), LMPG-6/Paka (Table F), LMPG-6/Pasa (Table G), LMPG-6/Popo (Table H), and LMPG-6/Romany (Table I).

A

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken13	4A	SNP516	SNP780	353.3	4.3	22.5	-4.0
	6A	SNP425	SNP5	221.6	2.9	16.2	-2.5
StP12	2D	SNP219	SNP116	166.8	4.0	8.1	3.3
	6D	SNP456	SNP194	75.2	3.8	4.8	2.6
StP13	1B	SNP385	SNP382	10.0	2.6	0.4	-1.0
	3B	SNP4	SNP813	713.8	3.7	3.4	-1.8
	4A	SNP891	SNP428	240.7	3.3	0.7	0.8
	4B	SNP782	SNP159	393.9	3.4	0.7	-1.2
	5D	SNP977	SNP491	105.6	3.3	7.3	2.1
	6A	SNP425	SNP5	216.6	2.7	0.0	-0.3
	6B	SNP391	SNP394	407.7	2.9	1.8	1.6

B

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
SA12	2A	SNP416	SNP767	215.8	6.1	24.1	2.0
StP12	2B	SNP294	SNP970	439.9	3.0	5.7	2.0
	4A	SNP375	SNP90	196.8	2.5	0.0	0.2
	4B	SNP858	SNP117	177.3	2.6	0.8	0.7
	5B	SNP471	SNP440	10.0	3.8	2.8	-1.6
	7B	SNP22	SNP421	350.2	3.6	3.1	-1.4
StP13	1A	SNP437	SNP323	205.5	2.8	1.0	-1.2
	2B	SNP294	SNP970	434.9	3.0	6.5	3.2
	2D	SNP707	SNP438	275.9	2.8	1.8	1.8
	3B	SNP320	SNP359	806.0	2.5	1.7	1.6
	4A	SNP375	SNP90	196.8	2.8	0.2	0.4
	5B	SNP141	SNP765	575.9	2.6	1.2	-1.5

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
StP13	6A	SNP654	SNP697	338.2	2.9	0.4	0.6
	6D	SNP60	SNP456	65.7	2.6	2.4	-2.4
	7B	SNP134	SNP175	281.5	3.2	1.3	-1.0

C

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken12	3B	SNP983	SNP982	418.8	2.8	12.0	1.9
	7B	SNP498	SNP317	116.4	2.5	11.0	8.2
StP12	3B	SNP535	SNP866	607.5	3.1	13.1	3.0
StP13	3B	SNP884	SNP252	441.2	3.5	14.5	2.9
	4A	SNP383	SNP380	690.1	3.1	3.7	2.9

D

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken13	2A	SNP148	SNP744	108.7	3.7	2.8	2.0
	3B	SNP885	SNP790	684.5	3.5	2.7	1.5
	4A	SNP371	SNP375	164.3	3.3	1.2	1.5
	4B	SNP816	SNP895	94.3	2.8	0.1	0.3
SA12	4A	SNP445	SNP402	268.2	3.1	9.8	1.5
	6B	SNP27	SNP59	226.6	2.9	9.8	1.0
	7D	SNP957	SNP269	140.5	2.5	2.8	0.6
StP12	3B	SNP770	SNP401	42.9	2.9	7.2	3.6
	4A	SNP380	-	714.6	2.6	5.0	2.0
	7B	SNP355	SNP200	164.9	2.8	5.5	2.9

E

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
StP13	6A	SNP919	SNP506	158.0	2.8	18.4	-1.9

F

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
StP12	2B	SNP797	SNP646	100.5	3.0	10.2	3.8

G

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
SA12	2A	SNP580	SNP776	205.7	4.1	8.9	1.2
	4B	SNP177	SNP203	358.8	4.7	17.4	2.2
	5B	SNP88	SNP192	850.4	3.1	1.5	0.7
StP12	6B	SNP220	SNP217	801.3	3.6	13.7	3.7
StP13	2B	SNP294	SNP970	429.9	2.7	12.0	2.3
	6B	SNP220	SNP217	806.3	5.5	22.0	3.6

H

Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken13	5B	SNP877	SNP42	728.9	3.1	13.5	-3.2
	6B	SNP500	SNP441	627.4	2.8	10.0	-1.7
SA12	5D	SNP714	SNP481	143.0	3.7	15.6	0.7
StP12	5B	SNP142	SNP784	629.2	3.1	12.9	-3.2

I

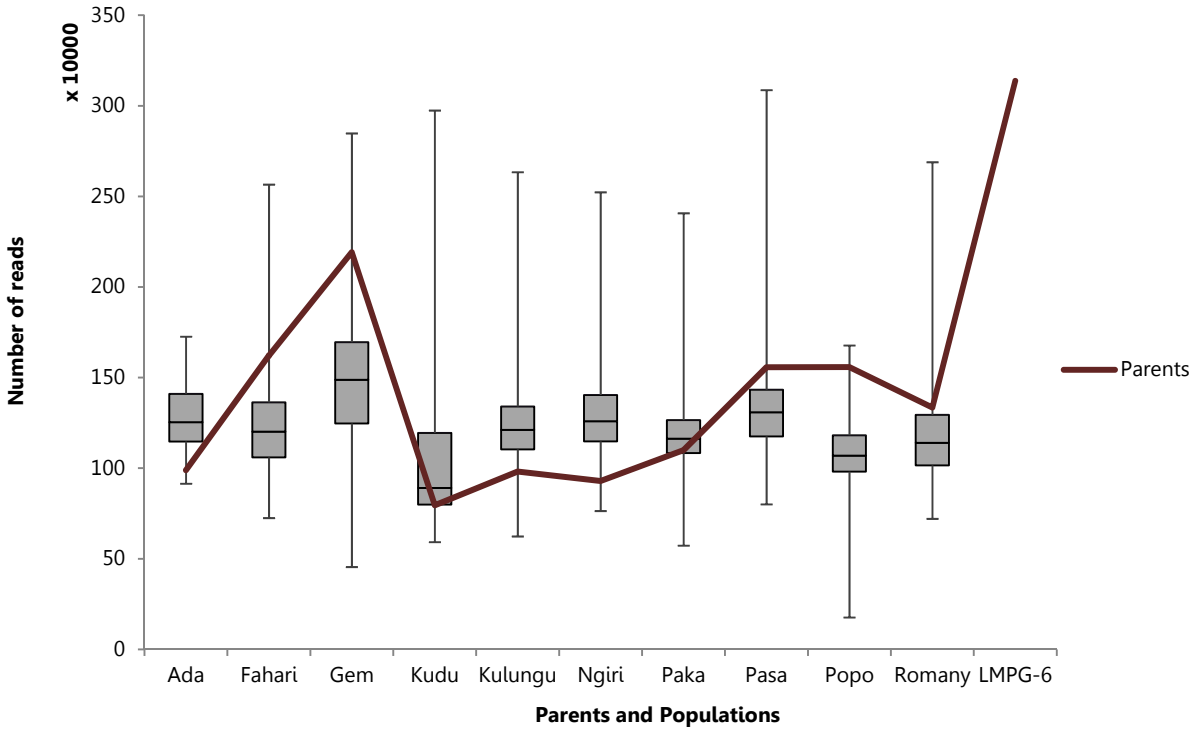
Environment	Chrom	Left SNP	Right SNP	QTL Pos (cM)	LOD[a]	R^{2b}	Add[c]
Ken13	6A	SNP935	SNP425	210.3	2.6	10.2	-1.8
	6D	SNP493	SNP147	136.7	2.9	7.7	-1.5
SA12	7B	SNP268	SNP337	43.0	7.6	24.6	4.4
StP12	3B	SNP988	SNP824	659.4	2.9	7.6	1.8
	7A	SNP599	SNP202	141.1	3.0	8.0	-2.5

[a] LOD values are the peak logarithm of odds score for the given QTL

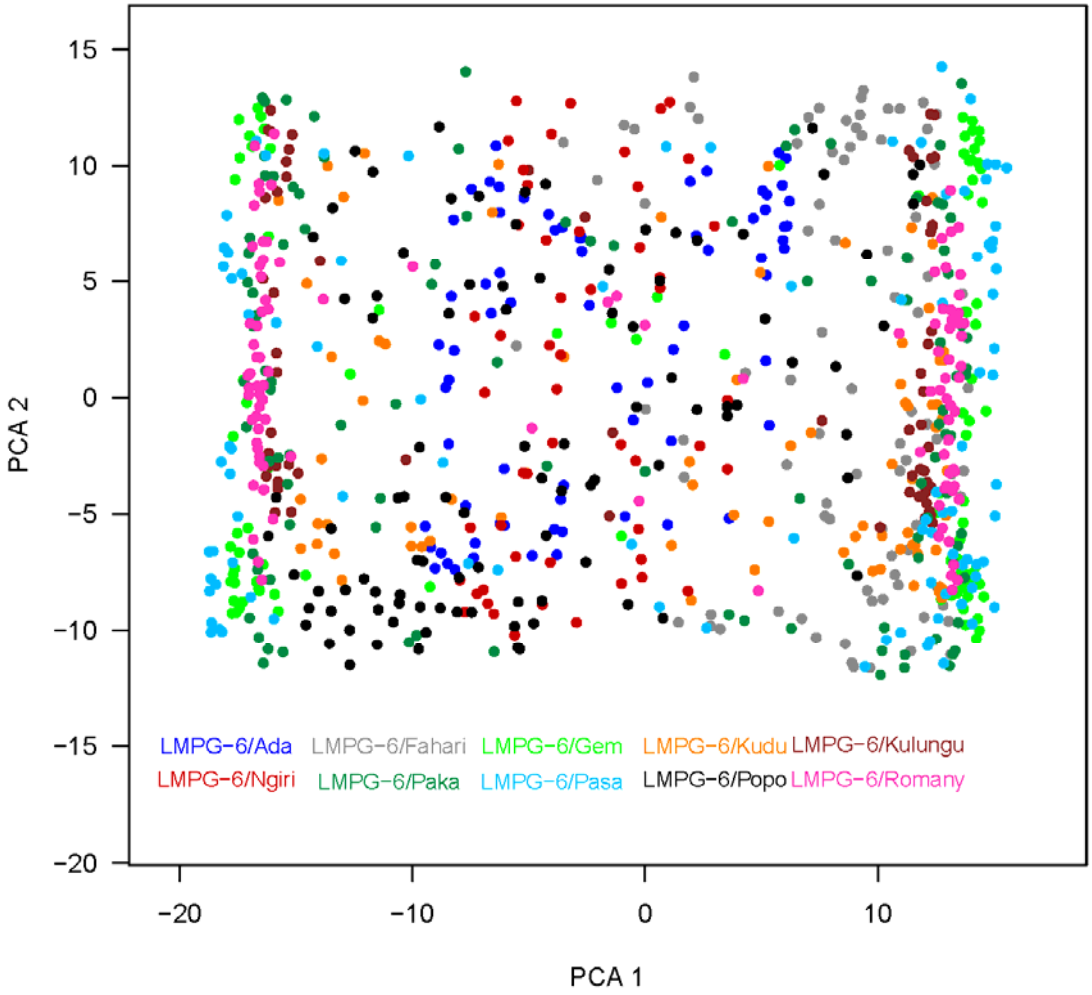
[b] Value indicates the phenotypic variation explained by the QTL

[c] Value indicates the estimated additive effect of the QTL; negative value means that the allele was contributed by RB07

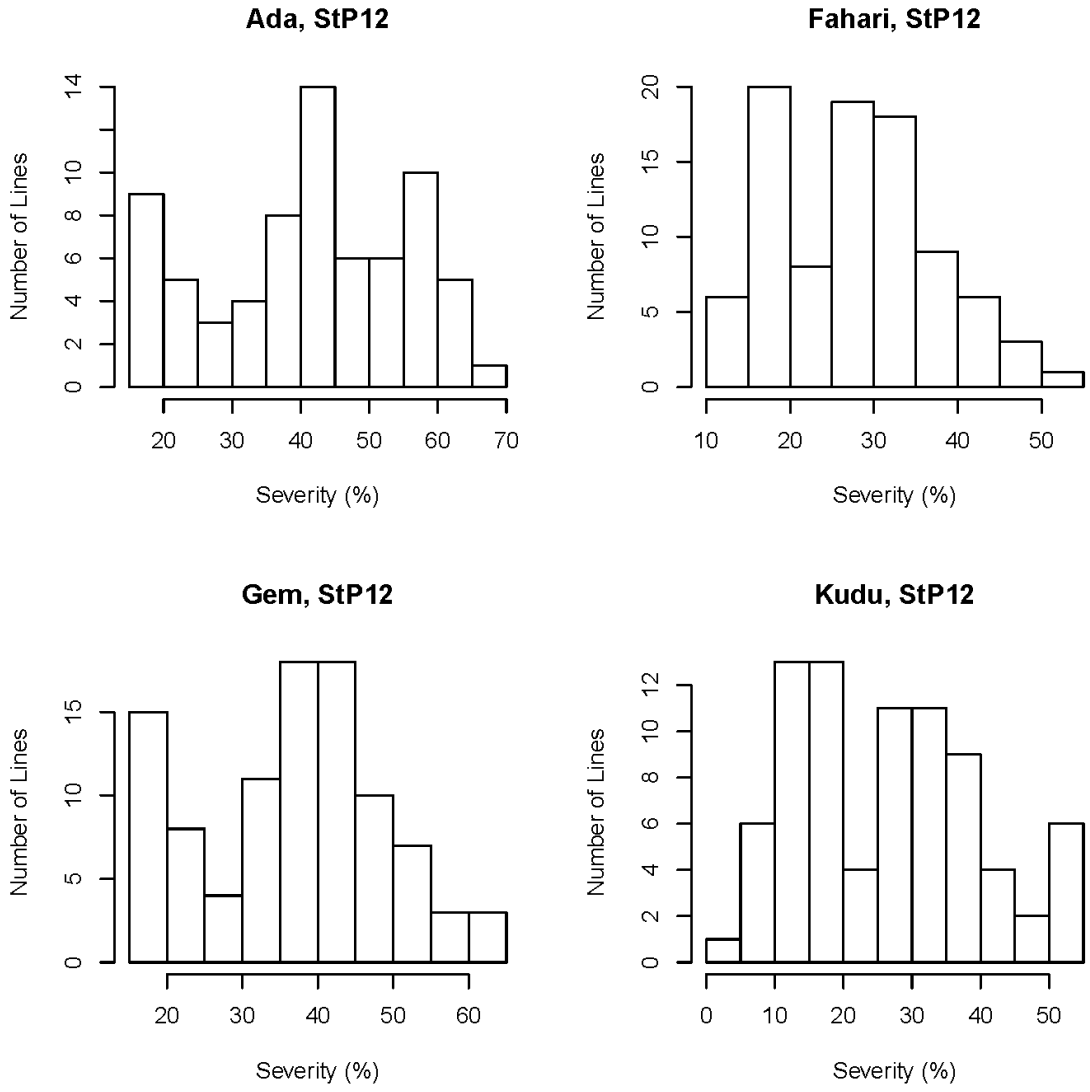
Appendix IV: Distribution of sequence reads by each population and parent (scaled down to 1/10) generated using the GBS approach.



