



Article

Identification of Secreted Protein Gene-Based SNP Markers Associated with Virulence Phenotypes of *Puccinia striiformis* f. sp. *tritici*, the Wheat Stripe Rust Pathogen

Qing Bai ¹ , Meinan Wang ¹, Chongjing Xia ^{1,2} , Deven R. See ^{1,3} and Xianming Chen ^{1,3,*}

¹ Department of Plant Pathology, Washington State University, Pullman, WA 99164-6430, USA; qing.bai@wsu.edu (Q.B.); meinan_wang@wsu.edu (M.W.); chongjing.xia@wsu.edu (C.X.); deven.see@usda.gov (D.R.S.)

² Wheat Research Institute, School of Life Sciences and Engineering, Southwest University of Science and Technology, Mianyang 621010, China

³ U.S. Department of Agriculture, Agricultural Research Service, Wheat Health, Genetics, and Quality Research Unit, Pullman, WA 99164-6430, USA

* Correspondence: xianming.chen@usda.gov; Tel.: +1-509-335-8086

Abstract: Stripe rust caused by *Puccinia striiformis* f. sp. *tritici* (*Pst*) is a destructive disease that occurs throughout the major wheat-growing regions of the world. This pathogen is highly variable due to the capacity of virulent races to undergo rapid changes in order to circumvent resistance in wheat cultivars and genotypes and to adapt to different environments. Intensive efforts have been made to study the genetics of wheat resistance to this disease; however, no known avirulence genes have been molecularly identified in *Pst* so far. To identify molecular markers for avirulence genes, a *Pst* panel of 157 selected isolates representing 126 races with diverse virulence spectra was genotyped using 209 secreted protein gene-based single nucleotide polymorphism (SP-SNP) markers via association analysis. Nineteen SP-SNP markers were identified for significant associations with 12 avirulence genes: *AvYr1*, *AvYr6*, *AvYr7*, *AvYr9*, *AvYr10*, *AvYr24*, *AvYr27*, *AvYr32*, *AvYr43*, *AvYr44*, *AvYrSP*, and *AvYr76*. Some SP-SNPs were associated with two or more avirulence genes. These results further confirmed that association analysis in combination with SP-SNP markers is a powerful tool for identifying markers for avirulence genes. This study provides genomic resources for further studies on the cloning of avirulence genes, understanding the mechanisms of host–pathogen interactions, and developing functional markers for tagging specific virulence genes and race groups.

Keywords: correlation coefficient; *Puccinia striiformis* f. sp. *tritici*; secreted protein gene; SNP markers; wheat stripe rust; virulence



Citation: Bai, Q.; Wang, M.; Xia, C.; See, D.R.; Chen, X. Identification of Secreted Protein Gene-Based SNP Markers Associated with Virulence Phenotypes of *Puccinia striiformis* f. sp. *tritici*, the Wheat Stripe Rust Pathogen. *Int. J. Mol. Sci.* **2022**, *23*, 4114. <https://doi.org/10.3390/ijms23084114>

Academic Editors: Akira Ishihama and Franklin W. N. Chow

Received: 5 March 2022

Accepted: 7 April 2022

Published: 8 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Stripe rust (yellow rust), caused by *Puccinia striiformis* Westend. f. sp. *tritici* Erikss. (*Pst*), is a destructive disease that occurs throughout the major wheat-growing regions of the world [1–5]. This obligate biotrophic fungal pathogen is highly variable due to the capacity of virulent races to undergo rapid changes in order to circumvent resistance in wheat cultivars and genotypes and to adapt to different environments [1,3,6–12]. In our group, more than 320 virulent races have been identified with *Pst* collections from the U.S. since the 1960s and ten other countries from 2007 to 2020 [13–16]. Furthermore, some new molecular groups specific to one or more countries have been identified through the analysis of population structure and differentiation, and these currently small groups have the potential to threaten wheat production in other countries [17]. Therefore, it is important to study virulence and genotype changes and the molecular mechanisms behind the rapid changes.

