Genome-wide mapping of resistance to stripe rust caused by Puccinia striiformis f. sp. tritici in hexaploid winter wheat

Kebede T. Muleta1,#a, Xianming Chen2 and Michael Pumphrey1*

1Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164-6420, USA
2USDA-ARS, Wheat Health, Genetics and Quality Research Unit and Dep. of Plant Pathology, Washington State University, Pullman, WA 99164-6430.

#ACurrent Address: Department of Agronomy, Kansas State University, Manhattan, Kansas, United States of America

*Corresponding author: Michael Pumphrey, email: m.pumphrey@wsu.edu

SHORT RUNNING TITLE: Mapping of stripe rust resistance in winter wheat
ABSTRACT

Stripe rust, caused by the fungus *Puccinia striiformis* f. sp. *tritici* (*Pst*) is a widespread disease and major limit to wheat production worldwide. The USDA-ARS National Small Grains Germplasm Collection (NSGC) has proven to be a rich source of genetic diversity to improve stripe rust resistance in wheat breeding programs. The objective of the present study was to investigate the genetic diversity and sources of resistance to stripe rust in 441 accessions from the collection representing globally-sourced winter wheat germplasm, as a complement to a previous study. The genome-wide association study (GWAS) was conducted to identify loci conferring resistance to *Pst* based on phenotypic data from four field experiments and greenhouse seedling resistance screening against three races of the pathogen. A total of 5,831 single nucleotide polymorphism (SNP) markers were used to investigate population structure, linkage disequilibrium and marker-trait associations. Our results showed 12 and 7 genomic regions significantly associated with *Pst* resistance based on field adult-plant and greenhouse seedling responses, respectively, at False Discovery Rate adjusted *P* value <0.1. Four of the significantly associated genomic regions were mapped far from previously identified *Pst* resistance genes and QTL, indicating that they represent potentially new *Pst* resistance loci. The present study provides additional insight into the usefulness of wheat germplasm collections as an important resistance source that can be used to incorporate diverse resistance genes into adapted wheat cultivars. Further work should aim at validating the identified genomic regions and the associated molecular markers to enhance their utility in marker-assisted breeding.

Key words: Stripe rust, winter wheat, disease resistance, association mapping, germplasm collection
NSGC, National Small Grains Germplasm Collection; *Pst, Puccinia striiformis* f. sp. *tritici*; FDR, False Discovery Rate; GWAS, Genome-wide Association Study; QTL, Quantitative Trait Loci; IT, Infection Types; SEV, Severity; MAF, Minor Allele Frequency; PCA, Principal Component Analysis; LD, Linkage, Disequilibrium; ANOVA, Analyses of Variance; BLUPs, Best Linear Unbiased Estimates; MLMM, Multi-Locus Mixed-Model, WL, Whitlow; CF, Centeral Ferry
INTRODUCTION

Hexaploid wheat (*Triticum aestivum* L.) is a major staple food for mankind in many parts of the world (Shewry and Hey, 2015). It provides nearly 21% of the food calories and 20% of the protein for more than 4.5 billion people in more than 100 countries (Shiferaw et al., 2013; Goutam et al., 2015; Shewry and Hey, 2015). The fungal pathogen *Puccinia striiformis* Westend. f. sp. *tritici* Eriks. (*Pst*), the causal agent of wheat stripe rust, is a serious threat to global wheat production (Boyd, 2005; Chen, 2005; Wellings, 2011). Virulence shifts in *Pst* population due to the genetic changes in existing local pathogen populations or migration of new races are major challenges to resistance breeding in wheat (Chen, 2007; Hovmøller et al., 2011; Burdon et al., 2014; Wan and Chen, 2014; Wan et al., 2016). Nearly US$1 billion are lost to *Pst* annually (Beddow et al. 2015). Controlling *Pst* through the use of chemical fungicides is less preferred by the growers due to its several limitations such as cost, safety to the environment and the lack of complete control under severe epidemics. Therefore, use of resistant wheat varieties is the most effective and economic method to cope with the changing *Pst* population and to mitigate stripe rust-associated losses (Chen, 2007; Burdon et al., 2014; Ellis et al., 2014).

For decades, wheat rust resistance breeding has been based on the deployment of a few race-specific resistance genes that are sufficiently effective to prevent rust spore production and epidemic development even if the plant is protected by a single resistance gene (Burdon et al., 2014; Ellis et al., 2014; Mundt, 2014). Race-specific resistance in most cases is detected at the seedling stage and remains effective at all stages of plant growth, and therefore, is often referred to as seedling or all-stage resistance, the inheritance of which is often qualitative or Mendelian. The resistance controlled by a single race-specific seedling gene is usually short-lived and it functions only if the *Pst* population is of one or more races that lack virulence against the
specific gene (Chen, 2007; Burdon et al., 2014; Chen et al., 2014; Singh et al., 2015). The rapid
decreases in the genetic make-up of wheat rust pathogens towards virulence against deployed
resistance genes have led to frequent replacement of agronomically superior cultivars due to their
susceptibility to virulent races, which have also been responsible for the occurrence of many
devastating epidemics throughout the world (Boyd, 2005; Chen, 2005, 2007; Milus et al., 2009;
Hovmøller et al., 2011; Wellings, 2011; Chen et al., 2014; Hulbert and Pumphrey, 2014).

In contrast, race non-specific resistance, which is usually controlled by multiple genes
each with a minor effect, provides partial resistance, but effective against all races of the
pathogen (Lagudah, 2011; Chen, 2013; Singh et al., 2014). In many cases, race non-specific
resistance loci become effective or more effective in the adult-plant stage, and hence are referred
to as adult-plant resistance (Parlevliet, 2002). Race non-specific quantitative resistance is
considered to be more durable or remain effective even when used over large acreage for many
years than resistance conferred by the race-specific major genes (Lindhout, 2002). Despite their
durability, however, race non-specific resistance loci have been shown to vary in the timing and
levels of defense across environments and are often inadequate under severe epidemics (Chen,
2014; Singh et al., 2015).