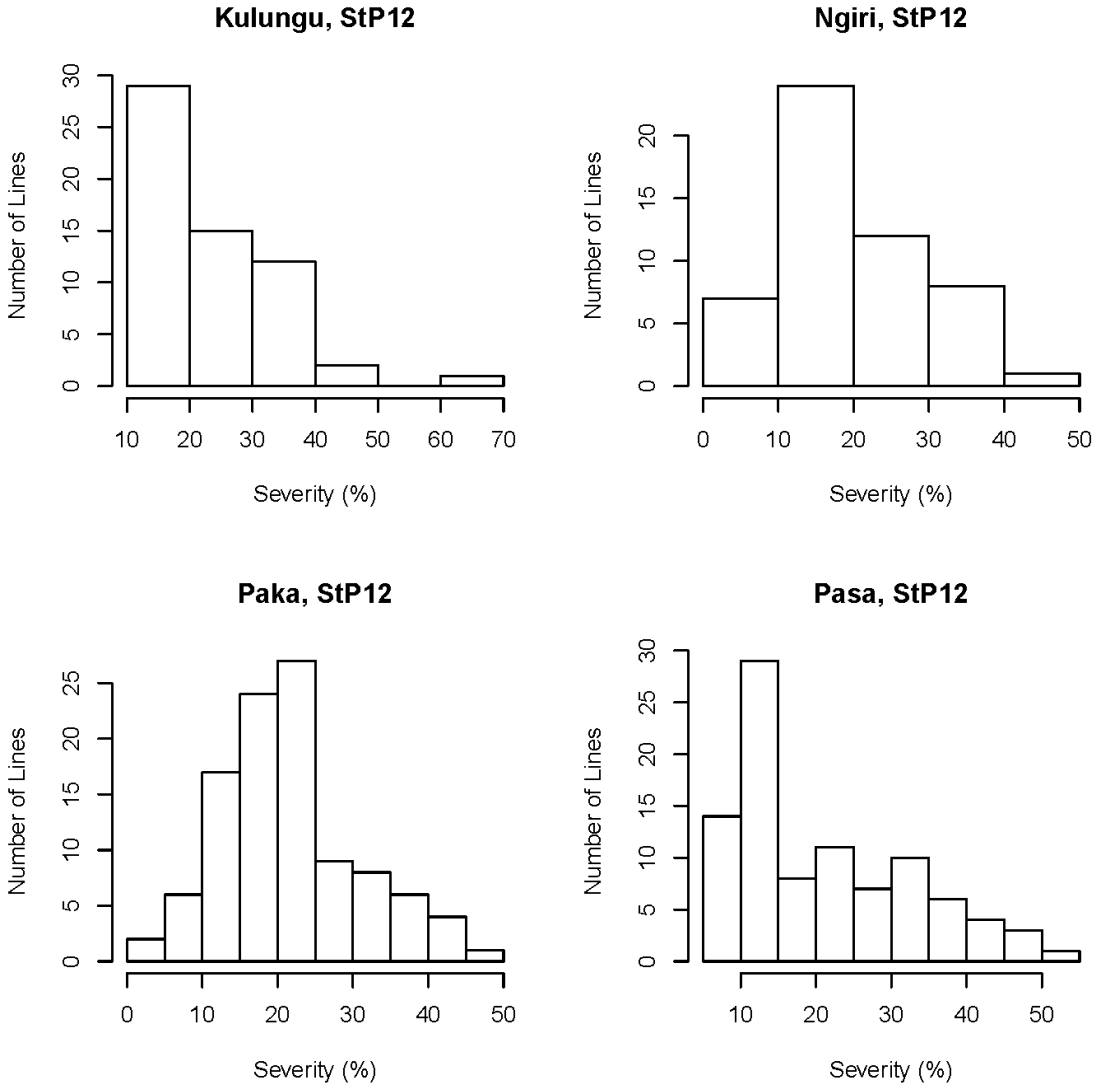
Appendix V: A scatter plot of principal component 1 (PC1) plotted against principal component 2 (PC2) of the NAM population.



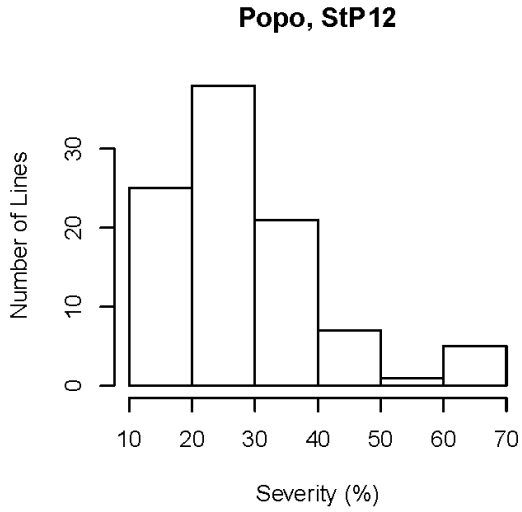
Appendix VI: Frequency distribution of stem rust severity (%) for the ten NAM populations in four environments – St. Paul 2012 (StP12), South Africa 2012 (SA12), St. Paul 2013 (StP13), Kenya 2013 (Ken13).

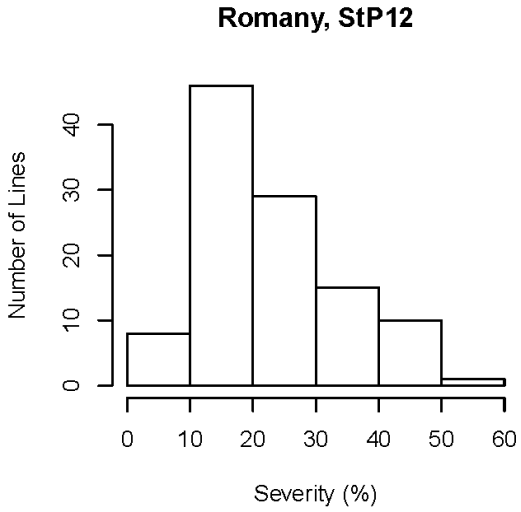


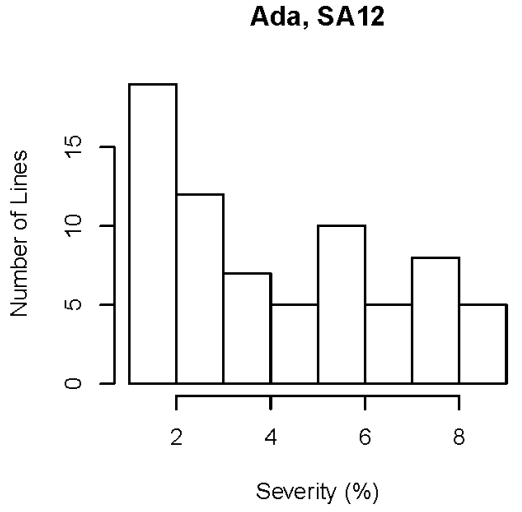
Appendix VI *continued*

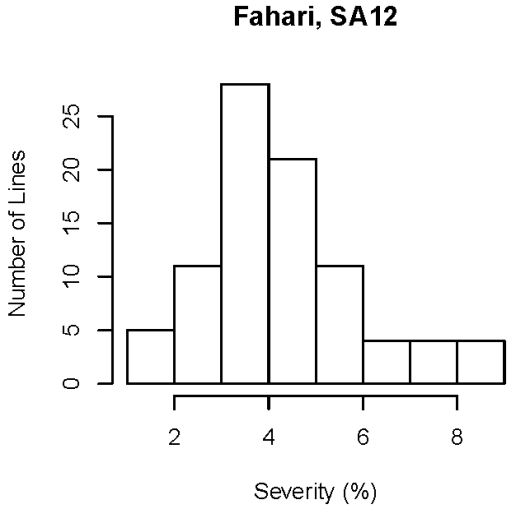


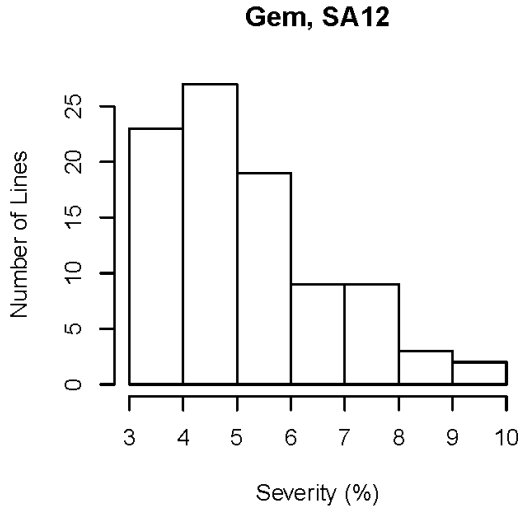
Appendix VI *continued*

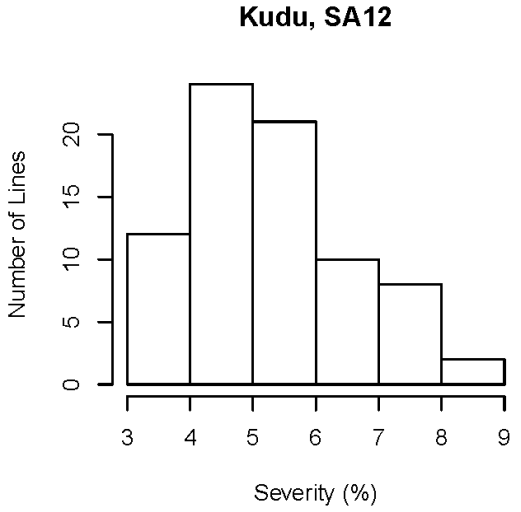




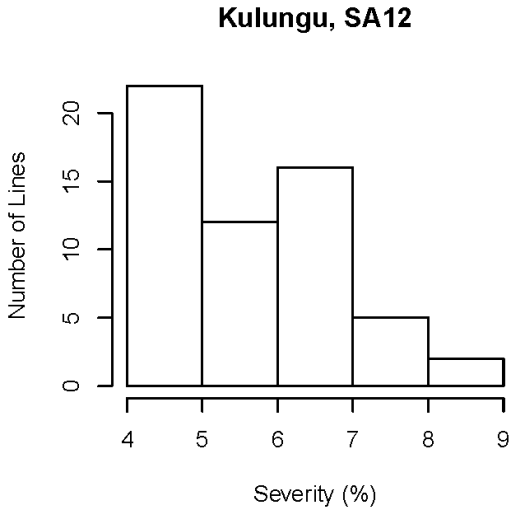


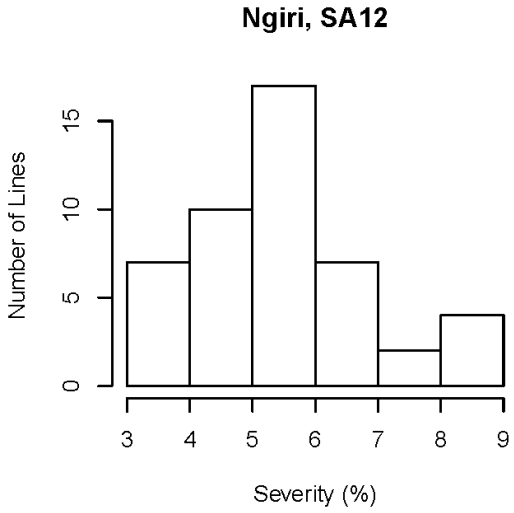


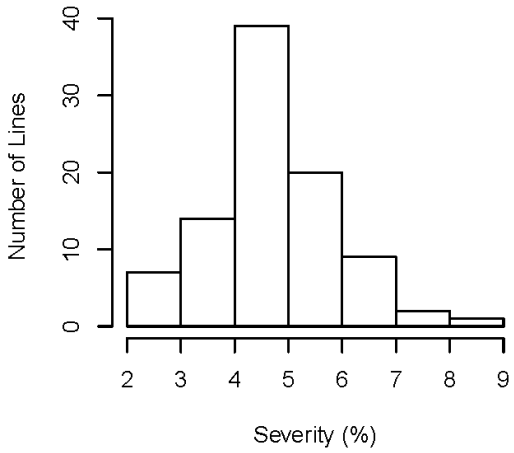


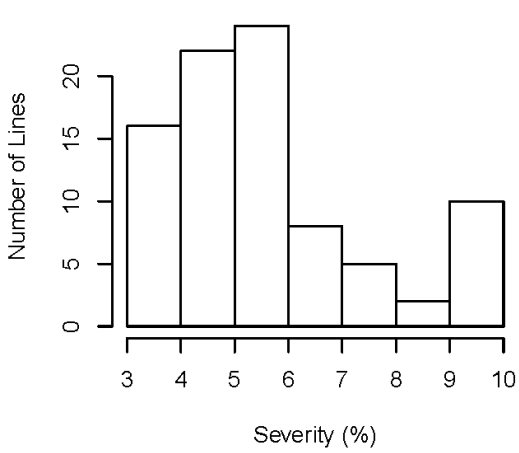


Appendix VI *continued*

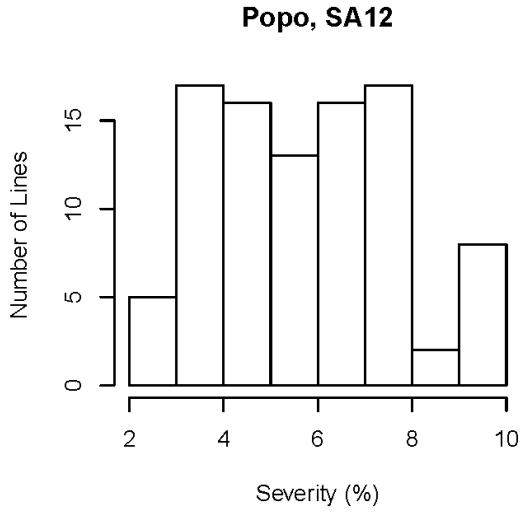


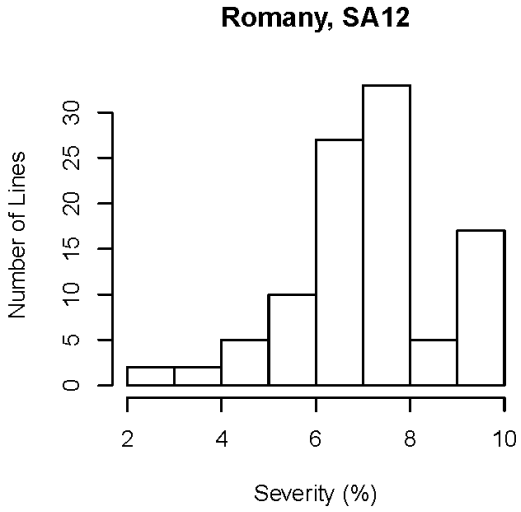


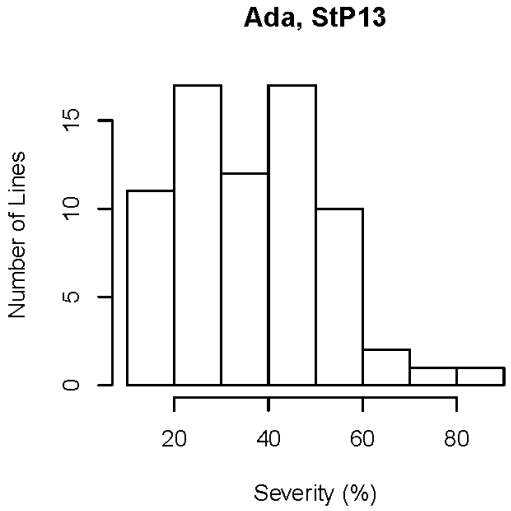


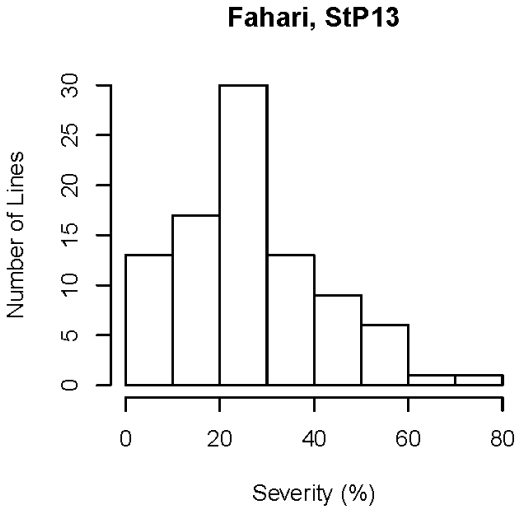


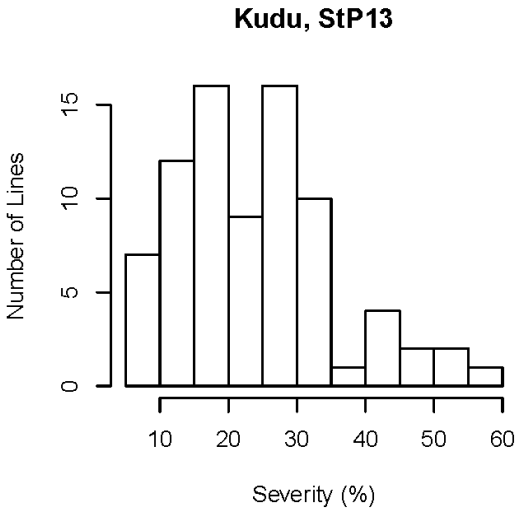
Appendix VI *continued*



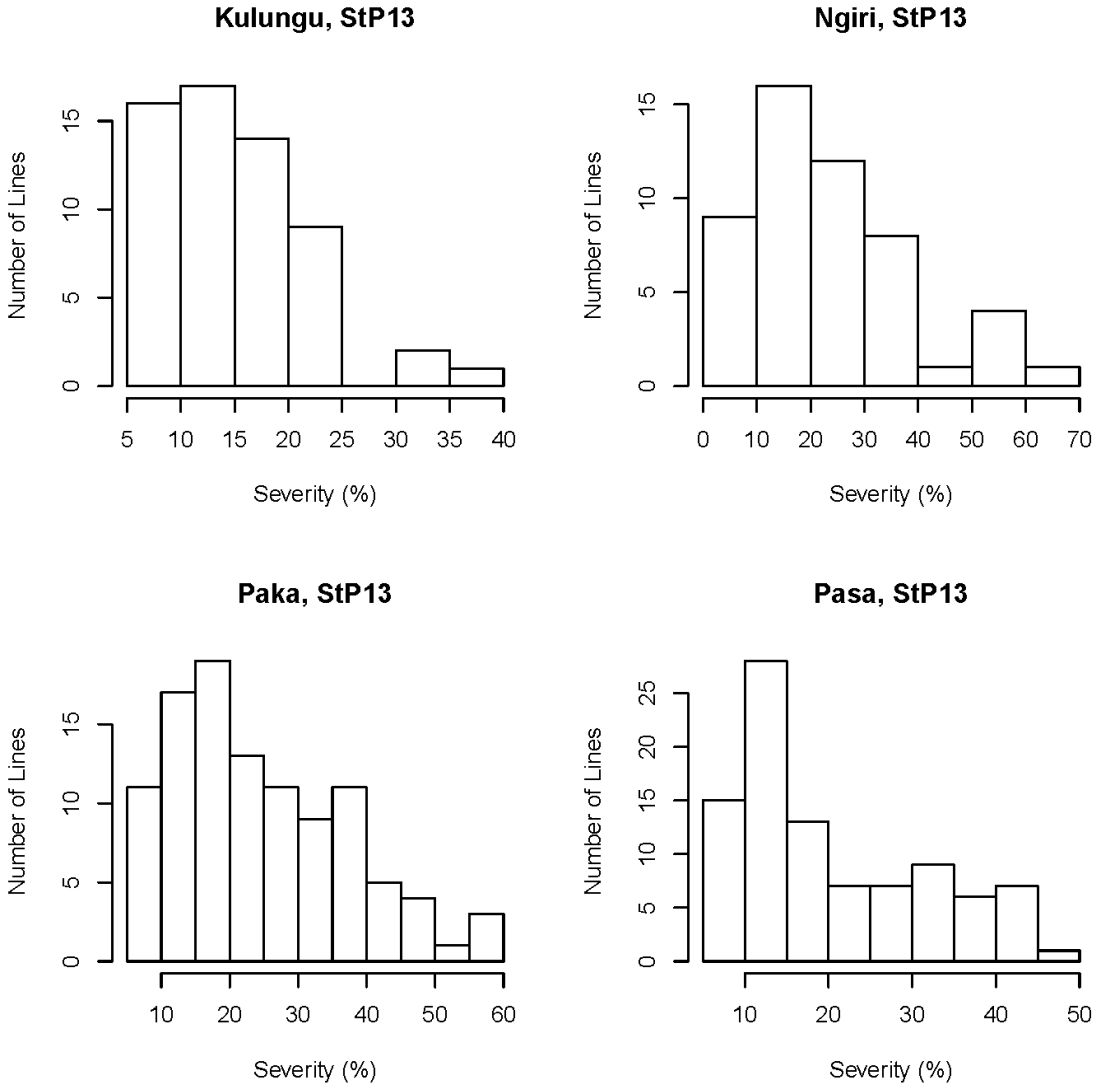




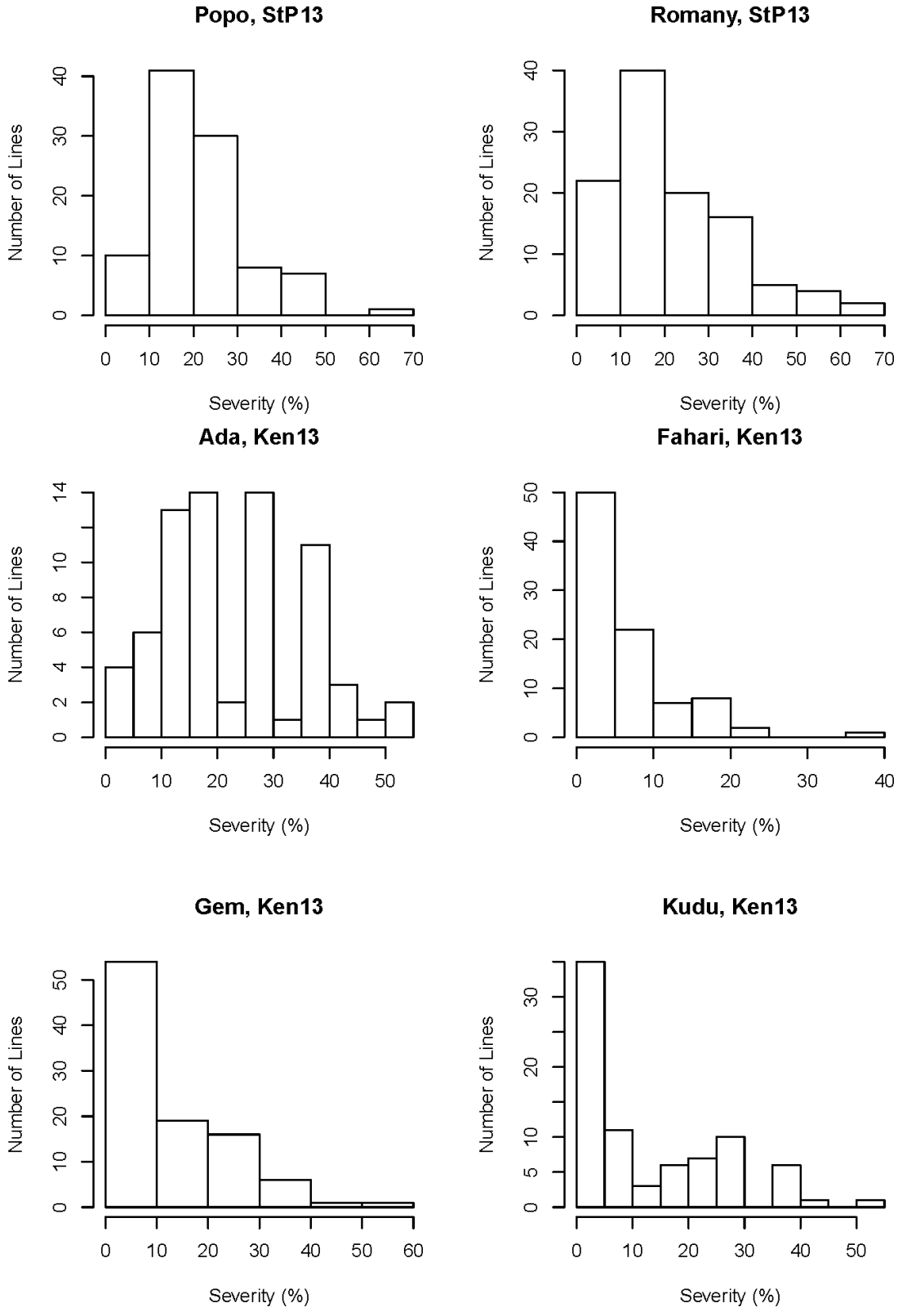




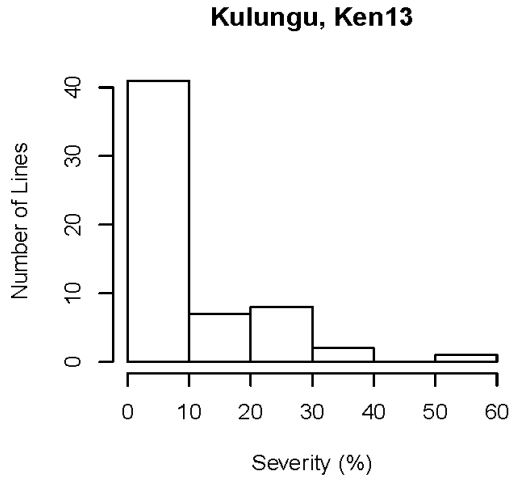
Appendix VI *continued*

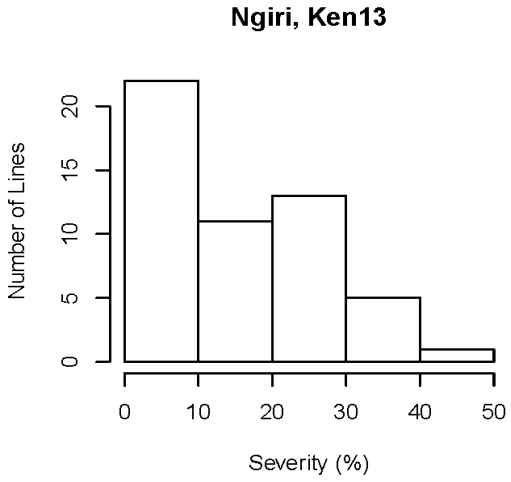


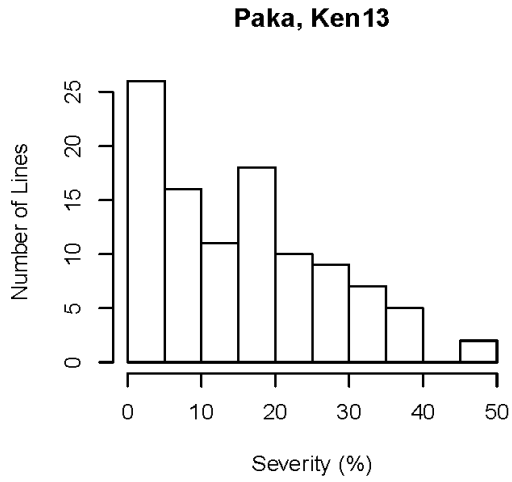
Appendix VI *continued*

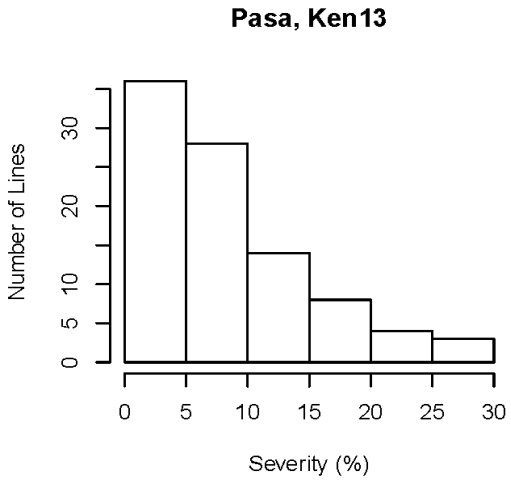


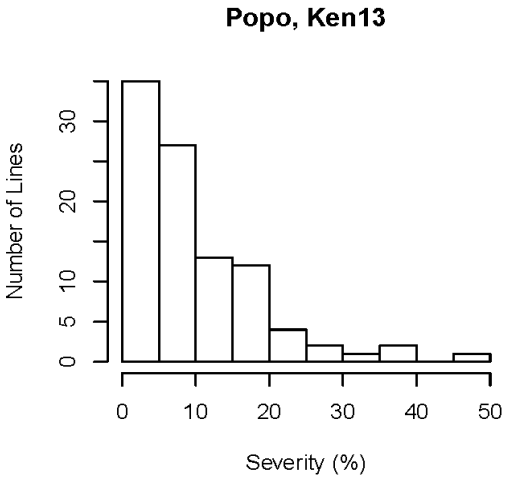
Appendix VI *continued*

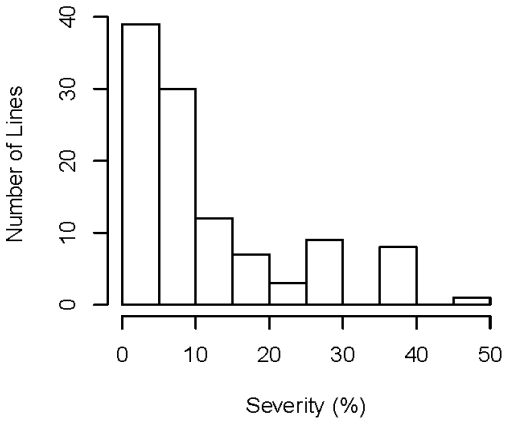




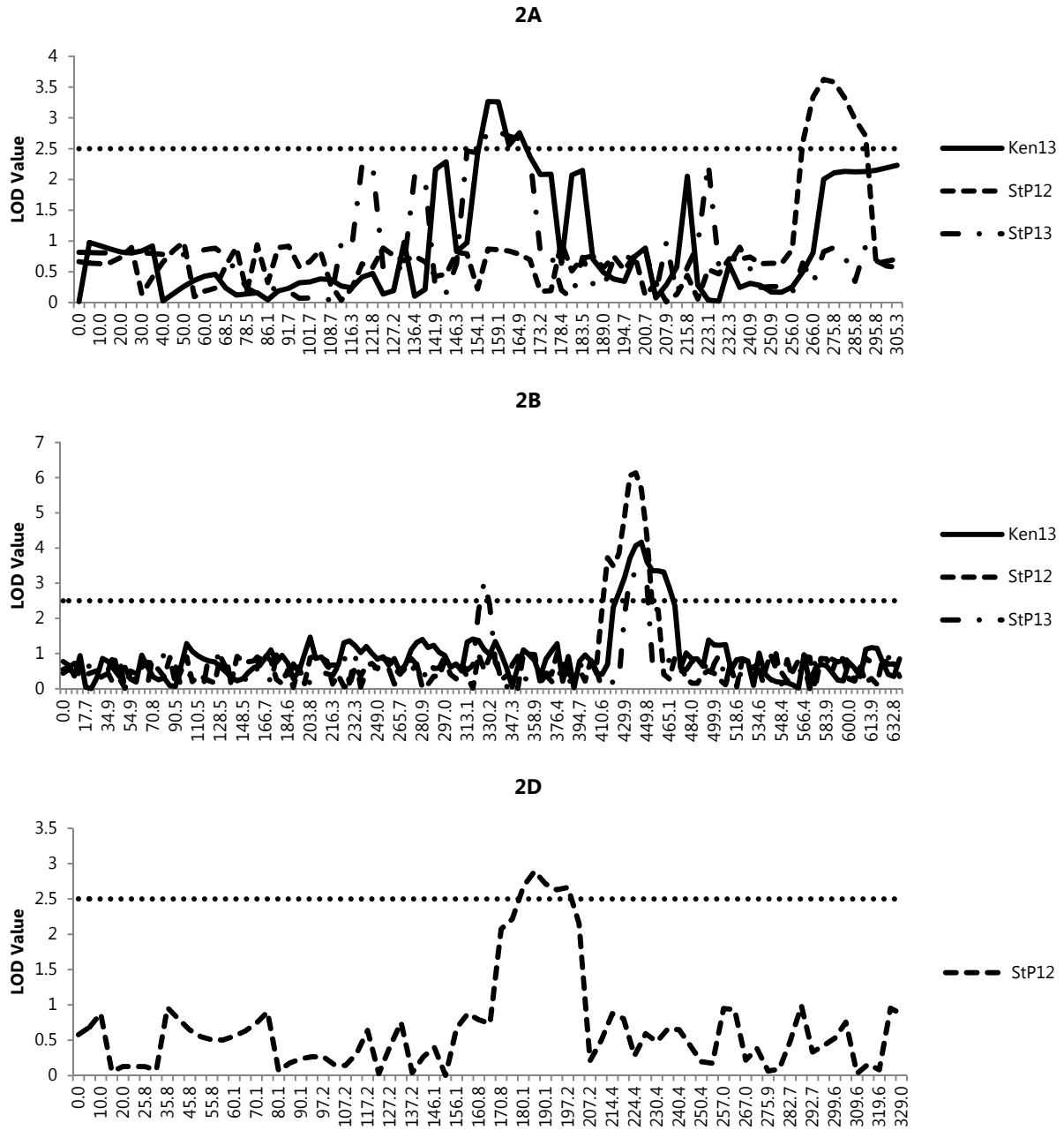




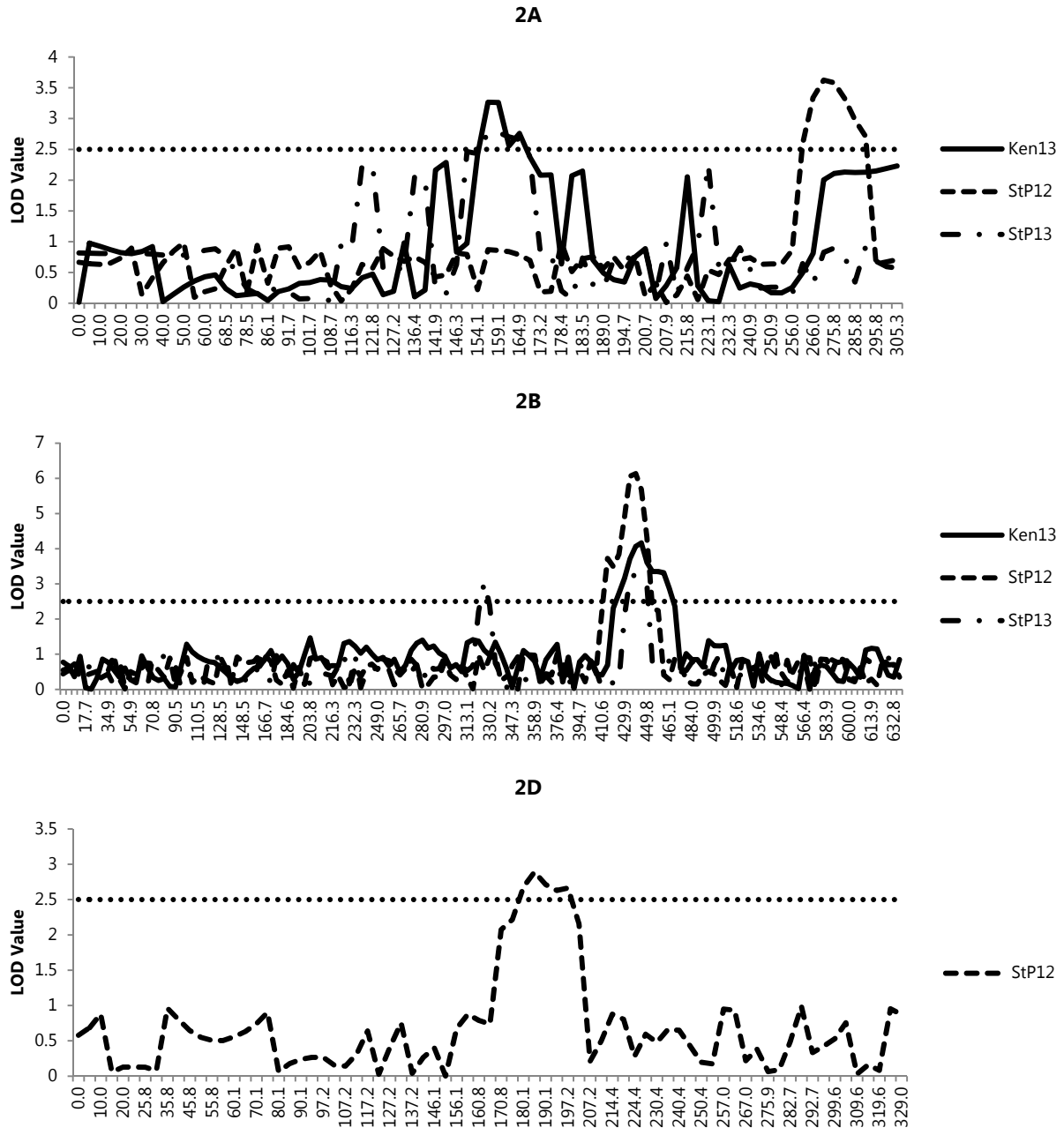




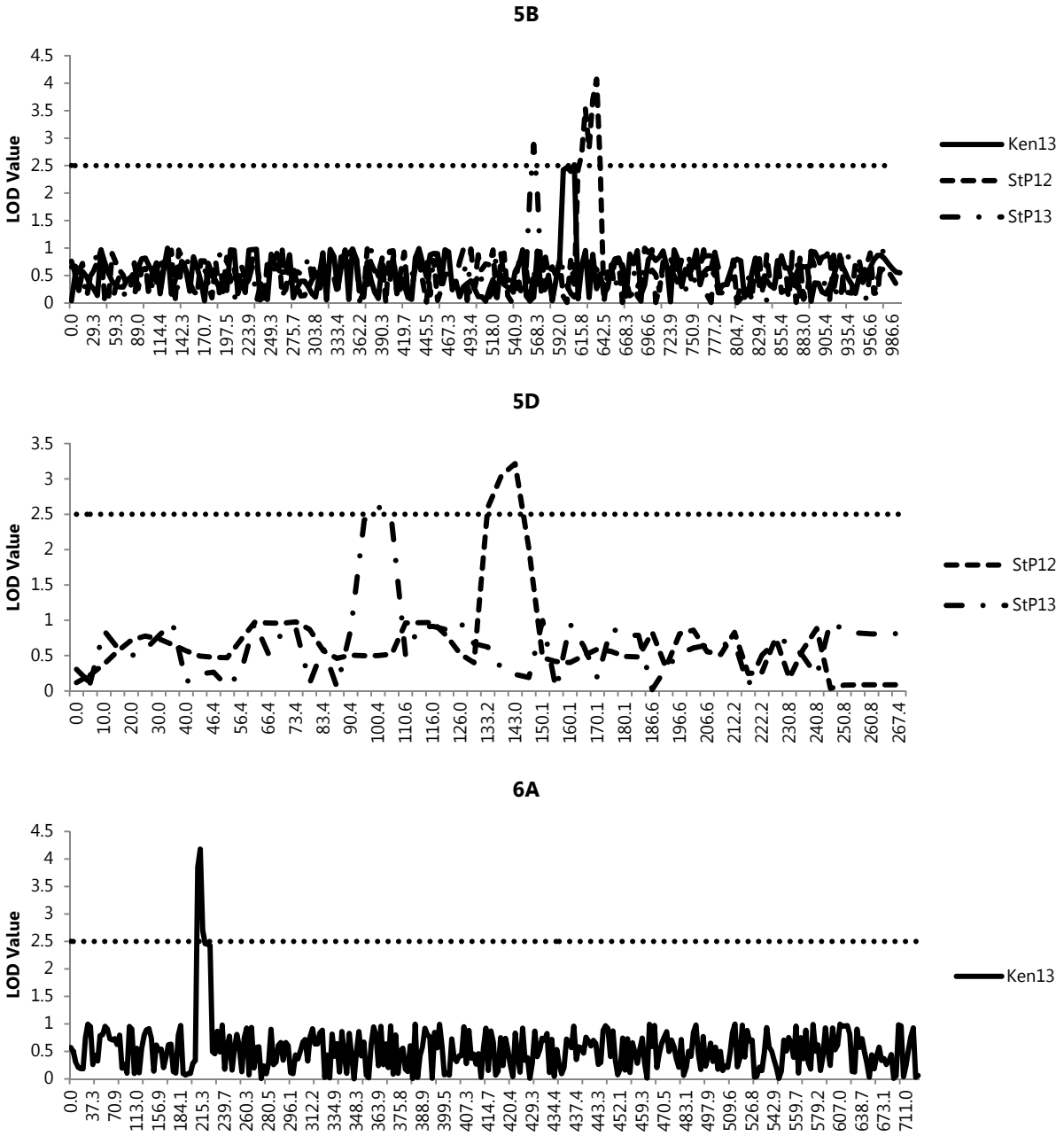
Appendix VII: Quantitative trait loci (QTL) detected in 11 chromosomes using the iQTLm method of joint mapping approach in four environments. LOD values of the QTL peaks are shown on the Y-axis; positions (cM) of the QTL are shown on X-axis. StP12, SA12, StP13, and Ken13 represent the environments St. Paul 2012, South Africa 2012, St. Paul 2013, and Kenya 2013, respectively. Chromosomes with no QTL detected are not shown. Traits with no significant QTL are not shown.