From an applied breeding point of view, both race specific and non-specific resistance have
been, and continue to be used by breeding programs to develop resistant varieties. For durable
varietal performance, breeding strategies can be built on pyramiding effective race-specific
major genes and the utilization of race non-specific minor genes in combinations. The prospect
of developing varieties with complex resistance backgrounds is increasingly facilitated by the
capacity for developing high-density molecular markers to perform genome-wide association
mapping that has a higher mapping resolution and genome-wide prediction, which are more
feasible due to the recent availability of genotyping-by-sequencing (Davey et al., 2011; Poland et al., 2012) and high-density single nucleotide polymorphism (SNP) genotyping platforms (Cavanagh et al., 2013; Wang et al., 2014). In particular, the Illumina® iSelect 9K and 90K wheat SNP chip assays have enabled robust and affordable genotyping of wheat and were successfully used for genome-wide association mapping, linkage mapping and genome-wide predictions of various traits (Bajgain et al., 2015; Kertho et al., 2015; Maccaferri et al., 2015b; Wu et al., 2015; Bulli et al., 2016; Naruoka et al., 2016; Muleta et al., 2017a; b; Turner et al., 2017). The objectives of this study were to: 1) explore genetic diversity among accessions in a world-wide collection of winter wheat accessions and identify new sources of seedling and adult plant resistance to stripe rust; and 2) identify genomic regions underlying resistance to stripe rust in the winter wheat accessions through genome-wide association analysis. Bulli et al. (2016) reported a similar study on 1,175 accessions of the NSGC winter wheat core collection. Studying the remaining 441 accessions in additional stripe rust prone environments with unique stripe rust race compositions and for seedling responses to single isolates under controlled conditions, the present study completed the genome-wide association analysis of the entire core collection of winter wheat accessions. We report the identification of chromosome regions harboring newly discovered putative resistance loci as well previously reported genes/QTL for resistance to Pst.

MATERIALS AND METHODS

Plant materials

The germplasm panel used in the present study is comprised of 441 hexaploid winter wheat (Triticum aestivum ssp. aestivum) accessions that were obtained from the USDA-ARS Small
Grains and Potato Germplasm Research Unit (Aberdeen, Idaho). The accessions were originally from diverse global production environments, including 89 countries in the six continents [Asia (55.8%), Africa (3.4%), South America (5.8%), North America (6.4%), Europe (25.6%); and Australia and New Zealand (3.0%)] (Figure 1). Furthermore, the panel contains landraces (62.0%), cultivated lines (12.5%), breeding lines (9.5%) and cultivars (16.0%). Genetically redundant accessions were sorted out based on Identity-By-Descent (IBD) kinship analysis, and represented by only one individual accession. Accordingly, 391 non-duplicate accessions were used for the analyses. **Table S1** shows the name, plant ID, country of origin and improvement status of the 391 winter wheat accessions.

**Field-based evaluation of stripe rust resistance**

The 441 winter wheat accessions in the diversity panel were evaluated for resistance to stripe rust under field conditions at two locations in the Pacific Northwest (PNW) of the US during 2013-2014. The locations were Central Ferry (46°38′9.816″N 117°47′24.0792″W) and the Whitlow Farm near Pullman (46°43′16.59″N 117°0′0.612″W), Washington, USA. The Whitlow farm represent a rain fed semi-arid wheat belt area of the PNW, while Central Ferry is an irrigated farm. Both nursery locations are conducive for naturally occurring *Pst* inoculum and stripe rust development, hence natural infection was relied on for stripe rust development in the field. Accessions were evaluated in non-replicated plots of 1 m-row spaced 35 cm apart. About five grams of seed of each accession were sown to individual plots. Winter wheat line ‘PS 279’, known to be susceptible to all *Pst* races, was used as a susceptible check and planted every 20 rows and on each side of the plot to insure uniform disease pressure across the experimental plots. Host reaction, i.e. infection types (IT), to stripe rust was estimated using a 0–9 scale (Line and Qayoum, 1992). Stripe rust severity (SEV) was assessed visually as percentage of infected
leaf area when stripe rust severities on flag leaves of the susceptible check (‘PS 279’) reached 80-100%. Disease ratings were done on a whole-plot basis.

**Seedling evaluation of stripe rust resistance**

Three races of *Pst* (PSTv-14, PSTv-37 and PSTv-40), representing prevalent races of the stripe rust pathogen population in Washington and across the US (Wan and Chen, 2014) were used to screen seedlings of the accessions under controlled conditions in a greenhouse. The virulence/avirulence profile of the *Pst* races (Table S2) were based on reactions on seedlings of standard differentials used in the United States (Wan and Chen, 2014). Five seeds of each accession were planted in 72-cell trays containing Sunshine mix #1 (Sungro Horticulture Distribution Inc., Quincy, MI, USA). Slow-release commercial fertilizer, Osmocote 15-9-12, N-P-K (Everris NA Inc., Dublin, OH, USA), was applied during planting. Susceptible check ‘Avocet S’ and the stripe rust single-gene differential lines were included in each tray. Seedlings at the two-leaf stage (approximately at 10 days after planting) were spray-inoculated with fresh urediniospores suspended in Soltrol-170 oil (Phillips Petroleum, Bartlesville, OK, U.S.A) at a rate of 0.01 g/mL. Inoculated seedlings were placed in a clean dark dew chamber for 16–24 hours at 13°C and 98% humidity and then incubated in a growth chamber at 17°C/12°C (day/night) with a 16-hour photoperiod. Disease reaction was assessed 16–18 days post-inoculation on a scale of 0-9 (Line and Qayoum, 1992).

**Genotyping**

The accessions were genotyped at the USDA-ARS genotyping laboratory in Fargo, ND, USA using the Illumina iSelect beadchip assay for wheat having 9K SNPs. Markers with minor allele frequency (MAF) less than 0.05 and missing data greater than 10% were removed. Similarly,
accessions with greater than 10% missing genotypic data were excluded from the GWAS analysis. Only markers with known genetic map positions were used. The genetic map information was based on the wheat consensus map developed from Illumina iSelect 9K wheat array (Cavanagh et al. 2013). A total of 5,831 high quality SNPs were used for the GWAS and population structure analyses. The molecular marker information was also used to identify genetically identical accessions and eliminate duplicates.

**Population structure and linkage disequilibrium analyses**

The Bayesian model-based clustering algorithm implemented in the STRUCTURE software version 2.2.3 (Pritchard et al. 2000) was used to infer population structure in the germplasm panel. A burn-in of 50,000 iterations and 100,000 Monte Carlo Markov Chain (MCMC) replicates were set to determine K values (number of subpopulations) in the range of 1 to 10. For each K, five independent runs were carried out. Principal component analyses (PCA) were also conducted using the PRINCOMP procedure in JMP Genomics v6.0 (SAS Institute, Cary, NC). Estimates of the principal components were used as covariates in mixed models for association analysis. An identity-by-decent matrix (K-matrix) estimated as a centered relatedness matrix in JMP Genomics v6.0 was used to estimate population relatedness and used as covariates in mixed models for association analyses. The distribution of phenotypic trait values across the sub-populations were assessed by conducting correlation analyses between the percentage of sub-population membership determined by structure analysis and stripe rust SEV and IT.