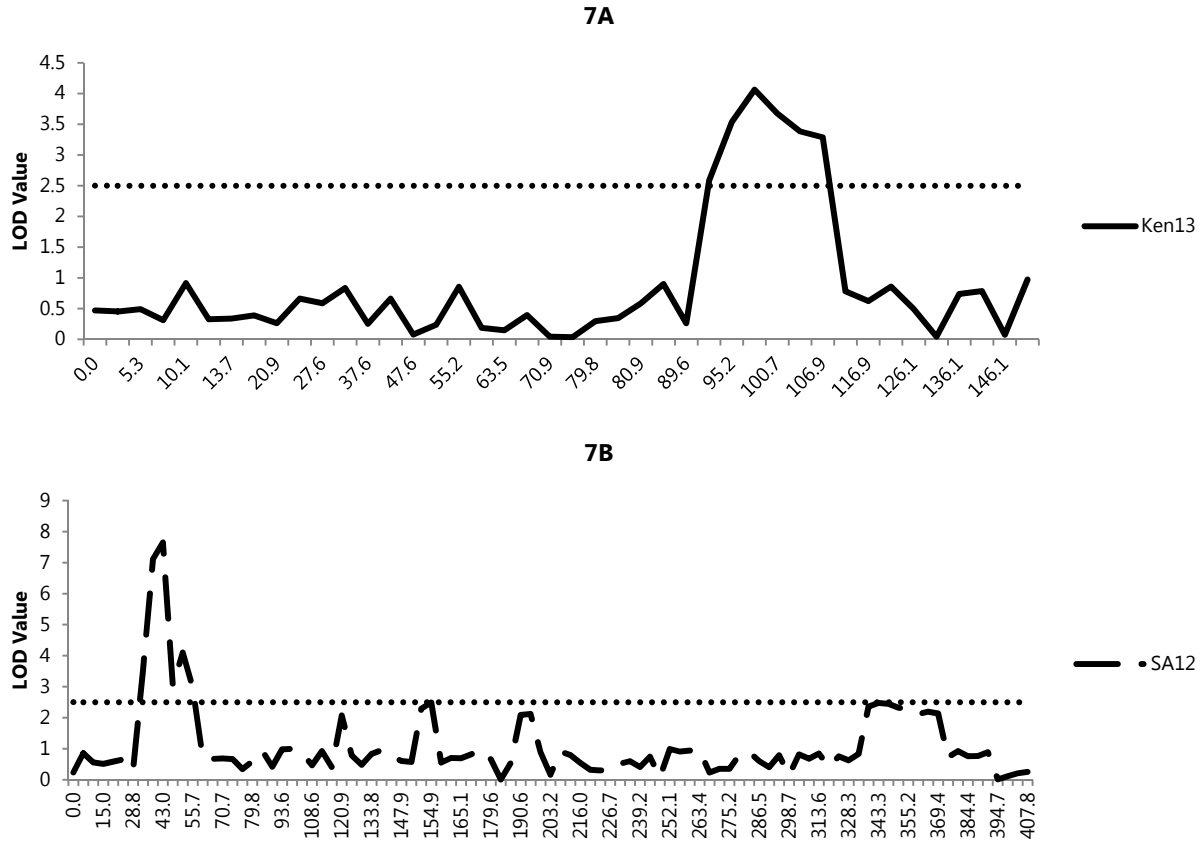
Appendix VII *continued*



Appendix VII *continued*



Appendix VII *continued*



Appendix VIII: Field reaction to stem rust observed on the RB07/MN06113-8 RIL population.

Environment	Kenya 2012		Kenya 2013		Ethiopia 2013		St. Paul 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
MN06_1	17.5	MS	-	-	-	-	25.0	MRMS
MN06_2	35.0	MSS	-	-	-	-	10.0	MRMS
MN06_4	10.0	MS	-	-	-	-	35.0	MS
MN06_5	17.5	MSMR	-	-	-	-	20.0	MRMS
MN06_6	15.0	MS	-	-	50.0	S	20.0	RMR
MN06_7	20.0	MS	-	-	-	-	35.0	MRMS
MN06_8	25.0	MSMR	-	-	-	-	15.0	MS
MN06_9	5.0	MRMS	-	-	-	-	10.0	MRMS
MN06_10	30.0	MS	35.0	MS	60.0	S	40.0	MSS
MN06_11	10.0	MRMS	25.0	MS	40.0	MSMR	35.0	MSS
MN06_12	45.0	MS	-	-	60.0	S	35.0	MSS
MN06_13	5.0	MRMS	-	-	45.0	MSS	20.0	RMR
MN06_14	25.0	MS	-	-	50.0	S	40.0	MS
MN06_15	17.5	MRMS	-	-	-	-	20.0	R
MN06_16	40.0	MSMR	-	-	60.0	S	25.0	MS
MN06_17	2.5	MS	-	-	30.0	MS	15.0	MRMS
MN06_18	47.5	MSS	-	-	-	-	10.0	MRMS
MN06_20	50.0	MS	-	-	70.0	S	40.0	MS
MN06_21	25.0	MRMS	-	-	60.0	S	20.0	RMR
MN06_22	25.0	MRMS	-	-	35.0	-	30.0	MRMS
MN06_23	17.5	MSS	45.0	MSS	50.0	S	25.0	MRMS
MN06_24	15.0	MRMS	-	-	40.0	MSMR	20.0	MRMS
MN06_25	30.0	MRMS	37.5	MSS	60.0	S	20.0	MRMS
MN06_26	10.0	MRMS	-	-	15.0	MS	15.0	MRMS
MN06_27	27.5	MRMS	-	-	60.0	S	15.0	RMR
MN06_28	20.0	MRMS	27.5	MRMS	30.0	MRMS	10.0	RMR
MN06_29	5.0	MRMS	-	-	-	-	15.0	R
MN06_30	7.5	MRMS	37.5	MSS	40.0	MSS	10.0	RMR
MN06_31	30.0	MRMS	27.5	MSMR	40.0	MRMS	15.0	MRMS
MN06_33	37.5	MRMS	-	-	60.0	S	35.0	MS
MN06_34	20.0	MRMS	22.5	MSS	45.0	MSS	15.0	MRMS
MN06_35	25.0	MRMS	-	-	60.0	S	20.0	MRMS
MN06_36	17.5	MRMS	-	-	60.0	S	30.0	MSS
MN06_37	12.5	MRMS	-	-	15.0	S	25.0	MRMS
MN06_38	17.5	MS	-	-	45.0	MSS	20.0	MRMS
Environment	Kenya 2012		Kenya 2013		Ethiopia 2013		St. Paul 2013	

Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
MN06_39	1.0	MRMS	-	-	-	-	15.0	MRMS
MN06_40	12.5	MRMS	-	-	35.0	S	10.0	R
MN06_41	42.5	MRMS	-	-	50.0	SMS	25.0	MRMS
MN06_42	47.5	MS	-	-	60.0	S	30.0	MS
MN06_43	42.5	MRMS	-	-	60.0	S	20.0	R
MN06_44	15.0	MRMS	-	-	-	-	15.0	MRMS
MN06_45	20.0	MSMR	-	-	-	-	15.0	MRMS
MN06_46	40.0	MRMS	-	-	60.0	SMS	20.0	R
MN06_47	15.0	MS	40.0	S	40.0	S	10.0	MRMS
MN06_49	50.0	S	37.5	S	55.0	SMS	25.0	R
MN06_50	27.5	MS	-	-	-	-	35.0	MS
MN06_51	30.0	MRMS	37.5	S	60.0	S	15.0	R
MN06_52	15.0	MS	17.5	MSS	40.0	MSS	15.0	MRMS
MN06_53	7.5	MRMS	-	-	25.0	MS	15.0	R
MN06_54	5.0	MRMS	35.0	MSS	50.0	MSS	10.0	MRMS
MN06_55	5.0	MS	-	-	50.0	SMS	20.0	MRMS
MN06_56	45.0	MRMS	-	-	45.0	S	20.0	MRMS
MN06_57	7.5	MRMS	27.5	MSMR	35.0	MSMR	15.0	R
MN06_58	27.5	MS	-	-	-	-	30.0	MS
MN06_59	22.5	MRMS	-	-	20.0	S	15.0	R
MN06_60	12.5	MRMS	-	-	40.0	S	10.0	MRMS
MN06_61	17.5	MRMS	15.0	MRMS	30.0	MS	10.0	MRMS
MN06_62	37.5	MRMS	30.0	MS	45.0	MSS	20.0	MRMS
MN06_63	15.0	MRMS	32.5	MSS	40.0	SMS	20.0	MRMS
MN06_64	12.5	MRMS	37.5	SMS	30.0	MSS	20.0	R
MN06_65	37.5	MRMS	37.5	MSS	-	-	35.0	MS
MN06_66	27.5	MS	20.0	SMS	45.0	MSS	20.0	MRMS
MN06_67	50.0	MSMR	40.0	SMS	55.0	SMS	25.0	MS
MN06_68	1.0	MRMS	17.5	S	40.0	S	20.0	MRMS
MN06_69	42.5	MSS	-	-	-	-	20.0	MRMS
MN06_70	35.0	MSMR	-	-	50.0	MS	15.0	MRMS
MN06_71	7.5	MRMS	-	-	40.0	MSS	15.0	MRMS
MN06_72	5.0	MRMS	12.5	MRMS	35.0	MS	20.0	MRMS
MN06_73	40.0	MRMS	45.0	S	60.0	S	25.0	MSS
MN06_74	7.5	MSS	22.5	S	25.0	S	15.0	MRMS
MN06_75	45.0	MS	42.5	SMS	70.0	S	20.0	MRMS
MN06_76	40.0	MSMR	40.0	SMS	-	-	10.0	R
MN06_78	15.0	MRMS	-	-	-	-	10.0	R
MN06_79	20.0	S	-	-	-	-	15.0	MRMS
Environment	Kenya 2012		Kenya 2013		Ethiopia 2013		St. Paul 2013	

Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
MN06_80	7.5	MRMS	27.5	MSMR	-	-	15.0	MRMS
MN06_81	42.5	MSMR	-	-	60.0	MS	15.0	R
MN06_82	37.5	MS	25.0	S	50.0	SMS	20.0	MRMS
MN06_83	37.5	MS	-	-	45.0	SMS	25.0	MS
MN06_84	32.5	MSMR	45.0	S	60.0	S	30.0	MS
MN06_85	37.5	MS	-	-	-	-	25.0	MRMS
MN06_86	37.5	MSMR	40.0	SMS	-	-	30.0	MS
MN06_87	35.0	MS	30.0	S	55.0	S	25.0	MRMS
MN06_88	7.5	MRMS	-	-	-	-	15.0	R
MN06_89	10.0	MRMS	-	-	-	-	15.0	R
MN06_91	35.0	MS	17.5	MSS	-	-	20.0	MRMS
MN06_93	32.5	MRMS	-	-	60.0	S	30.0	MS
MN06_94	35.0	MS	-	-	50.0	S	20.0	MRMS
MN06_95	7.5	MRMS	30.0	MS	45.0	S	25.0	R
MN06_96	15.0	MRMS	-	-	45.0	MSS	20.0	MRMS
MN06_97	30.0	MRMS	32.5	MS	45.0	S	30.0	MS
MN06_98	17.5	MS	37.5	SMS	60.0	S	35.0	MSS
MN06_99	37.5	MRMS	-	-	-	-	25.0	MR
MN06_100	17.5	MRMS	-	-	-	-	15.0	R
MN06_101	10.0	MRMS	-	-	-	-	20.0	MRMS
MN06_102	35.0	MS	20.0	S	50.0	S	25.0	MRMS
MN06_103	37.5	MSMR	30.0	MSMR	35.0	MS	25.0	MRMS
MN06_104	7.5	MRMS	-	-	50.0	S	20.0	R
MN06_105	7.5	MRMS	-	-	-	-	20.0	R
MN06_106	12.5	MSS	27.5	MSS	40.0	S	25.0	MS
MN06_107	22.5	S	-	-	-	-	5.0	MS
MN06_108	17.5	MRMS	-	-	-	-	20.0	MRMS
MN06_109	17.5	MSMR	27.5	MSS	45.0	MSS	15.0	MRMS
MN06_110	47.5	MRMS	40.0	MSS	45.0	S	25.0	MRMS
MN06_111	27.5	MSMR	-	-	-	-	30.0	MS
MN06_112	22.5	MSMR	-	-	-	-	20.0	MRMS
MN06_113	17.5	MS	-	-	-	-	20.0	MRMS
MN06_114	47.5	MSMR	-	-	-	-	40.0	S
MN06_115	42.5	MSMR	-	-	-	-	20.0	MRMS
MN06_116	35.0	MRMS	-	-	-	-	20.0	R
MN06_117	40.0	MRMS	-	-	-	-	25.0	MRMS
MN06_118	50.0	MS	-	-	-	-	15.0	R
MN06_119	10.0	MRMS	-	-	-	-	15.0	R
MN06_120	37.5	MSMR	-	-	60.0	S	20.0	MRMS
Environment	Kenya 2012		Kenya 2013		Ethiopia 2013		St. Paul 2013	

Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
MN06_121	22.5	MRMS	-	-	-	-	15.0	MRMS
MN06_122	20.0	MRMS	-	-	60.0	MSS	35.0	MSS
MN06_123	1.0	S	-	-	-	-	10.0	MS
MN06_124	7.5	MS	-	-	-	-	20.0	MRMS
MN06_125	5.0	MSMR	-	-	30.0	MRMS	15.0	MRMS
MN06_126	37.5	MSMR	-	-	-	-	20.0	MRMS
MN06_127	2.5	S	-	-	-	-	15.0	MRMS
MN06_128	30.0	MSMR	-	-	-	-	30.0	MS
MN06_129	2.5	MS	-	-	-	-	15.0	MS
MN06_130	5.0	MSMR	-	-	-	-	15.0	MS
MN06_131	37.5	MRMS	-	-	-	-	20.0	MRMS
MN06_132	22.5	MRMS	20.0	S	60.0	S	20.0	MRMS
MN06_133	35.0	MSMR	-	-	55.0	SMS	30.0	MS
MN06_134	35.0	MSS	-	-	-	-	15.0	MS
MN06_135	17.5	MRMS	-	-	-	-	20.0	MRMS
MN06_137	20.0	MRMS	-	-	-	-	25.0	MRMS
MN06_138	10.0	MRMS	-	-	-	-	15.0	MS
MN06_139	17.5	MRMS	-	-	-	-	15.0	MS
MN06_140	12.5	MRMS	17.5	MS	40.0	MS	10.0	MS
MN06_141	42.5	MSS	-	-	-	-	20.0	MRMS
MN06_142	30.0	MS	-	-	-	-	25.0	MRMS
MN06_143	45.0	MS	-	-	-	-	15.0	MRMS
MN06_144	35.0	MS	-	-	-	-	20.0	MRMS
MN06_145	32.5	MS	27.5	MS	50.0	MSS	20.0	MRMS
MN06_146	22.5	MR	-	-	-	-	20.0	MRMS
MN06_147	7.5	MRMS	-	-	-	-	20.0	MRMS
MN06_148	17.5	MRMS	-	-	55.0	S	15.0	MRMS
MN06_149	20.0	MRMS	-	-	60.0	S	20.0	R

[a] Disease severity (%) was measured according to modified Cobb scale (Peterson et al. 1948)

[b] Represents infection response of the plants to stem rust

Appendix IX: Field reaction to stem rust observed on the 10 NAM RIL populations.

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Ada_1	55	S	3	-	90	S	20	MSS
Ada_2	40	S	4	-	45	MSS	5	MSS
Ada_3	20	MSS	4	-	20	MSS	5	MSS
Ada_5	25	MSS	2	-	50	S	30	S
Ada_6	65	S	3	-	60	S	50	MSS
Ada_7	30	MSS	1	-	25	MSS	15	MS
Ada_8	55	MR	2	-	65	S	45	S
Ada_9	45	MRMSS	6	-	55	S	15	M
Ada_11	15	MR	3	-	30	MRMS	15	M
Ada_12	15	MSS	6	-	25	MRMS	1	R
Ada_13	30	MSS	2	-	35	MSS	30	S
Ada_14	60	S	4	-	50	MSS	40	MSS
Ada_15	60	S	2	-	45	MRMS	30	M
Ada_17	40	MSS	2	-	15	MRMS	10	MSS
Ada_18	35	MSS	2	-	40	MSS	1	R
Ada_19	45	MSS	4	-	45	MSS	30	MSS
Ada_21	55	S	2	-	75	S	30	S
Ada_22	40	MSS	8	-	65	S	10	MS
Ada_23	45	MSS	3	-	55	MSS	15	M
Ada_24	55	S	5	-	55	MSS	15	M
Ada_26	55	S	9	-	50	S	20	MSS
Ada_27	25	MSS	3	-	15	MRMS	15	MSS
Ada_28	45	S	6	-	45	MRMS	20	S
Ada_29	50	S	9	-	55	MSS	55	S
Ada_30	15	RMR	3	-	15	MRR	40	MSS
Ada_31	50	S	9	-	50	MRMS	55	S
Ada_32	60	S	4	-	40	MSS	10	M
Ada_33	60	S	5	-	50	MSS	20	M
Ada_34	35	S	8	-	35	MSS	20	MSS
Ada_35	60	S	3	-	55	S	40	S
Ada_37	45	MSS	2	-	50	MSS	30	S
Ada_38	35	S	2	-	10	MSS	20	MSS
Ada_41	50	S	3	-	50	MSS	40	MSS
Ada_42	65	S	2	-	45	MSS	20	M
Ada_43	60	S	3	-	35	MRMS	30	M
Ada_44	20	MSS	1	-	10	MSS	20	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Ada_45	25	MS	1	-	25	MRMS	30	MSS
Ada_46	20	MSS	5	-	10	MR	20	S
Ada_47	55	S	3	-	35	MRMS	40	S
Ada_48	65	S	2	-	55	MSS	40	MSS
Ada_49	45	S	5	-	55	MSS	15	M
Ada_50	60	S	6	-	50	MSS	30	MSS
Ada_51	45	S	6	-	40	S	30	MSS
Ada_52	45	MSS	6	-	40	S	15	M
Ada_54	20	M	6	-	20	MRMS	10	M
Ada_56	60	S	7	-	30	MRMS	15	M
Ada_57	45	MSS	9	-	25	MRMS	30	MSS
Ada_58	25	S	3	-	20	MSS	10	MSS
Ada_59	20	M	6	-	25	MRMS	25	MSS
Ada_61	65	S	7	-	30	MRMS	20	MSS
Ada_62	25	MSS	4	-	40	S	30	MR
Ada_63	40	MSS	3	-	25	MR	40	MSS
Ada_64	30	MSS	4	-	30	MSS	20	S
Ada_65	65	S	2	-	50	MSS	40	MSS
Ada_66	50	MSS	2	-	60	MSS	45	S
Ada_67	60	S	2	-	25	S	30	S
Ada_68	45	MSS	2	-	45	MSS	40	M
Ada_69	50	S	2	-	45	MSS	15	M
Ada_70	50	S	9	-	35	MRMS	25	S
Ada_71	45	S	8	-	20	S	15	MSS
Ada_72	35	MSS	7	-	40	MSS	15	M
Ada_73	60	S	8	-	50	S	45	S
Ada_74	40	MSS	6	-	20	MRMS	10	MSS
Ada_75	70	S	8	-	55	MSS	40	MSS
Ada_76	40	MSS	8	-	25	MSS	20	S
Ada_77	40	S	7	-	25	MSS	20	MSS
Ada_78	45	S	6	-	30	MRMS	35	M
Ada_79	45	M	5	-	30	MRMS	40	MSS
Ada_81	20	RMR	8	-	25	MR	15	M
Ada_82	40	MSS	7	-	40	S	30	MSS
Ada_84	45	S	8	-	25	RMR	20	MSS
Fahari_1	30	MSS	8	-	10	MSS	1	R
Fahari_2	20	RMR	3	-	15	RMR	1	R
Fahari_3	30	MSS	6	-	25	MSS	5	RMR
Fahari_4	35	MSS	5	-	25	MRMS	10	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Fahari_5	35	M	4	-	20	MRMS	10	MSS
Fahari_6	35	M	3	-	40	MRMS	10	M
Fahari_7	20	MR	2	-	25	MRMS	1	R
Fahari_8	10	MSS	2	-	10	RMR	1	R
Fahari_9	20	MSS	6	-	10	RMR	5	MSS
Fahari_10	30	M	7	-	10	RMR	5	MS
Fahari_12	30	RMRMS	4	-	45	MRMS	10	M
Fahari_13	40	MSS	4	-	30	MRMS	1	S
Fahari_14	40	M	6	-	20	MRMS	5	MS
Fahari_16	55	S	4	-	80	S	5	M
Fahari_18	20	RMR	5	-	10	R	1	S
Fahari_19	20	RMR	5	-	25	RMR	1	R
Fahari_22	15	RMR	7	-	10	RMR	1	S
Fahari_23	30	RMRMS	6	-	25	MRMS	5	M
Fahari_25	30	MS	2	-	35	MRMS	5	M
Fahari_26	20	RMR	4	-	20	RMR	1	R
Fahari_27	20	MSS	4	-	5	RMR	5	MS
Fahari_28	30	M	4	-	30	MRMS	5	M
Fahari_29	25	RMR	4	-	25	RMR	25	M
Fahari_30	40	MS	5	-	65	S	5	RMR
Fahari_31	20	MR	3	-	25	RMR	20	M
Fahari_32	30	M	5	-	30	MRMS	5	R
Fahari_33	20	MSS	3	-	30	MRMS	1	R
Fahari_34	30	MS	4	-	30	MRMS	20	MSS
Fahari_35	25	MS	4	-	45	MRMS	5	RMR
Fahari_36	40	MSS	5	-	35	MSS	10	MSS
Fahari_37	30	M	3	-	40	MSS	1	MS
Fahari_38	20	M	6	-	30	MRMS	10	MS
Fahari_40	20	RMR	4	-	20	RMR	20	MR
Fahari_41	25	M	6	-	25	RMR	5	M
Fahari_42	20	RMR	4	-	25	MRR	10	MSS
Fahari_43	30	MSS	5	-	5	MRMS	1	S
Fahari_44	35	MS	6	-	30	MRMS	15	MSS
Fahari_45	15	RMR	5	-	15	R	5	M
Fahari_46	40	MSS	8	-	30	MSS	5	MSS
Fahari_47	35	MSS	3	-	20	MRMS	5	MSS
Fahari_48	45	S	4	-	55	S	20	M
Fahari_49	20	MSS	4	-	10	S	5	MS
Fahari_50	30	M	6	-	40	MRMS	10	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Fahari_51	15	M	6	-	5	S	5	MS
Fahari_52	35	MSS	3	-	35	MSS	10	MSS
Fahari_53	45	S	4	-	60	MSS	15	M
Fahari_54	20	RMR	3	-	45	MRMS	10	MSS
Fahari_55	20	RMR	5	-	20	RMR	5	M
Fahari_57	25	M	8	-	15	RMR	5	M
Fahari_58	30	MS	5	-	30	MSS	1	R
Fahari_59	35	MSS	4	-	25	MRMS	1	R
Fahari_60	35	MS	5	-	40	MRMS	10	M
Fahari_61	40	M	9	-	35	MSS	15	MSS
Fahari_62	30	MSS	4	-	25	S	10	MSS
Fahari_63	35	MSS	4	-	30	MSS	10	M
Fahari_64	30	RMRMS	5	-	20	MRMS	10	M
Fahari_65	20	RMR	4	-	15	MRMS	10	M
Fahari_66	35	MSS	3	-	25	MSS	5	M
Fahari_67	20	M	7	-	30	RMR	5	M
Fahari_68	15	MSS	1	-	5	MR	1	S
Fahari_69	50	MSS	4	-	30	MSS	5	M
Fahari_70	45	MSS	9	-	60	MSS	20	MSS
Fahari_71	50	S	4	-	45	MRMS	25	M
Fahari_72	35	M	6	-	45	MRMS	5	M
Fahari_73	25	RMR	5	-	25	MRMS	15	M
Fahari_76	30	MSS	5	-	20	MRMS	0	R
Fahari_77	20	RMR	9	-	15	RMR	5	RMR
Fahari_78	35	MS	4	-	40	MRMS	1	R
Fahari_79	45	MSS	4	-	40	MRMS	15	MS
Fahari_80	40	MSS	5	-	30	MRMS	10	M
Fahari_81	35	MSS	5	-	10	S	5	MSS
Fahari_83	35	MSS	5	-	55	S	10	MSS
Fahari_85	30	M	-	-	40	MRMS	10	M
Fahari_86	35	MS	3	-	30	MRMS	10	RMR
Fahari_87	25	RMR	4	-	25	RMR	20	M
Fahari_88	40	MS	2	-	20	MSS	5	MSS
Fahari_89	50	S	4	-	50	S	15	M
Fahari_90	20	RMR	4	-	45	S	5	M
Fahari_91	35	MSS	-	-	50	S	5	M
Fahari_92	40	MS	4	-	55	S	20	M
Fahari_93	45	S	9	-	60	MSS	10	M
Fahari_94	20	RMR	5	-	20	RMR	10	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Fahari_95	15	RMR	8	-	15	RMR	10	MSS
Fahari_96	30	M	5	-	30	MRMS	20	MSS
Fahari_97	30	M	5	-	40	RMR	15	MS
Fahari_98	25	MSS	4	-	10	MSS	5	MSS
Fahari_99	45	MSS	5	-	50	MSS	40	S
Fahari_100	35	MSS	3	-	20	MSS	1	S
Fahari_101	35	M	6	-	35	MRMS	5	MS
Fahari_102	25	M	7	-	30	MRMS	1	S
Gem_1	40	MSS	9	-	40	MRMS	30	S
Gem_2	45	S	5	-	15	MRMS	5	MS
Gem_3	60	S	6	-	75	S	40	MSS
Gem_4	40	MSS	10	-	15	MRMS	1	R
Gem_6	35	MSS	8	-	30	MSS	10	MSS
Gem_7	40	S	7	-	45	S	10	MSS
Gem_8	50	MSS	8	-	50	S	20	MSS
Gem_9	20	RMR	7	-	25	MSS	1	MS
Gem_10	35	MSS	6	-	45	MSS	10	MSS
Gem_11	65	S	3	-	50	S	35	M
Gem_12	55	S	5	-	35	MSS	30	S
Gem_13	20	MSS	4	-	30	MSS	10	MSS
Gem_14	45	MSS	8	-	25	MSS	10	MS
Gem_15	40	M	5	-	60	S	15	MSS
Gem_16	40	MSS	8	-	45	S	20	MSS
Gem_17	25	M	8	-	25	MRMS	1	R
Gem_18	45	S	6	-	55	S	35	MSS
Gem_19	40	MS	5	-	40	S	10	M
Gem_20	40	MSS	5	-	40	S	10	M
Gem_21	40	MSS	5	-	30	MRMS	1	S
Gem_22	35	M	9	-	50	S	25	MSS
Gem_23	15	RMR	5	-	50	S	5	MS
Gem_24	20	RMR	-	-	35	MRMS	5	MSS
Gem_25	45	MS	5	-	45	S	30	S
Gem_26	25	M	4	-	40	MRMS	20	M
Gem_27	40	MSS	3	-	40	MRMS	10	MSS
Gem_28	40	MSS	7	-	40	MRMS	10	M
Gem_29	40	MRMSS	5	-	35	MRMS	5	M
Gem_30	50	S	3	-	45	MSS	20	MSS
Gem_31	25	M	6	-	40	MRMS	30	MSS
Gem_32	50	S	-	-	50	S	35	S