Pairwise measures of linkage disequilibrium (LD) between pairs of markers were estimated as squared allele frequency correlations ($r^2$) between pairs of intra-chromosomal SNPs using the software package JMP GENOMICS version 6.1. To determine the average pattern of genome-wide LD decay over genetic distance, a scatterplot of $r^2$ values against the
corresponding genetic distance between markers were constructed. The second-degree locally
weighted polynomial regression (LOESS) based curve was fitted to estimate the extent of LD
decay (Cleveland, 1979).

**Phenotypic data analysis**

Stripe rust IT and SEV values from the field experiments were subjected to analyses of variance
(ANOVA) using the PROC MIXED COVTEST statement in SAS v.9.3. ANOVA was calculated
by including the genotypes, environment and genotype by environment interactions as random
factors. Genotype adjusted means were computed based on best linear unbiased estimates
(BLUPs). Pearson correlation coefficients between different locations and years as well as
between BLUP values and each environment were calculated to determine the consistency of IT
and SEV across the environments. The broad-sense heritability ($H^2$) estimates were calculated
for each location and across all environments according to the following formula:

$$ H^2 = \frac{\sigma^2_G}{\sigma^2_G + (\sigma^2_E/y) + (\sigma^2_{GXE}/y) + \sigma^2_{error}/y} $$

Where $\sigma^2_G$ is the genotypic variance, $\sigma^2_E$ is the environment variance, $\sigma^2_{GXE}$ is the genotype by
environment interaction variance, and $\sigma^2_{error}$ is the residual error variance and $y$ is the number of
years within each location.

**Marker-trait association**

Genome-wide scan of loci governing seedling and adult-plant resistance to stripe rust in the 391-
winter wheat association mapping panel was conducted using the multi-locus mixed-model
(MLMM) (Segura et al., 2012). MLMM performs stepwise mixed-model regression with
forward inclusion and backward elimination based on the significant SNP markers included as
fixed-effect cofactors. The 5,831 informative SNPs were tested for genome-wide significance of
associations with stripe rust resistance by including the first two PCs (explaining 21.6% of the cumulative variation) and the kinship matrix as covariates in the MLMM analysis. Marker-trait associations were considered significant at a threshold of false discovery rate (FDR) value of less than 0.1. Multiple co-segregating significant markers were assigned to a unique QTL region based on the criteria that inter-marker genetic distance of less than 6.8 cM (genetic distance beyond which LD is due to physical linkage) and LD $r^2 > 0.1$ among the markers with the QTL region. Among the multiple co-segregating markers within a confidence interval, ones that showed the strongest association were used to represent the QTL.

RESULTS

Estimates of variance components and heritability

The mixed model analysis of variance components for stripe rust IT and SEV are summarized in Table S3. Both IT and SEV showed highly significant ($P < 0.0001$) differences among the genotypes and displayed significant genotype × environment interactions. Based on the BLUP values of the stripe rust response data, 42 accessions (10.7%) were highly resistant (IT = 0-3), while 100 accessions (25.6%) were highly susceptible (IT = 7-9) across all environments. The remaining 249 accessions (63.7%) showed either an intermediate reaction (IT = 4-6) across all environments or variable stripe rust responses across the different environments. Twenty-nine accessions (7.4%) showed high resistance to all the three races of $Pst$ at seedling stage. Twelve accessions (3.1%) were identified to be highly resistant to all the three races at the seedling stage and in all field environments (Figure 2, Table S5).

Broad-sense heritability ($H^2$) estimates for stripe rust IT ranged from 0.79 at Whitlow to 0.86 at Central Ferry. For SEV, heritability estimates were in the range of 0.82 at Whitlow to
0.79 at Central Ferry (Table S3). Correlation coefficients between the different locations and years for stripe rust IT and SEV are summarized in Table S4. The Pearson correlation coefficients between stripe rust IT and SEV between the multiple locations over multiple years averaged 0.75 and 0.73, respectively. Average correlations between years within locations were 0.74 and 0.78 for IT, and 0.74 and 0.73 for SEV at Whitlow and Central Ferry, respectively. Correlation between Whitlow and Central Ferry (over multiple years) averaged 0.75 for IT and 0.73 for SEV, respectively) (Table S4).

**Genome-wide distribution of SNP markers and LD analysis**

Among the 8,631 SNP markers included in the wheat 9K iSelect assay, 5,831 (67.6%) were polymorphic with MAF >5% in the 391 winter wheat association mapping panel. Among the 5,831 SNPs, 47.4% were mapped to the A genome, 45.9% to the B genome and 6.7% to the D genome. For all the three genomes, more than 99% of the markers were mapped with a distance shorter than 10 cM. For the A and B genomes, none of the markers were mapped with a gap larger than 20 cM. However, the D genome had ten large gaps of greater than 20 cM. The A genome had seven gaps larger than 10 cM, while the B genome had 3 gaps larger than 10 cM (Figure S1).

Linkage disequilibrium was estimated by squared correlation coefficient ($r^2$) from all pairs of SNPs along each chromosome. Average genome-wide LD was estimated to be 0.06. The critical $r^2$ value, determined by calculating the 95\(^{th}\) percentile of the square root transformed $r^2$ of unlinked markers, was estimated to be 0.11. Based on the locally weighted polynomial regression-based smoothening curve fitted onto the scatterplot of $r^2$ vs genetic distance, LD decayed below the critical $r^2$ value at about 6.8 cM (Figure 3). Chromosome 1A contained the largest percentage (18.4%) of marker pairs in LD due to linkage ($r^2$ greater than the critical value
0.11), while chromosome 4D showed the lowest percentage of such marker pairs (<0.01%). The inter-marker genetic distance at which the smoothening curve intersected the population specific critical $r^2$ value (i.e. 6.8 cM) was used to establish QTL confidence intervals.

**Population structure of the germplasm panel**

The admixture model with correlated allele frequency in the STRUCTURE program was used to determine the model based estimates of population structure in the germplasm panel. The 391 winter wheat accessions were structured into 2 subpopulations (Figure 4; Figure S2). The grouping of the accessions into two subpopulations revealed a distinct pattern related to geographic origin and types of accessions. One subgroup (called cluster 1 hereafter) contained landraces originating predominantly from Asia. The second subgroup (called cluster 2 hereafter) contained accessions originating largely from Europe and few from North America, South America and Australia (hereafter called cluster 2). The model based estimate of population structure was also supported by the principal component analysis (PCA) revealing a grouping pattern that was largely in agreement with the STRUCTURE analysis. The first and second principal components (PCs) explained of 17.2% and 4.4% variation in the germplasm panel, respectively (Figure 4).