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Gem_33	65	S	6	-	45	MS	5	MSS
Gem_34	20	MSS	9	-	10	MSS	1	R
Gem_35	45	S	5	-	45	MRMS	5	MSS
Gem_36	55	S	5	-	60	S	30	MSS
Gem_37	50	S	4	-	60	S	20	MSS
Gem_39	50	S	4	-	40	MRMS	10	MSS
Gem_40	40	MSS	6	-	50	S	30	MSS
Gem_41	25	M	7	-	10	RMR	0	R
Gem_42	20	MS	6	-	20	MSS	5	MSS
Gem_43	30	MS	4	-	25	MRMS	1	MS
Gem_44	40	MSS	5	-	35	MRMS	5	MS
Gem_45	35	S	7	-	30	MSS	10	MS
Gem_46	35	S	5	-	25	MRMS	1	RMR
Gem_47	45	S	4	-	40	MSS	10	MSS
Gem_48	45	S	4	-	45	MSS	25	MSS
Gem_49	25	MS	3	-	35	MSS	1	S
Gem_50	40	S	6	-	30	MSS	10	MSS
Gem_51	30	MRMSS	10	-	50	MRMS	10	M
Gem_52	20	RMR	5	-	20	MRMS	5	M
Gem_53	50	S	6	-	50	MRMS	10	M
Gem_54	45	MSS	5	-	40	MRMS	30	MSS
Gem_55	45	MSS	-	-	45	MRMS	15	M
Gem_56	45	MSS	6	-	30	MSS	15	M
Gem_57	15	MSS	5	-	10	MRMS	5	M
Gem_58	30	M	6	-	35	MSS	1	R
Gem_59	20	M	5	-	10	MSS	1	R
Gem_60	35	S	8	-	25	MSS	10	MSS
Gem_62	40	S	8	-	40	MRMS	10	MSS
Gem_63	25	M	6	-	45	MRMS	5	M
Gem_65	45	S	5	-	45	MRMS	15	MS
Gem_66	45	S	5	-	25	S	5	MS
Gem_67	20	RMR	5	-	25	RMR	15	M
Gem_68	45	S	4	-	40	MRMS	30	MSS
Gem_69	40	MS	7	-	30	MRMS	20	MSS
Gem_70	50	MSS	-	-	40	MRMS	1	R
Gem_71	55	S	4	-	45	MRMS	20	M
Gem_72	55	MSS	8	-	45	MSS	35	MSS
Gem_73	45	S	5	-	40	MSS	30	MSS
Gem_74	40	MSS	6	-	30	MRMS	10	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Gem_75	55	S	3	-	50	S	35	MSS
Gem_77	50	S	4	-	45	MSS	10	MSS
Gem_78	55	MSS	7	-	55	S	30	M
Gem_79	40	MSS	4	-	35	MSS	30	MSS
Gem_80	25	M	4	-	25	MSS	5	MS
Gem_81	45	S	7	-	45	MSS	60	S
Gem_82	60	S	6	-	55	S	50	S
Gem_83	35	MSS	6	-	25	S	15	MSS
Gem_84	20	RMR	5	-	25	MRMS	1	R
Gem_85	35	MSS	5	-	25	MRMS	20	MSS
Gem_86	20	RMR	3	-	15	MRMS	20	MSS
Gem_87	20	RMR	4	-	30	RMR	5	MR
Gem_88	35	M	8	-	30	MRMS	10	M
Gem_90	35	MS	3	-	30	MSS	1	R
Gem_91	45	MSS	4	-	50	S	25	MSS
Gem_92	20	RMR	5	-	30	RMR	5	MR
Gem_93	45	MSS	6	-	45	S	20	M
Gem_94	35	MSS	5	-	35	MRMS	1	R
Gem_95	45	S	7	-	45	S	15	MSS
Gem_96	60	S	6	-	40	MSS	20	M
Gem_98	55	S	4	-	35	MSS	5	MS
Gem_99	65	S	6	-	45	MRMS	30	S
Gem_100	50	S	5	-	35	MRMS	10	MSS
Gem_101	25	RMR	6	-	50	S	25	MSS
Gem_102	50	S	4	-	40	MRMS	10	MSS
Gem_103	20	RMR	-	-	30	MRMS	20	MSS
Gem_104	30	M	5	-	25	MRMS	10	M
Kudu_1	15	MR	6	-	10	MRMS	10	M
Kudu_2	20	RMR	6	-	25	RMR	5	M
Kudu_3	20	RMR	6	-	20	MRMS	5	M
Kudu_4	35	M	7	-	35	MRMS	20	M
Kudu_5	40	S	6	-	35	MRMS	30	S
Kudu_6	35	MS	8	-	35	MRMS	10	MSS
Kudu_7	35	MS	9	-	35	MRMS	40	MSS
Kudu_8	20	RMR	7	-	20	RMR	5	M
Kudu_9	15	RMR	6	-	20	RMR	1	R
Kudu_10	30	M	5	-	30	MSS	5	MSS
Kudu_11	35	M	9	-	40	MSS	20	MSS
Kudu_12	45	S	8	-	35	MRMS	30	S

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Kudu_13	10	MR	7	-	10	MSS	5	RMR
Kudu_14	25	M	6	-	15	MSS	10	MSS
Kudu_15	25	M	7	-	30	MRMS	1	R
Kudu_16	40	MSS	6	-	20	MRMS	10	M
Kudu_17	35	MS	5	-	30	RMR	20	M
Kudu_19	25	M	4	-	45	S	45	S
Kudu_20	30	M	-	-	50	S	15	M
Kudu_21	20	RMR	-	-	25	MRMS	0	R
Kudu_22	15	RMR	5	-	25	MRMS	1	RMR
Kudu_23	40	MS	5	-	60	S	25	M
Kudu_24	30	MR	6	-	30	MRMS	5	RMR
Kudu_25	20	RMR	-	-	20	MRMS	5	MR
Kudu_26	15	RMR	5	-	30	MRMS	0	R
Kudu_28	35	MSS	6	-	30	MRMS	30	MSS
Kudu_29	20	RMR	5	-	10	MRMS	1	R
Kudu_30	10	RMR	6	-	10	MRMS	1	MS
Kudu_31	40	S	8	-	50	S	40	S
Kudu_32	15	RMR	8	-	10	RMR	1	R
Kudu_33	55	S	8	-	45	MSS	25	MSS
Kudu_34	30	MS	7	-	35	MSS	25	M
Kudu_35	30	MR	7	-	30	MSS	10	MSS
Kudu_37	55	S	5	-	45	S	55	S
Kudu_38	35	MSS	8	-	30	MSS	25	MSS
Kudu_39	20	MR	7	-	30	MSS	1	S
Kudu_40	20	MR	7	-	15	RMR	10	MSS
Kudu_43	10	RMR	5	-	20	MRMS	1	R
Kudu_44	25	RMR	5	-	30	MRMS	15	M
Kudu_45	45	S	5	-	55	S	30	MSS
Kudu_47	10	RMR	4	-	25	MRMS	1	R
Kudu_48	40	S	5	-	20	MRMS	30	S
Kudu_50	20	M	6	-	20	MRMS	5	M
Kudu_51	15	RMR	6	-	15	MRMS	1	R
Kudu_52	30	M	7	-	20	MRMS	40	S
Kudu_53	55	S	5	-	45	S	25	MSS
Kudu_54	35	MSS	5	-	25	MRMS	40	S
Kudu_55	30	M	4	-	30	MSS	1	R
Kudu_56	30	MS	5	-	35	MSS	15	M
Kudu_57	45	MSS	4	-	30	MRMS	40	S
Kudu_58	15	RMR	4	-	10	RMR	0	R