The effect of population structure on stripe rust resistance was assessed by performing correlation analysis between the percent sub-population membership (from structure analysis) and BLUP values of IT and SEV from the field experiments. Heat maps of stripe rust IT and SEV produced by individual accessions were also compared to that of population sub-grouping based on structure analysis, identity-by-descent kinship matrix and Ward clustering. Significant correlations between population structure and responses to stripe rust were investigated (Table S4). Most of the accessions with high subpopulation membership in cluster 1 showed highly
susceptible reactions to stripe rust, while most of the accessions with high subpopulation membership in cluster 2 showed high levels of resistance to stripe rust (Figure S3).

Genomic regions associated with resistance to stripe rust

**Adult-plant resistance.** Statistically significant associations between SNP markers and response to stripe rust were identified by applying the multi-locus mixed-model analysis of marker-trait associations. A total of 12 genomic regions were significantly associated with resistance to stripe rust under field conditions based on the False Discovery Rate (FDR) adjusted probability ($P$) of $<0.1$ (Table 1). The significant genomic regions were detected on chromosomes 1A, 1B, 2B, 3B, 4A, 4B, 5A, 5B, and 6B and explained 27.6% and 25.8% of the total variation in stripe rust IT and SEV, respectively. The genomic region detected on the short arm of chromosome 1A was tagged by SNP *IWA5976* that showed highly significant and stable associations across environments. *IWA5976* explained 4.8% and 4.1% of the variation in stripe rust IT and SEV, respectively. Two other SNP markers were also mapped within the confidence interval of this genomic region (*IWA3635* and *IWA8001*) that showed significant associations at a marker-wise $P<0.01$ at multiple environments, but failed to meet the criteria for a significant association based on FDR $P<0.1$. LD between *IWA5976*, *IWA3635* and *IWA8001* ranged from 0.53, 0.42 and 0.03, indicating that they most likely represent the same putative QTL.

Three SNP markers on chromosome 1BS (located between 9.6 and 11.6 cM) were identified as highly significant loci for stripe rust resistance across years and locations. This genomic region was tagged by SNP *IWA4715*. Two other SNP markers (*IWA406* and *IWA7331*) that are mapped within the confidence interval of the QTL-tagging SNP also showed significant associations at marker-wise $P<0.01$ at multiple locations and years. LD values between *IWA4715* and the other associated SNPs (*IWA7331*and *IWA406*) were 0.6, and 0.1, respectively,
indicating that the three SNPs likely represent the same putative stripe rust resistance gene. *IWA4715* also showed a highly significant association with resistance to stripe rust at the seedling stage against race PSTv-14. On chromosome 2B, a genomic region significantly associated with field resistance to stripe rust at FDR adjusted $P < 0.1$ was identified at a position of 44 cM. This genomic region was identified by *IWA7800* and explained on average 2.9% and 3.2% of the variation in stripe rust IT and SEV, respectively.

On chromosome 3B, two genomic regions showed highly significant associations with field resistance to stripe rust. These genomic regions were tagged by SNPs *IWA1703* and *IWA1196* that were mapped at 89.4 and 97.1 cM, respectively. LD between these two genomic regions was less than 0.01, which indicates that the two SNPs likely represent two distinct genomic regions.

On chromosome 4A, a putative resistance locus tagged by the SNP *IWA3981* (mapped at 85.2 cM) showed a highly stable and strong association with stripe rust IT and SEV. The variation in stripe rust response explained by *IWA3981* was 2.6% for both IT and SEV. Two genomic regions on chromosome 4B (*IWA113* at 59.8 cM and *IWA1100* at 106.3 cM) were also significant at FDR $P < 0.1$. Based on their genetic map position and LD between them ($r^2 = 0.05$), *IWA113* and *IWA1100* represent different putative resistance loci. SNPs *IWA14* (on chromosome 5A at 11.2 cM), *IWA6627* (on chromosome 5B at 130.4 cM) and *IWA1721* (on chromosome 6B at 37.9 cM) were among the significantly associated SNPs with field-based resistance to stripe rust at multiple environments.

**Mapping of seedling resistance.** GWAS identified seven genomic regions significantly (FDR $P < 0.1$) associated with resistance to the three races of *Pst* (PSTv-14, PSTv-37 and PSTv-40) tested at seedling stage in controlled environments (*Table 2*). Four of the seven genomic regions detected for seedling resistance were detected for PSTv-14. These include, *IWA4715*...
(mapped to Chromosome 1B at 11.6 cM), *IWA6831* (mapped to chromosome 1B at 131.5 cM), *IWA3621* (mapped to chromosome 2B at 114.3 cM) and *IWA7815* (mapped to 5B at 121.8 cM).

Among the four genomic regions detected against PSTv-14, *IWA6831, IWA3621* also showed moderate associations in the field experiment. *IWA4715* showed a highly significant (FDR adjusted *P* <0.1) associations (*P* <0.001) both with field-based and seedling resistance to *Pst*. For PSTv-37, two genomic regions were significant at FDR *P* <0.1. These include *IWA6550* (mapped to chromosome 6A at 126.5 cM) and *IWA2105* (mapped to chromosome 7B at 19.0 cM). Both genomic regions showed moderate associations in field conditions. A single genomic region was detected significant against PSTv-40 (*IWA1375*), which was mapped to chromosome 1A at 7.1 cM. This genomic also showed highly significant marker-wise associations in the field tests.

**DISCUSSION**

**Identifying new sources of resistance to *Pst* in the winter wheat accessions**

The continual evolution of new virulent races of *Pst* has been a lingering challenge to global wheat production (Chen, 2005, 2007, Singh et al., 2005, 2015, Hovmøller et al., 2008, 2011; Milus et al., 2009; Wellings, 2011; Burdon et al., 2014; Sørensen et al., 2014; Chen et al., 2014; Hulbert and Pumphrey, 2014). Most of the available host genetic resistances are race-specific and easily overcome by pathogen evolution. This necessitates a constant search for new sources of resistance and development of varieties with durable resistance. Global efforts are currently underway to explore wheat germplasm collections maintained in gene banks for new sources of resistance and other agriculturally useful traits (Gurung et al., 2014; Kertho et al., 2015; Maccaferri et al., 2015b; a; Sehgal et al., 2015; Rahmatov et al., 2016; Muleta et al., 2017a). The
ability to efficiently unlock the genetic diversity of wheat germplasm collections and their utilization in wheat breeding has become more feasible with the recent improvements in genotyping technology and statistical models for quantitative trait analysis (Cavanagh et al., 2013; Wang et al., 2014). In the present study, we report the identification of resistance sources and genome-wide mapping of stripe rust resistance genes in globally-sourced winter habit hexaploid wheat accessions using high-density SNP markers and phenotypic data from field and greenhouse experiments.