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Kudu_59	50	S	8	-	30	MRMS	40	S
Kudu_60	0	R	6	-	15	MRMS	5	MSS
Kudu_61	50	S	8	-	25	MSS	10	M
Kudu_62	15	RMR	5	-	5	RMR	1	R
Kudu_63	55	S	4	-	55	S	30	MSS
Kudu_64	55	S	4	-	35	MSS	25	MSS
Kudu_65	40	MSS	5	-	35	MSS	20	MSS
Kudu_66	40	MS	5	-	30	MSS	30	MSS
Kudu_67	15	MSS	6	-	25	MSS	1	R
Kudu_69	15	RMR	4	-	25	MRMS	1	R
Kudu_70	20	RMRMS	4	-	25	MRMS	5	M
Kudu_71	30	MS	5	-	20	MRMS	20	MSS
Kudu_72	15	RMR	4	-	15	RMR	5	M
Kudu_73	35	M	4	-	30	MRMS	5	M
Kudu_74	35	MS	5	-	20	RMR	10	MSS
Kudu_75	45	MSS	5	-	30	MRMS	30	MSS
Kudu_76	55	S	5	-	35	MSS	30	MSS
Kudu_77	40	S	5	-	15	MRMS	20	MSS
Kudu_79	40	S	6	-	20	MRMS	30	MSS
Kudu_80	35	MSS	6	-	15	MRMS	10	MSSS
Kudu_81	30	MS	6	-	20	MRMS	5	M
Kudu_82	10	RMR	6	-	15	MRMS	0	R
Kudu_83	30	MSS	6	-	20	MRMS	10	MSS
Kudu_84	20	M	6	-	15	MRMS	25	MSS
Kudu_85	15	M	5	-	20	MRMS	10	M
Kudu_86	15	RMR	6	-	15	RMR	5	M
Kudu_87	20	RMR	7	-	15	RMR	1	R
Kudu_88	10	RMR	3	-	15	RMR	1	R
Kudu_89	20	M	5	-	20	RMR	5	M
Kulungu_1	30	MS	4	-	15	RMR	30	MSS
Kulungu_3	15	MS	7	-	10	R	5	MS
Kulungu_4	20	MS	7	-	10	R	15	MSS
Kulungu_5	65	S	7	-	20	MSS	60	S
Kulungu_6	15	RMR	5	-	15	MRMS	5	MSS
Kulungu_7	15	RMR	5	-	15	MRMS	10	MSS
Kulungu_8	20	S	6	-	20	MRMS	1	S
Kulungu_9	30	M	7	-	15	RMR	10	M
Kulungu_10	40	MSS	6	-	20	MSS	40	S
Kulungu_12	20	M	7	-	10	RMR	1	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Kulungu_13	15	M	7	-	10	RMR	5	MS
Kulungu_15	35	MS	6	-	20	MSS	20	MSS
Kulungu_17	30	MS	7	-	25	MSS	10	MSS
Kulungu_18	15	MSS	6	-	15	MSS	5	MS
Kulungu_19	25	MSS	5	-	20	MRMS	5	M
Kulungu_20	20	MSS	7	-	20	MSS	1	MR
Kulungu_21	30	M	7	-	20	MRMS	20	MSS
Kulungu_22	10	MS	6	-	5	R	10	M
Kulungu_23	30	MS	-	-	20	MRMS	25	MSS
Kulungu_24	15	MS	5	-	5	MS	5	M
Kulungu_25	30	MSS	8	-	15	MSS	30	MSS
Kulungu_26	20	M	5	-	10	MR	5	M
Kulungu_27	35	S	5	-	25	MSS	10	MSS
Kulungu_28	20	M	8	-	15	MRMS	1	S
Kulungu_29	30	MSS	4	-	15	MRMS	5	M
Kulungu_30	35	MSS	8	-	15	MSS	5	M
Kulungu_31	25	M	5	-	20	MRMS	5	M
Kulungu_32	20	M	6	-	20	RMR	10	MSS
Kulungu_35	35	MSS	4	-	15	MRMS	10	MSS
Kulungu_36	15	M	5	-	10	RMR	5	MSS
Kulungu_37	15	MS	7	-	10	RMR	0	R
Kulungu_39	15	M	7	-	15	MSS	5	M
Kulungu_42	35	M	7	-	25	S	30	S
Kulungu_43	20	MS	5	-	10	RMR	20	MSS
Kulungu_44	40	S	6	-	25	MSS	5	M
Kulungu_45	20	S	4	-	15	MRMS	20	MSS
Kulungu_46	20	M	6	-	10	RMR	5	M
Kulungu_47	30	S	8	-	15	MRMS	40	S
Kulungu_48	10	MS	5	-	15	MRMS	1	R
Kulungu_49	20	RMR	9	-	15	RMR	15	M
Kulungu_50	25	MS	8	-	25	MSS	5	MS
Kulungu_51	30	M	6	-	10	RMR	5	MSS
Kulungu_52	40	S	5	-	25	MRMS	5	M
Kulungu_53	40	S	7	-	20	MSS	20	M
Kulungu_54	20	M	7	-	10	RMR	10	M
Kulungu_55	15	M	9	-	10	RMR	0	R
Kulungu_56	30	MS	7	-	20	MRMS	5	M
Kulungu_57	50	S	5	-	25	MSS	30	S
Kulungu_58	40	S	5	-	20	MRMS	10	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Kulungu_59	25	M	7	-	15	MRMS	10	M
Kulungu_62	15	MSS	6	-	10	RMR	5	MR
Kulungu_65	15	M	5	-	25	MRMS	5	M
Kulungu_66	15	RMR	4	-	25	MRMS	1	R
Kulungu_67	45	S	6	-	35	MRMS	5	S
Kulungu_68	20	M	5	-	15	RMR	5	M
Kulungu_69	20	RMR	4	-	20	RMR	5	M
Kulungu_70	40	S	6	-	40	MSS	30	MS
Kulungu_71	35	MS	5	-	35	MRMS	30	MSS
Kulungu_72	30	M	-	-	10	RMR	25	MSS
Ngiri_2	20	M	4	-	25	RMR	5	RMR
Ngiri_3	15	M	6	-	15	RMR	0	R
Ngiri_4	0	R	5	-	10	R	5	MSS
Ngiri_7	20	RMR	4	-	25	RMR	5	M
Ngiri_8	15	MR	6	-	25	RMR	5	M
Ngiri_9	15	RMR	5	-	20	RMR	20	MSS
Ngiri_10	20	MS	6	-	15	RMR	5	MSS
Ngiri_11	25	RMR	6	-	25	RMR	15	S
Ngiri_12	40	MS	-	-	55	S	50	S
Ngiri_13	20	MR	3	-	30	MRMS	20	MSS
Ngiri_14	25	M	4	-	25	MRMS	40	S
Ngiri_15	20	MR	7	-	20	MRMS	10	MSS
Ngiri_16	30	MS	7	-	40	MRMS	10	M
Ngiri_17	15	RMR	5	-	15	RMR	1	R
Ngiri_18	15	RMR	6	-	15	RMR	10	MS
Ngiri_19	25	M	6	-	35	MRMS	30	MSS
Ngiri_21	10	MSS	5	-	10	MRMS	1	R
Ngiri_23	0	M	6	-	10	RMR	5	MSS
Ngiri_25	20	MR	-	-	20	RMR	5	R
Ngiri_26	20	MR	7	-	20	RMR	5	M
Ngiri_27	35	MSS	9	-	20	MRMS	35	S
Ngiri_28	20	RMR	9	-	10	MR	5	MSS
Ngiri_29	35	MS	8	-	50	S	15	MSS
Ngiri_30	20	M	9	-	25	MRMS	20	S
Ngiri_32	20	M	7	-	15	RMR	5	MSS
Ngiri_33	20	RMR	6	-	35	MRMS	15	M
Ngiri_34	20	MR	6	-	20	RMR	30	MSS
Ngiri_35	0	R	6	-	10	RMR	25	MSS
Ngiri_36	30	MS	-	-	60	S	30	S

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Ngiri_38	35	MSS	5	-	35	MRMS	30	S
Ngiri_39	35	MSS	5	-	60	S	20	M
Ngiri_40	30	MSS	7	-	40	MRMS	20	MSS
Ngiri_42	0	R	5	-	5	R	5	MSS
Ngiri_43	50	S	6	-	35	MSS	40	S
Ngiri_44	10	MSS	6	-	5	R	15	MSS
Ngiri_45	15	M	7	-	10	RMR	30	MSS
Ngiri_46	40	S	7	-	65	S	30	MSS
Ngiri_47	40	S	8	-	55	S	30	S
Ngiri_48	15	M	9	-	25	RMR	20	M
Ngiri_49	20	M	6	-	20	RMR	30	M
Ngiri_50	25	MRMSS	6	-	30	RMR	15	MSS
Ngiri_51	15	RMR	6	-	20	RMR	10	MSS
Ngiri_52	25	M	6	-	20	MRMS	25	MSS
Ngiri_53	20	RMR	5	-	10	RMR	10	MSS
Ngiri_54	15	MSS	4	-	-	-	1	R
Ngiri_55	30	MSS	3	-	15	MSS	10	MSS
Ngiri_57	20	M	5	-	25	MSS	30	S
Ngiri_58	35	MS	4	-	30	MRMS	40	S
Ngiri_59	10	M	6	-	15	RMR	1	R
Ngiri_60	30	MSS	5	-	40	MRMS	25	MS
Ngiri_61	25	M	-	-	40	MRMS	30	MS
Ngiri_64	25	MRMSS	-	-	30	RMR	35	MSS
Paka_2	25	MRMSS	6	-	10	MSS	5	M
Paka_3	45	S	5	-	25	MSS	10	MSS
Paka_4	20	RMR	3	-	30	RMR	10	M
Paka_5	45	S	-	-	60	S	30	S
Paka_6	25	MRMSS	5	-	25	RMR	10	MSS
Paka_7	25	MSS	6	-	25	MSS	1	M
Paka_8	20	MR	5	-	15	MRMS	30	S
Paka_9	20	MR	6	-	15	MRMS	1	R
Paka_10	35	MS	6	-	35	MRMS	15	MSS
Paka_12	25	MSS	5	-	60	S	25	MS
Paka_14	25	M	7	-	15	RMR	20	M
Paka_15	25	MS	5	-	15	RMR	10	MSS
Paka_16	15	M	2	-	20	RMR	5	MSS
Paka_17	20	M	4	-	20	RMR	15	M
Paka_18	35	MS	5	-	45	MSS	20	M
Paka_19	25	MS	5	-	15	RMR	15	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Paka_20	15	M	4	-	15	RMR	10	MR
Paka_22	20	M	6	-	15	RMR	5	MSS
Paka_23	15	MR	5	-	25	MRMS	1	R
Paka_24	15	MR	6	-	20	MRMS	5	MS
Paka_25	20	MR	5	-	25	MRMS	35	MS
Paka_26	25	M	4	-	20	RMR	50	S
Paka_27	25	MRMSS	5	-	10	RMR	25	S
Paka_28	20	MR	5	-	10	RMR	15	M
Paka_29	25	M	6	-	60	S	20	MR
Paka_30	25	M	-	-	35	MRMS	30	S
Paka_31	25	MS	6	-	20	RMR	5	MS
Paka_32	15	M	9	-	10	RMR	20	MSS
Paka_33	25	M	7	-	45	MRMS	20	M
Paka_34	20	M	7	-	15	RMR	20	M
Paka_35	30	MSS	6	-	35	MSS	30	MSS
Paka_36	15	MS	7	-	15	MRMS	5	MS
Paka_37	15	MR	7	-	25	RMR	15	M
Paka_38	10	MR	6	-	20	RMR	5	M
Paka_39	10	MR	5	-	15	RMR	1	R
Paka_40	15	M	6	-	25	RMR	10	RMR
Paka_41	25	MSS	6	-	50	MSS	15	M
Paka_42	15	MR	6	-	25	RMR	10	M
Paka_43	20	MR	5	-	20	RMR	20	MSS
Paka_44	50	S	5	-	30	MSS	25	MSS
Paka_45	25	MRMSS	-	-	50	S	5	RMR
Paka_47	20	MR	-	-	20	RMR	20	MSS
Paka_48	10	MR	-	-	20	RMR	5	M
Paka_49	20	MR	-	-	20	RMR	15	M
Paka_50	15	M	4	-	15	RMR	5	MSS
Paka_51	40	MS	-	-	30	MRMS	20	MS
Paka_52	15	M	4	-	10	RMR	5	MSS
Paka_53	25	M	5	-	45	MRMS	35	S
Paka_55	35	MSS	4	-	30	MSS	1	RMR
Paka_56	10	MR	6	-	10	RMR	1	S
Paka_57	20	M	5	-	20	MRMS	10	MSS
Paka_58	40	S	5	-	25	MSS	20	M
Paka_59	30	MSS	5	-	15	MRMS	10	MSS
Paka_60	20	M	6	-	30	RMR	5	M
Paka_61	20	M	5	-	20	RMR	5	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Paka_62	25	M	5	-	30	RMR	15	M
Paka_63	25	M	6	-	40	MRMS	20	MSS
Paka_64	40	S	5	-	40	MRMS	20	MS
Paka_65	20	M	6	-	15	MS	10	M
Paka_66	20	M	-	-	35	RMR	50	S
Paka_67	20	M	7	-	40	MS	15	MSS
Paka_68	20	M	6	-	20	RMR	5	RMR
Paka_69	15	MR	5	-	15	RMR	5	MS
Paka_70	30	M	5	-	40	RMR	30	S
Paka_71	20	MR	4	-	20	MSS	10	M
Paka_72	15	MR	5	-	25	RMR	10	MSS
Paka_73	25	MR	5	-	15	MRMS	5	M
Paka_74	0	R	4	-	5	RMR	5	M
Paka_75	35	MSS	4	-	35	MSS	10	M
Paka_76	25	M	-	-	40	MRMS	25	M
Paka_77	35	MSS	3	-	20	MRMS	25	MSS
Paka_78	20	M	6	-	50	MR	15	MSS
Paka_79	15	MR	7	-	30	MRMS	10	MS
Paka_80	30	MSS	4	-	25	MRMS	10	M
Paka_81	30	M	5	-	40	MRMS	20	MS
Paka_82	10	MSS	3	-	40	RMR	20	MSS
Paka_83	15	MR	4	-	20	RMR	5	M
Paka_84	35	M	5	-	40	MRMS	30	MSS
Paka_85	25	M	5	-	45	MSS	35	MSS
Paka_86	20	M	5	-	35	MRMS	20	MS
Paka_87	15	MR	4	-	20	MRMS	40	S
Paka_88	0	R	3	-	5	MSS	5	M
Paka_89	20	M	5	-	15	MSS	25	MSS
Paka_90	25	MSS	4	-	20	MRMS	30	S
Paka_92	15	MR	3	-	10	RMR	10	M
Paka_93	35	MSS	5	-	30	MSS	30	MSS
Paka_94	25	MSS	5	-	25	MRMS	35	S
Paka_95	20	MSS	-	-	35	MRMS	20	M
Paka_96	45	S	5	-	30	MSS	40	S
Paka_97	25	MR	-	-	30	MRMS	35	MSS
Paka_98	45	S	-	-	40	MRMS	35	S
Paka_99	30	MSS	8	-	55	MSS	40	MSS
Paka_103	25	M	5	-	20	MRMS	20	MSS
Paka_104	10	MR	3	-	10	MRMS	5	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Paka_105	30	MSS	5	-	40	MSS	35	S
Paka_107	40	S	5	-	50	MSS	40	S
Paka_108	40	S	5	-	45	MSS	25	MSS
Paka_109	35	M	4	-	35	MRMS	40	S
Paka_110	40	S	5	-	40	MSS	25	S
Paka_111	30	MSS	5	-	30	MRMS	30	MSS
Paka_112	20	M	8	-	10	RMR	25	M
Paka_114	25	M	6	-	15	RMR	15	M
Paka_115	30	S	7	-	35	MSS	25	S
Paka_116	25	MRMSS	7	-	25	RMR	20	MSS
Pasa_1	15	MSS	4	-	10	MRMS	1	S
Pasa_2	5	R	5	-	45	MSS	20	MSS
Pasa_3	15	MR	5	-	10	RMR	5	M
Pasa_4	30	M	7	-	40	MSS	5	MSS
Pasa_5	45	S	10	-	15	RMR	10	M
Pasa_6	10	RMR	8	-	35	MSS	20	MSS
Pasa_7	35	MS	7	-	45	MSS	15	M
Pasa_8	10	RMR	7	-	25	MSS	5	M
Pasa_9	25	M	10	-	35	MRMS	10	M
Pasa_10	50	S	7	-	15	MRMS	5	MSS
Pasa_11	35	S	6	-	10	RMR	1	R
Pasa_12	15	MR	8	-	25	MSS	10	M
Pasa_13	25	M	-	-	40	MSS	20	M
Pasa_15	50	S	6	-	20	MRMS	1	R
Pasa_18	35	MSS	5	-	15	MRMS	5	MS
Pasa_19	35	MS	6	-	10	MRMS	5	MSS
Pasa_21	15	MR	10	-	30	MSS	20	MSS
Pasa_22	15	MR	5	-	15	MRMS	10	MSS
Pasa_23	25	MS	6	-	40	MSS	20	MSS
Pasa_24	25	MS	4	-	25	MRMS	10	MSS
Pasa_25	30	MS	7	-	25	MSS	10	MSS
Pasa_26	20	MS	-	-	25	MSS	10	MSS
Pasa_27	20	MS	5	-	15	RMR	10	RMR
Pasa_28	15	R	3	-	15	RMR	10	MS
Pasa_29	25	MS	5	-	25	MRMS	15	M
Pasa_30	25	M	4	-	45	S	15	MSS
Pasa_31	15	M	8	-	20	MRMS	50	S
Pasa_32	15	M	10	-	45	S	20	MSS
Pasa_33	20	MR	5	-	10	MSS	10	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Pasa_34	35	MSS	4	-	15	MSS	1	MR
Pasa_35	40	MS	5	-	20	MRMS	10	MSS
Pasa_37	15	RMR	6	-	45	S	15	MSS
Pasa_38	10	RMR	6	-	45	S	15	MSS
Pasa_39	20	MR	5	-	35	MRMS	20	MSS
Pasa_41	15	MR	8	-	25	MRMS	1	MR
Pasa_42	15	R	8	-	20	MSS	1	R
Pasa_43	35	MS	10	-	20	RMR	10	M
Pasa_44	10	R	3	-	15	MRMS	1	R
Pasa_46	25	MSS	4	-	15	RMR	1	R
Pasa_48	45	S	6	-	30	MSS	10	M
Pasa_51	50	S	5	-	20	MSS	10	M
Pasa_52	30	M	7	-	15	RMR	5	M
Pasa_53	20	MR	7	-	25	MSS	10	M
Pasa_54	10	R	8	-	20	RMR	25	MSS
Pasa_55	10	R	8	-	15	RMR	5	MS
Pasa_56	15	MR	9	-	20	MSS	5	MS
Pasa_57	10	MR	10	-	15	S	1	M
Pasa_58	15	RMRMS	3	-	15	MRMS	5	M
Pasa_59	15	MSS	10	-	20	MRMS	20	MSS
Pasa_60	35	MSS	6	-	10	RMR	5	MS
Pasa_61	55	S	8	-	30	MSS	10	M
Pasa_62	25	M	10	-	30	MSS	5	M
Pasa_63	10	MSS	4	-	15	S	5	MS
Pasa_64	30	MSS	6	-	15	RMR	15	MSS
Pasa_66	25	M	5	-	35	S	15	MSS
Pasa_67	30	MSS	5	-	15	MRMS	1	R
Pasa_68	15	M	6	-	30	S	1	MR
Pasa_69	10	MSS	7	-	30	RMR	15	M
Pasa_70	15	MSS	10	-	15	MRMS	1	M
Pasa_71	45	S	-	-	35	RMR	10	M
Pasa_72	5	MR	10	-	40	MSS	5	M
Pasa_73	45	S	4	-	35	MSS	20	MSS
Pasa_74	40	S	7	-	20	MSS	5	MSS
Pasa_75	25	MSS	8	-	40	S	20	MSS
Pasa_76	35	MSS	4	-	15	RMR	5	MSS
Pasa_77	30	MS	5	-	25	MSS	5	MSS
Pasa_79	20	MS	3	-	15	RMR	5	RMR
Pasa_80	15	M	6	-	15	RMR	1	R