Wheat germplasm of the primary gene pool, including the wild and early domesticated relatives of wheat, landraces and breeding lines, are considered a promising source of resistance to rusts and other diseases in wheat (Maccaferri et al., 2015b; Bulli et al., 2016; Muleta et al., 2017a; b; Turner et al., 2017). The core subset of the USDA-NSGC winter wheat landraces collection characterized in this study showed substantial variation in stripe rust resistance observed on both adult plants in fields and seedlings in the greenhouse. By evaluating the resistance of the germplasm collection against predominant and highly virulent races at the seedling stage as well as at the adult-plant stage under field conditions, we were able to identify several effective sources of resistance to *Pst* in the winter wheat accessions. The winter wheat accessions with a higher percentage of stripe rust resistance-associated alleles can be utilized to broaden the genetic base of rust resistance in wheat breeding germplasm.

**Population structure and its relationships with response to *Pst***

Understanding the patterns of genetic diversity and population structure in the germplasm panel is essential for the utilization of plant genetic resources for genetic studies and germplasm development (Atwell et al., 2010). In this study, we performed population structure analyses based on a model-based clustering algorithm and distance-based hierarchical Fast Ward
clustering approaches to reveal genetic groups in the population. The analyses revealed two
distinct major sub-groups in the USDA-NSGC winter wheat population. The two sub-
populations reveal a major division between accessions from accessions from Asia (sub-
population 1) and Europe, North America and South America (sub-population 2) and grouping
between landrace lines and; advanced breeding lines and cultivars. Previous studies by
Maccaferri et al., 2015b; Bulli et al., 2016; and Muleta et al., (2017a) reported similar patterns of
clustering between accessions from Asia and the West (Europe and the North America) and has
been attributed to the effect of intensive natural selection and enrichment of favorable alleles for
key traits through breeding.

The effect of population structure on stripe rust response was further assessed by means of
correlation between BLUP values of stripe rust IT and SEV and population subgroup
membership. Modest to high influence of population structure was detected for both \( Pst \) IT and
SEV, with Pearson’s correlation coefficients from 0.15 to 0.38 and \( P < 0.0001 \). Sub-population 1
contained a higher proportion of susceptible accessions compared to sub-population 2. Among
the Asian sub-population, resistant accessions were largely landraces and originated mainly from
Iran and a few from the neighboring countries of the Middle East. Kertho et al., (2015) also
reported that landrace accessions originating from Iran showed a high level of resistance to
multiple races of leaf rust and PSTv-37. The identification of accessions with high levels and
broadly effective stripe rust resistance in the sub-population 1 (mainly breeding population) may
be due to the accumulation of favorable allele due to selection in breeding (Bulli et al., 2016).

**Marker-trait association and alignment of the significant QTLs to previously identified \( Yr \) genes/QTLs**
The multi-locus mixed-model analyses of associations between the SNP markers and response to stripe rust in the winter wheat accessions highlighted several chromosome regions harboring previously reported genes/QTL as well as newly identified putative resistance loci. Based on the integrated genetic map developed by Maccaferri et al. (2015b), fifteen of the nineteen genomic regions identified for the field-based and seedling resistance to stripe rust were mapped close to previously identified stripe rust resistance genes and QTL. The remaining four genomic regions, including those tagged by IWA14 on chromosome 5A, IWA6627 on chromosome 5B, IWA6550 on chromosome 6A and IWA2105 on chromosome 7B do not correspond to any of previously identified genes or QTL for resistance to Pst, and thus represent potentially novel loci that are highly stable and effective across multiple years and locations (Table S6). The association of the genomic region tagged by IWA3981 on chromosome 4A has also been reported by Bulli et al. (2016) in a larger, but distinct global collection of winter wheat accessions, indicating that this genomic region is broadly effective against Pst. The molecular markers associated with newly discovered resistance loci can be exploited for diversifying stripe rust resistance in wheat breeding programs.

Two genomic regions were identified on the short arm of chromosome 1A. These include IWA1375 mapped at 7 cM and IWA5976 mapped at 15 cM proximal to IWA1375. IWA1375 was detected at seedling stage for resistance to Pst race PSTv-40, while IWA5976 was effective for the field resistance to stripe rust across multiple environments. The confidence interval of the genomic regions tagged by IWA1375 and IWA5976 overlap with previously reported stripe rust resistance QTL, such QYr.cim-1A.1_GWAS, QYr.cim-1A.2_GWAS, QYr.tam-1A_Avocet, QYr.cim-1A.2_GWAS and QYr.sgi-4A.1_Kariega (Crossa et al., 2007; Prins et al., 2011; Basnet et al., 2014). It is possible that IWA1375 and IWA5976 represent alleles of the previously
reported stripe rust resistance QTL. Yet, allelism testing will be required to determine the
relationships between the detected loci and previously reported genes and/or QTL.

On the short arm of chromosome 1B, IWA4715 and IWA7331 showed highly significant
and consistent associations with field resistance to stripe rust at multiple environments as well
as for seedling resistance against multiple races. The association of IWA4715 and IWA7331
with resistance to stripe rust was also reported by Naruoka et al. (2015) in winter wheat breeding
germlasm from the PNW region of the United States and by Kertho et al., (2015) in winter
wheat landraces. This shows that the locus identified by IWA4715 and IWA7331 is highly
effective and widely distributed across different populations. The map positions of IWA4715 and
IWA7331 overlap with the highly effective stripe rust resistance gene Yr15. In a QTL analysis of
stripe rust resistance in a recombinant inbred line population derived from the club wheat ‘Coda’
by the PNW soft white winter wheat Brundage, Case et al., (2014) reported a QTL (Qyrco.wpg-
1B.1) that was mapped within the confidence interval of IWA4715 and IWA7331 and showed
effectiveness at seedling and adult-plant stages. Further studies will be required to establish the
relationship between IWA4715 and previously reported stripe rust genes/QTL.

Another genomic region on chromosome 1B was identified by SNP IWA6831 for
seedling stage resistance to race PSTv-14. IWA6831 maps to the gene-rich region of
chromosome 1B where a number of Yr genes have been mapped previously. These includes Yr29
(Rosewarne et al., 2012), YrExp1 (Lin and Chen, 2007), QYr.cim-1BL_Francolin (Lan et al.,
2014), QYr.tam-1B_Quaiu (Basnet et al., 2014), QYr-1B_Saar (Lillemo et al., 2008), QYr.sun-
1B_Wollaro (Bansal et al., 2014), QYr.sun-1B_Kukri (Bariana et al., 2010), and QYr_Pavon76
(William et al., 2006). Additional research is needed to fully characterize chromosome 1B
resistance and determine if \textit{IWA6831} is related to one of the previously mapped stripe rust
resistance genes or QTL.

On chromosome 2B, a genomic region associated with field resistance to \textit{Pst} was
identified at FDR-adjusted probability \((P) < 0.1\) at multiple locations and years. This locus was
tagged by the SNP \textit{IWA7800} at the map position of 44 cM. The stripe rust resistance gene \textit{Yr27}
(Rosewarne et al. 2008) and two other QTL, \textit{QYrlu.cau-2BS1_Luke} (Guo et al. 2008) and
\textit{QYr.cim-2B.2_GWAS} (Crossa et al. 2007), have been previously mapped within the confidence
interval of \textit{IWA7800} (McDonald et al., 2004; Crossa et al., 2007; Guo et al., 2008). An additional
genomic region on chromosome 2B was identified by \textit{IWA3621} for highly significant association
with seedling resistance to \textit{Pst}. The association of the genomic region tagged by \textit{IWA3621} on
chromosome 2B has also been reported by Muleta et al. (2017) in a global collection of spring
wheat. \textit{IWA3621} is mapped at 114 cM, which is within the confidence interval of several
previously identified stripe rust QTL. These QTL include \textit{YrKK}, \textit{QYr.inra-2B.1_Camp}
\textit{Remy}, \textit{QYr.ucw-2B_UC1110}, \textit{QYr-2B_Opata 85}, \textit{QYr.tam-2BL2_TAM111}, \textit{Qyrlo.wpg-2B_Louise} and \textit{QYrid.ui-2B.2_IDO444} (Boukhatem et al., 2002; Mallard et al., 2005; Carter et
al., 2009; Lowe et al., 2011; Li et al., 2013; Basnet et al., 2014). It is therefore likely
that \textit{IWA3621} represents a genomic region associated to some of the previously identified QTL.

On 3B, SNPs \textit{IWA1703} (mapped at 89.4 cM) and \textit{IWA1196} (mapped at 97.1 cM)
identified genomic regions significantly associated with the response to \textit{Pst} under field
conditions. Stripe rust QTL \textit{QYrpi.vt-3BL_VA00W-38} (Christopher et al., 2013), \textit{QYr.inra-3Bcentr_Renan} (Dedryver et al., 2009) and \textit{QYr.cim-3B_Pastor} (Rosewarne et al., 2012) have
been previously mapped within the confidence interval of \textit{IWA1703} and \textit{IWA1196}, suggesting
that they may represent alleles of the same genes/QTL. The SNP loci \textit{IWA1113} and \textit{IWA1100} on
chromosome 4B at 59.8 cM and 106.5 cM, respectively, were significantly associated with response to *Pst* under field conditions. *IWA113* is mapped close to several previously identified stripe rust resistance genes and QTL, including, *QYr-4B_Avocet* (William et al., 2006), *QYr.sun-4B_Janz* (Zwart et al., 2010), *QYr.ufs-4B_Palmiet* (Agenbag et al., 2012) and *QYrrb.ui-4B_Rio Blanco* (Chen et al., 2012). On the other hand, *IWA1100* is mapped close to *Yr50* (Liu et al., 2013), *Yr62* (Lu et al. 2014) and *QYr-4B_Oligoculm* (Suenaga et al., 2003). Further studies are required to determine the relationships between these genomic regions and previously mapped stripe rust genes and QTL.

Two additional genomic regions were identified on chromosome 5B, which include *IWA4790* (mapped at 12.1 cM) and *IWA7815* (mapped at 121.8 cM). The position of *IWA4790* overlaps the map position of *Yr47* (Bansal et al., 2011). *IWA7815* was detected only at seedling stage against PSTv-14 and did not show resistance under field conditions and against PSTv-37 and PSTv-40, indicating its ineffectiveness to other races of stripe rust. The genomic region detected on the short arm of chromosome 6B (tagged by *IWA1721* at 37.9 cM) was mapped close to *QYr.sun-6B_Janz* (Bariana et al., 2010), *QYrst.wgp-6B.2_Stephens* (Santra et al., 2008), *QYr.tam-6BS_TAM111* (Basnet et al., 2014), *QYr.caas-6B_Naxos* (Ren et al., 2012) and *QYr.ufs-6B_Kariega* (Prins et al., 2011).

**CONCLUSION**

The results of our genome-wide association study highlighted the presence of valuable sources of resistance to stripe rust that could be exploited by breeders to broaden the genetic basis of resistance in elite wheat germplasm. The chromosome regions identified for significant association with stripe rust resistance in the USDA winter collection represent both newly discovered putative resistance loci as well as regions previously reported to be associated with
stripe rust resistance. To enhance the utilization of resistance sources and the corresponding closely linked molecular markers in wheat breeding, additional experimental validation will be required to identify the accessions carrying the favorable SNP alleles associated with resistance genes and determine which of the QTL that mapped close to previously known resistance genes represent novel resistance genes and which ones are alleles of previously mapped genes.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS’ CONTRIBUTIONS

KTM: performed the field experiments, data analysis, and drafted the manuscript. MP: supervised the experiments from start to finish, and critically revised the manuscript. XC: contributed to the design and supervision of the field experiment, and revised the manuscript. All authors have read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This project was supported by the National Research Initiative Competitive Grant 2011-68002-30029 (Triticeae-CAP) from the United State Department of Agriculture National Institute of Food and Agriculture. We thank John Kuehner, Gary Shelton, Dr. Anmin Wan, and Dr. Kent Evans for providing technical support. The funders had no role in study design, data collection, decision to publish, or preparation of the manuscript.

REFERENCE


uncovers multiple targets of selection for improvement in hexaploid wheat landraces and

Chen, X.M. 2005. Epidemiology and control of stripe rust [ Puccinia striiformis f. sp. tritici ] on


Chen, X.M. 2014. Integration of cultivar resistance and fungicide application for control of

Genome-wide identification of QTL conferring high-temperature adult-plant (HTAP)
resistance to stripe rust (Puccinia striiformis f. sp. tritici) in wheat. Mol. Breed. 29(3): 791–
800.


Christopher, M.D., S. Liu, M.D. Hall, D.S. Marshal, M.O. Fountain, J.W. Johnson, E.A. Milus,
K.A. Garland -Campbel, X. Chen, and C.A. Griffely. 2013. Identi fi cation and mapping of
adult-plant stripe rust resistance in soft red winter wheat cultivar ‘ USG 3555 .’ Plant Breed.

Cleveland, W.S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. J. Am.

Crossa, J., J. Burgueno, S. Dreisigacker, M. Vargas, S.A. Herrera-Foessel, M. Lillemo, R.P.

26

2007. Association analysis of historical bread wheat germplasm using additive genetic


Genome-wide genetic marker discovery and genotyping using next-generation sequencing.