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Pasa_81	40	S	6	-	20	MRMS	5	RMR
Pasa_82	10	M	10	-	50	S	25	MSS
Pasa_83	35	S	4	-	35	MRMS	10	MSS
Pasa_84	40	S	3	-	15	RMR	5	MSS
Pasa_85	15	MSS	4	-	25	MRMS	1	R
Pasa_86	20	M	7	-	15	R	15	MSS
Pasa_87	15	MSS	7	-	40	MRMS	15	MSS
Pasa_88	15	M	4	-	45	S	25	MSS
Pasa_89	15	M	8	-	30	MSS	20	MSS
Pasa_90	40	S	6	-	35	MSS	5	M
Pasa_91	40	S	10	-	10	RMR	5	RMR
Pasa_92	15	M	5	-	15	MSS	10	MSS
Pasa_94	10	RMR	10	-	20	S	10	MSS
Pasa_95	15	R	4	-	25	MRMS	10	MSS
Pasa_96	10	R	-	-	25	MRMS	5	RMR
Pasa_97	15	MR	10	-	40	S	15	MSS
Pasa_98	15	M	6	-	15	MRMS	5	RMR
Pasa_99	15	M	7	-	0	R	5	M
Pasa_100	15	MS	6	-	30	MSS	20	M
Pasa_101	30	M	8	-	35	MSS	15	MSS
Pasa_102	15	M	7	-	20	MSS	10	MSS
Pasa_103	20	M	6	-	45	MRMS	10	M
Pasa_104	35	MSS	7	-	10	RMR	5	M
Pasa_105	15	M	10	-	15	RMR	1	MS
Pasa_106	25	M	6	-	25	MRMS	40	S
Popo_1	30	MS	9	-	15	RMR	5	MS
Popo_2	45	S	7	-	20	MRMS	10	MSS
Popo_3	30	MSS	10	-	10	RMR	1	MS
Popo_4	40	S	6	-	15	RMR	5	MS
Popo_5	15	MSS	7	-	25	S	15	M
Popo_6	40	MSS	6	-	10	RMR	5	MSS
Popo_7	25	MRMSS	5	-	15	RMR	10	S
Popo_8	20	MS	5	-	15	MRMS	5	MSS
Popo_9	20	M	8	-	40	S	30	MSS
Popo_10	20	M	4	-	15	MSS	5	MS
Popo_11	25	M	7	-	25	MSS	5	M
Popo_12	15	MSS	6	-	20	MSS	10	MS
Popo_13	50	S	9	-	15	MSS	35	S
Popo_14	25	MS	7	-	35	S	10	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Popo_15	25	MSS	10	-	30	MSS	20	M
Popo_16	15	MSS	7	-	25	MSS	15	MSS
Popo_17	35	MSS	5	-	10	MRMS	5	MSS
Popo_18	60	S	5	-	25	MSS	5	MSS
Popo_19	15	M	5	-	15	MSS	5	MS
Popo_20	40	MSS	5	-	20	RMR	5	MSS
Popo_21	30	MSS	8	-	20	MRMS	18	MSS
Popo_22	25	MSS	8	-	20	MSS	10	MSS
Popo_23	25	MSS	6	-	5	RMR	20	MSS
Popo_25	10	MS	8	-	20	RMR	20	MSS
Popo_26	30	MSS	3	-	10	RMR	5	MS
Popo_27	30	S	5	-	25	MRMS	20	MSS
Popo_28	25	MS	5	-	15	MSS	1	M
Popo_29	20	MS	3	-	30	MSS	10	S
Popo_30	45	MSS	5	-	45	MSS	15	MSS
Popo_31	40	S	5	-	20	MRMS	10	M
Popo_32	15	M	6	-	10	RMR	10	M
Popo_33	30	MS	2	-	10	R	3	S
Popo_34	25	MSS	6	-	45	S	25	MSS
Popo_35	10	M	8	-	15	S	15	M
Popo_36	40	MSS	6	-	20	MSS	25	MSS
Popo_39	70	S	5	-	15	MRMS	10	MS
Popo_40	30	S	5	-	20	MRMS	10	M
Popo_41	25	MSS	5	-	25	MSS	10	M
Popo_42	40	MSS	7	-	20	MSS	10	S
Popo_43	30	MS	6	-	15	MRMS	15	MSS
Popo_44	15	M	2	-	15	MR	5	MSS
Popo_45	25	MSS	-	-	20	MS	10	M
Popo_46	30	MSS	5	-	35	MSS	10	MSS
Popo_47	20	M	5	-	20	MRMS	25	MSS
Popo_48	20	M	-	-	25	MRMS	5	RMR
Popo_49	10	M	5	-	30	MSS	10	MSS
Popo_50	35	MSS	4	-	25	S	15	MSS
Popo_51	40	S	6	-	10	RMR	1	R
Popo_52	20	M	8	-	50	S	20	M
Popo_53	40	S	5	-	30	MSS	15	MSS
Popo_54	30	S	8	-	45	S	40	S
Popo_56	40	S	-	-	40	MRMS	10	RMR
Popo_57	20	M	4	-	15	S	5	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Popo_58	70	S	3	-	10	RMR	10	M
Popo_59	35	MSS	6	-	70	S	10	MSS
Popo_60	30	M	5	-	20	MRMS	10	M
Popo_61	40	S	7	-	50	S	25	S
Popo_63	25	MS	4	-	10	MRMS	5	M
Popo_66	25	M	4	-	10	S	5	MSS
Popo_68	20	M	4	-	15	MRMS	10	M
Popo_69	35	MSS	6	-	15	MR	10	MSS
Popo_70	40	S	6	-	15	MRMS	15	M
Popo_71	25	M	8	-	30	MSS	30	S
Popo_72	15	MS	5	-	20	S	25	MSS
Popo_73	25	MS	7	-	35	MRMS	30	S
Popo_74	25	MS	8	-	25	S	10	M
Popo_76	35	S	6	-	15	MRMS	10	M
Popo_77	20	MR	10	-	25	MRMS	10	M
Popo_78	35	S	4	-	20	S	10	MSS
Popo_79	25	M	3	-	15	RMR	1	MR
Popo_81	45	S	4	-	25	MR	10	MSS
Popo_82	65	S	5	-	20	MRMS	10	MSS
Popo_83	65	S	4	-	20	MRMS	15	MSS
Popo_84	20	MSS	5	-	15	MRMS	0	R
Popo_85	65	S	5	-	20	MSS	15	MSS
Popo_86	45	MSS	3	-	15	MRMS	5	MSS
Popo_87	25	MSS	4	-	15	MRMS	20	S
Popo_88	15	MR	6	-	10	RMR	10	MSS
Popo_90	25	MS	4	-	15	MRMS	10	MSS
Popo_91	45	MSS	-	-	10	RMR	1	R
Popo_92	30	MS	4	-	15	RMR	10	MSS
Popo_93	30	MSS	6	-	50	S	30	MSS
Popo_94	20	M	8	-	25	MRMS	15	M
Popo_95	35	MS	-	-	30	MRMS	10	M
Popo_96	25	MSS	7	-	30	MSS	15	MSS
Popo_97	20	M	3	-	15	MRMS	5	MSS
Popo_98	45	S	6	-	25	MSS	15	MSS
Popo_101	30	MS	5	-	40	MSS	15	MSS
Popo_102	35	S	6	-	15	MSS	5	M
Popo_103	30	MSS	6	-	25	MRMS	10	M
Popo_104	25	MSS	5	-	35	MSS	30	S
Popo_105	40	S	6	-	15	RMR	5	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Popo_106	30	MSS	7	-	30	MRMS	15	MSS
Popo_108	35	MSS	6	-	25	MRMS	20	MSS
Popo_109	25	M	9	-	10	MRMS	5	MS
Popo_110	25	M	4	-	10	MRMS	1	RMR
Popo_111	10	MR	8	-	10	MRMS	5	MSS
Romany_1	25	MSS	-	-	40	MRMS	5	MR
Romany_2	20	M	6	-	20	RMR	5	M
Romany_4	20	MR	7	-	5	R	10	MSS
Romany_5	25	M	7	-	40	MRMS	15	M
Romany_6	20	M	8	-	15	MRMS	15	M
Romany_7	40	S	8	-	30	S	30	MSS
Romany_8	20	MR	8	-	35	MRMS	10	M
Romany_9	20	MR	8	-	15	MRMS	10	MSS
Romany_10	15	M	8	-	40	MRMS	15	M
Romany_12	20	MR	7	-	35	MRMS	10	RMR
Romany_13	25	M	6	-	20	RMR	1	M
Romany_14	30	M	8	-	15	RMR	5	M
Romany_15	30	MS	7	-	30	MRMS	10	MSS
Romany_16	15	M	7	-	15	RMR	5	M
Romany_18	25	MSS	10	-	20	S	5	MS
Romany_19	15	M	7	-	20	RMR	5	MS
Romany_20	30	MSS	8	-	25	S	30	MSS
Romany_21	20	MS	10	-	30	MRMS	15	M
Romany_22	25	M	6	-	15	RMR	5	M
Romany_23	0	R	6	-	0	R	1	MR
Romany_24	45	MSS	7	-	65	S	30	S
Romany_25	35	MSS	8	-	35	S	25	MSS
Romany_26	25	M	10	-	20	RMR	25	MSS
Romany_27	15	MR	8	-	15	RMR	5	M
Romany_28	0	R	8	-	0	R	1	S
Romany_29	0	R	7	-	5	R	5	MS
Romany_30	10	MSS	8	-	15	S	1	MS
Romany_31	25	MSS	8	-	55	MRMS	10	M
Romany_34	0	R	7	-	5	R	1	MR
Romany_35	15	M	8	-	10	RMR	10	MSS
Romany_37	15	M	7	-	15	RMR	5	RMR
Romany_38	15	MSS	4	-	10	MR	1	MR
Romany_39	20	M	8	-	30	MRMS	15	M
Romany_40	30	M	7	-	10	RMR	5	M

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Romany_41	25	M	10	-	15	RMR	10	M
Romany_42	35	MSS	9	-	30	MRMS	5	M
Romany_43	35	MS	8	-	30	RMR	5	RMR
Romany_44	15	MSS	9	-	20	MSS	5	M
Romany_45	25	M	9	-	15	RMR	1	M
Romany_46	20	MR	7	-	20	RMR	5	RMR
Romany_48	15	MR	8	-	20	RMR	5	M
Romany_49	20	MR	8	-	25	MRMS	5	M
Romany_50	40	S	8	-	20	RMR	10	M
Romany_51	35	MSS	8	-	45	S	15	MSS
Romany_52	50	S	8	-	45	S	20	MSS
Romany_53	25	M	2	-	10	MSS	1	R
Romany_54	40	S	7	-	20	MRMS	5	RMR
Romany_55	40	S	8	-	40	MRMS	20	MS
Romany_56	25	M	10	-	25	MRMS	5	MS
Romany_57	45	S	9	-	40	MSS	30	MSS
Romany_58	25	M	8	-	30	RMR	30	M
Romany_59	15	M	8	-	20	RMR	20	MSS
Romany_60	20	MR	10	-	15	RMR	10	M
Romany_62	45	S	7	-	60	S	10	MSS
Romany_63	20	M	-	-	35	RMR	20	M
Romany_64	30	MR	7	-	35	RMR	30	S
Romany_65	15	MR	7	-	20	RMR	20	MSS
Romany_66	15	MSS	10	-	5	MRMS	20	S
Romany_67	20	M	7	-	20	RMR	10	M
Romany_68	30	MSS	10	-	25	RMR	0	R
Romany_69	35	MSS	10	-	25	MSS	15	MSS
Romany_71	25	M	7	-	25	RMR	20	M
Romany_72	30	MRMSS	2	-	30	MSS	10	MS
Romany_73	15	MR	7	-	5	R	5	MSS
Romany_74	20	M	4	-	15	RMR	10	MSS
Romany_75	20	M	6	-	15	MRMS	5	MSS
Romany_76	15	MR	7	-	15	RMR	15	MSS
Romany_77	20	M	10	-	30	RMR	10	M
Romany_79	5	MSS	5	-	10	MR	10	MSS
Romany_80	25	M	10	-	30	RMR	40	S
Romany_82	15	MR	10	-	10	MR	10	S
Romany_83	20	MSS	10	-	10	MSS	5	MSS
Romany_84	20	M	6	-	10	MRMS	5	MSS

Environment	St. Paul 2012		South Africa 2012		St. Paul 2013		Kenya 2013	
Line	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]	Severity[a]	IR[b]
Romany_85	60	S	8	-	70	S	40	MSS
Romany_86	25	MRMSS	7	-	15	MRMS	30	MSS
Romany_87	30	MSS	8	-	35	MSS	10	MS
Romany_88	15	MR	-	-	15	RMR	10	M
Romany_90	25	MSS	7	-	20	MRMS	5	MS
Romany_91	20	MSS	-	-	35	MSS	10	MS
Romany_92	45	MSS	-	-	20	RMR	40	MSS
Romany_93	20	M	5	-	10	RMR	10	MSS
Romany_94	30	M	8	-	30	MRMS	50	MSS
Romany_95	20	MR	10	-	20	RMR	30	M
Romany_96	25	M	10	-	20	RMR	25	MSS
Romany_97	20	MR	8	-	15	RMR	40	S
Romany_98	15	MR	-	-	20	RMR	40	S
Romany_99	25	M	8	-	15	RMR	10	MSS
Romany_101	20	MR	7	-	15	RMR	5	M
Romany_102	35	MSS	7	-	45	S	15	MSS
Romany_104	50	S	7	-	60	S	15	M
Romany_105	20	MR	7	-	10	R	10	MSS
Romany_106	45	S	5	-	45	S	10	M
Romany_107	20	M	8	-	10	S	40	S
Romany_109	45	MSS	8	-	60	S	40	MSS
Romany_110	20	MR	5	-	15	MRMS	10	M
Romany_111	15	MR	6	-	20	RMR	10	MSS
Romany_113	30	MS	6	-	25	S	15	S
Romany_114	20	M	6	-	20	RMR	10	M
Romany_115	35	MSS	-	-	40	MRMS	40	MSS
Romany_116	30	M	-	-	25	S	10	S
Romany_117	40	MSS	8	-	30	S	15	M
Romany_119	35	RMR	9	-	5	R	10	M
Romany_120	35	MSS	7	-	45	S	5	RMR
Romany_121	50	S	8	-	35	S	5	M
Romany_122	15	RMR	8	-	10	R	5	MS
Romany_123	10	MR	5	-	10	MRMS	5	MS
Romany_124	5	MS	6	-	10	R	10	MS
Romany_125	45	S	10	-	40	MRMS	30	MSS
Romany_126	35	MSS	10	-	35	RMR	5	RMR

[a] Disease severity (%) was measured according to modified Cobb scale (Peterson et al. 1948)

[b] Represents infection response of the plants to stem rust