Jahier. 2009. Characterization of genetic components involved in durable resistance to


Ellis, J.G., E.S. Lagudah, W. Spielmeyer, and P.N. Dodds. 2014. The past, present and future of

breeding rust resistant wheat. Front Plant Sci 5(November): 641Available at


high-temperature adult-plant and slow-rusting resistance to Puccinia striiformis f. sp. tritici

in wheat cultivars. Phytopathology 98(7): 803–809.


Genome-Wide Association Study Reveals Novel Quantitative Trait Loci Associated with


Table 1. Genomic regions significantly associated with response to stripe rust infection under field conditions. Significant marker-trait associations were determined based on FDR-adjusted $P < 0.1$ in at least one environment.

<table>
<thead>
<tr>
<th>QTL-tagging SNP</th>
<th>Associated SNP in the QTL region</th>
<th>Chr$^3$.</th>
<th>Position (cM)$^4$</th>
<th>MAF$^5$</th>
<th>Major</th>
<th>Minor</th>
<th>IT</th>
<th>SEV</th>
<th>IT</th>
<th>SEV</th>
<th>Environments where the SNPs showed significant association$^7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IWA5976</td>
<td>IWA8001, IWA3635</td>
<td>1A</td>
<td>21.7</td>
<td>0.17</td>
<td>T</td>
<td>C</td>
<td>2.9–5.7</td>
<td>3.4–4.0</td>
<td>4.8</td>
<td>4.1</td>
<td>wl_IT_13, wl_SEV_13, wl_SEV_14, cf_SEV_13, cf_IT_13, cf_IT_14</td>
</tr>
<tr>
<td>IWA4715</td>
<td>IWA406, IWA7331</td>
<td>1B</td>
<td>11.6</td>
<td>0.09</td>
<td>A</td>
<td>G</td>
<td>2.5–3.8</td>
<td>2.5–4.6</td>
<td>3.8</td>
<td>4.3</td>
<td>wl_SEV_13, wl_SEV_14, cf_IT_13, cf_SEV_13, cf_SEV_14</td>
</tr>
<tr>
<td>IWA7800</td>
<td></td>
<td>2B</td>
<td>44</td>
<td>0.15</td>
<td>A</td>
<td>G</td>
<td>2.4–3.1</td>
<td>2.8–5.7</td>
<td>2.9</td>
<td>3.2</td>
<td>wl_IT_13, wl_SEV_13, wl_SEV_14, cf_IT_13, cf_SEV_13, cf_SEV_14</td>
</tr>
<tr>
<td>IWA1703</td>
<td>IWA239, IWA4721, IWA5178, IWA1704</td>
<td>3B</td>
<td>89.4</td>
<td>0.31</td>
<td>G</td>
<td>A</td>
<td>2.2–3.5</td>
<td>2.2–4.1</td>
<td>3.9</td>
<td>3.1</td>
<td>wl_IT_13, wl_SEV_14, cf_SEV_14</td>
</tr>
<tr>
<td>IWA1196</td>
<td>IWA1148, IWA4847, IWA4428, IWA4427</td>
<td>3B</td>
<td>97.1</td>
<td>0.06</td>
<td>C</td>
<td>T</td>
<td>2.9–3.6</td>
<td>2.1–4.1</td>
<td>3.1</td>
<td>2.6</td>
<td>wl_IT_13, cf_IT_13, cf_SEV_13, cf_IT_14, cf_SEV_14</td>
</tr>
<tr>
<td>IWA3981</td>
<td></td>
<td>4A</td>
<td>85.2</td>
<td>0.06</td>
<td>A</td>
<td>G</td>
<td>2.5–3.7</td>
<td>2.2–6.2</td>
<td>2.6</td>
<td>2.4</td>
<td>wl_IT_13, cf_IT_13, cf_SEV_13, cf_IT_14, cf_SEV_14</td>
</tr>
<tr>
<td>IWA1113</td>
<td>IWA7641, IWA58</td>
<td>4B</td>
<td>59.8</td>
<td>0.14</td>
<td>G</td>
<td>A</td>
<td>2.4–6.6</td>
<td>5.4–5.6</td>
<td>4.7</td>
<td>3.6</td>
<td>wl_SEV_13, cf_IT_13, cf_SEV_13, cf_IT_14, cf_SEV_14</td>
</tr>
<tr>
<td>IWA1100</td>
<td>IWA4490, IWA2031</td>
<td>4B</td>
<td>106.5</td>
<td>0.2</td>
<td>C</td>
<td>T</td>
<td>2.1–6.0</td>
<td>2.0–2.8</td>
<td>4.2</td>
<td>3.9</td>
<td>wl_IT_13, wl_SEV_13, wl_SEV_14, cf_IT_13, cf_SEV_14, cf_IT_14</td>
</tr>
<tr>
<td>IWA14</td>
<td></td>
<td>5A</td>
<td>11.2</td>
<td>0.45</td>
<td>A</td>
<td>G</td>
<td>2.2–2.4</td>
<td>2.2–4.1</td>
<td>1.6</td>
<td>1.5</td>
<td>wl_IT_13, wl_SEV_14</td>
</tr>
<tr>
<td>IWA4790</td>
<td>IWA7400</td>
<td>5B</td>
<td>12.1</td>
<td>0.46</td>
<td>A</td>
<td>C</td>
<td>2.0–5.5</td>
<td>2.6–4.3</td>
<td>5.2</td>
<td>4.6</td>
<td>wl_IT_13, wl_SEV_13, wl_IT_14, wl_SEV_14, cf_IT_13, cf_SEV_13, cf_SEV_14</td>
</tr>
<tr>
<td>IWA6627</td>
<td>IWA1755, IWA4829, IWA4566</td>
<td>5B</td>
<td>130.4</td>
<td>0.23</td>
<td>A</td>
<td>G</td>
<td>3.6–5.5</td>
<td>3.9–5.5</td>
<td>3.3</td>
<td>5.4</td>
<td>wl_IT_13, wl_SEV_13, wl_SEV_14</td>
</tr>
<tr>
<td>IWA1721</td>
<td></td>
<td>6B</td>
<td>37.9</td>
<td>0.17</td>
<td>G</td>
<td>A</td>
<td>2.5–7.4</td>
<td>2.1–3.9</td>
<td>3.7</td>
<td>4.8</td>
<td>wl_IT_13, wl_SEV_13, wl_IT_14, wl_SEV_14, cf_IT_13, cf_SEV_13, cf_SEV_14</td>
</tr>
</tbody>
</table>

$^1$SNP index from the wheat 9K iSelect assay, $^2$Other significant SNPs identified within the confidence interval of the QTL, $^3$Chromosome, $^4$based on the consensus map of the wheat 9K iSelect assay by Cavanagh et al. 2013, $^5$Minor Allele Frequency, $^6$Underline indicates favorable allele, $^7$wl = Whitlow near Pullman and cf = Central Ferry, Washington; IT = infection type and SEV = severity; 13 = 2013 and 14 = 2014.
**Table 2.** Genomic regions significantly associated with seedling response to stripe rust infection under greenhouse conditions. Significant marker-trait associations were determined based on FDR-adjusted $P < 0.1$ in at least one environment.

<table>
<thead>
<tr>
<th>QTL-tagging SNP$^1$</th>
<th>Associated SNP in the QTL region$^2$</th>
<th>Chr.$^3$</th>
<th>Position (cM)$^4$</th>
<th>MAF$^5$</th>
<th>Alleles (Major/Minor)$^6$</th>
<th>$P$-values</th>
<th>R square (%)</th>
<th>$PsT$ race</th>
<th>Significant association under field condition ($P &lt; 0.001$)$^7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IWA1375</td>
<td></td>
<td>1A</td>
<td>7.1</td>
<td>0.06</td>
<td>C/T</td>
<td>6.39</td>
<td>5.4</td>
<td>PSTv-40</td>
<td>wl_IT_13, cf_SEV_13, cf_IT_14, cf_SEV_14,</td>
</tr>
<tr>
<td>IWA4715</td>
<td>IWA7331</td>
<td>1B</td>
<td>11.6</td>
<td>0.09</td>
<td>A/G</td>
<td>12.55</td>
<td>4.8</td>
<td>PSTv-14</td>
<td>wl_SEV_13, wl_SEV_14, cf_IT_13, cf_SEV_13, cf_SEV_14,</td>
</tr>
<tr>
<td>IWA6831</td>
<td>–</td>
<td>1B</td>
<td>131.5</td>
<td>0.39</td>
<td>T/C</td>
<td>4.23</td>
<td>3.4</td>
<td>PSTv-14</td>
<td>cf_IT_13, cf_SEV_13</td>
</tr>
<tr>
<td>IWA3621</td>
<td></td>
<td>2B</td>
<td>114.3</td>
<td>0.32</td>
<td>A/G</td>
<td>4.56</td>
<td>2.5</td>
<td>PSTv-14</td>
<td>wl_IT_13, wl_SEV_13, cf.IT_14,</td>
</tr>
<tr>
<td>IWA7815</td>
<td>IWA4280, IWA6111b, IWA6112, IWA5478</td>
<td>5B</td>
<td>121.8</td>
<td>0.08</td>
<td>G/A</td>
<td>16.35</td>
<td>8.1</td>
<td>PSTv-14</td>
<td>–</td>
</tr>
<tr>
<td>IWA6550</td>
<td>–</td>
<td>6A</td>
<td>126.5</td>
<td>0.42</td>
<td>A/C</td>
<td>4.88</td>
<td>4.0</td>
<td>PSTv-37</td>
<td>wl_SEV_13, wl_SEV_114</td>
</tr>
<tr>
<td>IWA2105</td>
<td></td>
<td>7B</td>
<td>19.0</td>
<td>0.19</td>
<td>C/T</td>
<td>6.13</td>
<td>2.6</td>
<td>PSTv-37</td>
<td>cf_SEV_13, cf_IT_13, wl_IT_13, wl_SEV_14</td>
</tr>
</tbody>
</table>

$^1$SNP index from the wheat 9K iSelect assay, $^2$Other significant SNPs identified within the confidence interval of the QTL, $^3$Chromosome, $^4$based on the consensus map of the wheat 9K iSelect assay by Cavanagh et al. 2013, $^5$Minor Allele Frequency, $^6$Underline indicates favorable allele, $^7$wl = Whitlow near Pullman and cf = Central Ferry, Washington; IT = infection type and SEV = severity; 13 = 2013 and 14 = 2014.
Figure captions

Figure 1. Geographic origin of the 441 winter wheat accessions.

Figure 2. Box plots showing the distribution of the responses of adult plants and seedlings of the winter wheat accessions to stripe rust infection. A) Stripe rust infection type from the field experiments, B) Stripe rust severity from the field experiments, C) Infection type responses under greenhouse condition.

Figure 3. Pairwise measure of genome-wide linkage disequilibrium (LD) for the winter wheat accessions using 5,831 polymorphic single nucleotide polymorphism (SNP) markers. LD decay rate was estimated from a plot of $r^2$ (between pairs of polymorphic marker loci) against the genetic distance (cM). The red colored line indicates a LOESS smoothening curve fitted to the LD decay plot, while the blue lines indicate the point at which the smoothening curve decayed below the critical $r^2$ value.

Figure 4. Genetic relatedness and population structure of the winter wheat accessions. Neighbour-joining tree based on genetic similarities among the winter wheat accessions (upper panel), and scatterplot of the two first principal components explaining the genomic variation among the winter wheat accessions (lower panel).

Fig 5. Marker-trait associations and quantile-quantile plots of mixed model analyses for stripe rust resistance at Whitlow near Pullman. The negative base 10 logarithms of $p$-value (vertical axis) from the mixed-linear model GWAS was plotted against the genomic position of each SNP on the wheat chromosomes (horizontal axis). Top to bottom: IT in 2013, IT in 2014, SEV in 2013, and SEV, 2014.

Fig 5. Marker-trait associations and quantile-quantile plots of mixed model analyses for stripe rust resistance at Central Ferry. The negative base 10 logarithms of $p$-value (vertical axis) from the mixed-linear model GWAS was plotted against the genomic position of each SNP on the wheat chromosomes (horizontal axis). Top to bottom: IT in 2013, IT in 2014, SEV in 2013, and SEV, 2014.
Supplemental information

**Table S1.** Accession name, ID, improvement status and origin for the 441 winter wheat accessions.

**Table S2.** Virulence/avirulence formula of the stripe rust isolates used for seedling resistance screening.

**Table S3.** Mean responses of the winter wheat accessions to *Puccinia striiformis* f. sp. *tritici* infection and estimates of broad sense heritability.

**Table S4.** Pearson’s correlation coefficients between the five test environments for stripe rust infection types (IT) and severity (SEV).

**Table S5:** winter wheat accessions that showed highly resistant responses across all locations, to all races of the Pst and combined resistance to both seedling and field based resistance.

**Table S6.** Comparisons of the putative genomic regions significantly associated with stripe rust resistance in the present study with previously characterized gene/QTL.

**Figure S1.** Genome-wide distribution of SNP markers used in the GWAS tests.

**Figure S2:** the likely number of population sub-structure in the global collection of winter wheat collection.

**Figure S3.** Genetic relatedness and population structure of the winter wheat accessions and its effect on response to stripe rust. A) Dandrogram based on Ward clustering algorithm, B) Heat map of the pairwise measure of kinship based on identity by decent (IBD), C) Effect of population structure on stripe rust infection types.