To investigate *Pst* race evolution, researchers have used a combination of the classical theory and modern sequencing technology. Since the gene-for-gene recognition between

host resistance and pathogen avirulence genes was phenotypically demonstrated by Flor in the middle 20th century [18], research on identification and function demonstration of avirulence (*Avr*) genes has been conducted in rust pathogens [19–21]. In the flax rust fungus *Melampsora lini*, the avirulence gene *AvrL567* was first molecularly characterized through genetic mapping along with a cosegregating cDNA probe [20,21]. In recent years, a combination of different technologies, including genetic mapping and genomic approaches, has become popular for studying *Avr* genes in rust fungi. Using this approach, high-density genetic maps were constructed for poplar rust fungus (*Melampsora larici-populina*) [22] and *Pst* [23,24]. In Xia et al. (2020), the QTL analysis of a sexual population mapped six *Avr* genes in three linkage groups and identified a genomic cluster at a single contig containing four *Avr* genes (*AvYr7*, *AvYr43*, *AvYr44*, and *AvYrEpx2*) [23].

Correlation analysis between genetic variations and virulence/avirulence phenotypes has been used to study host–pathogen interactions [25,26]. In rust fungi, highly expressed secreted protein (SP) genes have been demonstrated to include some proteins with important pathogenicity functions [19]. The SP gene-derived SNP markers (SP-SNP) have been used to study population structures and to tag specific virulence genes. In our group, Xia et al. (2016) attempted to identify *Avr* candidate genes in *Pst* and indicated that association analysis of genetic variations (SP-SNPs) and virulence/avirulence phenotypes can be used to identify markers for *Avr* genes. Xia et al. (2017) further took advantage of comparative genomics and correlation analysis and identified more than 900 *Pst*-specific SP genes and 73 *Avr* candidate genes. In a most recent study, 62 additional avirulence candidate genes significantly associated with 16 avirulence genes were identified by means of a comparison of the genomic variations in 30 mutant isolates derived from ethyl methanesulfonate (EMS) mutagenesis with respect to their progenitor isolate [27,28]. In addition, some virulence factors have also been reported in *Pst*, such as *PstSCR1* activating immunity in non-host plants [29], *Pst_8713* involved in enhancing *Pst* virulence [30], and *Pst18363* as an important pathogenicity factor in *Pst* [31]. However, no known *Avr* genes have been identified in *Pst* so far.

In our group, we have utilized SNPs from sequence data to distinguish different *P. striiformis* isolates [32], develop SNP markers for SP genes [33], and use them for characterizing populations, constructing linkage maps, studying virulence, and determining mechanisms for variations [24,27,33–35]. Xia et al. (2017) identified more than 900 *Pst*-specific SP genes and we designed hundreds of primers based on the SP genes. Therefore, the objectives of the present study were to (1) develop more SNP markers using genomic sequences of *Pst* SP genes, (2) further characterize the U.S. and international *Pst* isolates using the new SP-SNP markers, and (3) identify SP-SNPs associated with avirulence genes.

2. Results

2.1. Distribution of Avirulence/Virulence Phenotypes

The infection type (IT) data of the selected 157 *Pst* isolates are provided in Table S1 and the distribution of avirulence and virulence phenotypes determined for the 18 *Yr* single-gene lines are shown in Figure 1. Of the 18 *Yr* gene lines, 2 (*Yr5* and *Yr15*) had only the avirulence phenotype, and thus the identification of markers for virulence in these two resistance genes was not possible. Therefore, these genes were excluded from further analyses. The remaining 16 *Yr* genes had the less frequent phenotype, all above the 0.05 frequency value, and therefore they were suitable for association analysis.

2.2. SP-SNP Markers

After eliminating SP-SNP markers with a minor allele frequency (MAF) <5% and a missing rate >50%, 209 SP-SNPs were retained for subsequent analyses (Table S2). The MAF of the 209 SP-SNPs based on the 157 isolates ranged from 0.05 to 0.50 with a mean of 0.33. The results indicated that the 209 markers were suitable for genotyping *Pst* isolates.

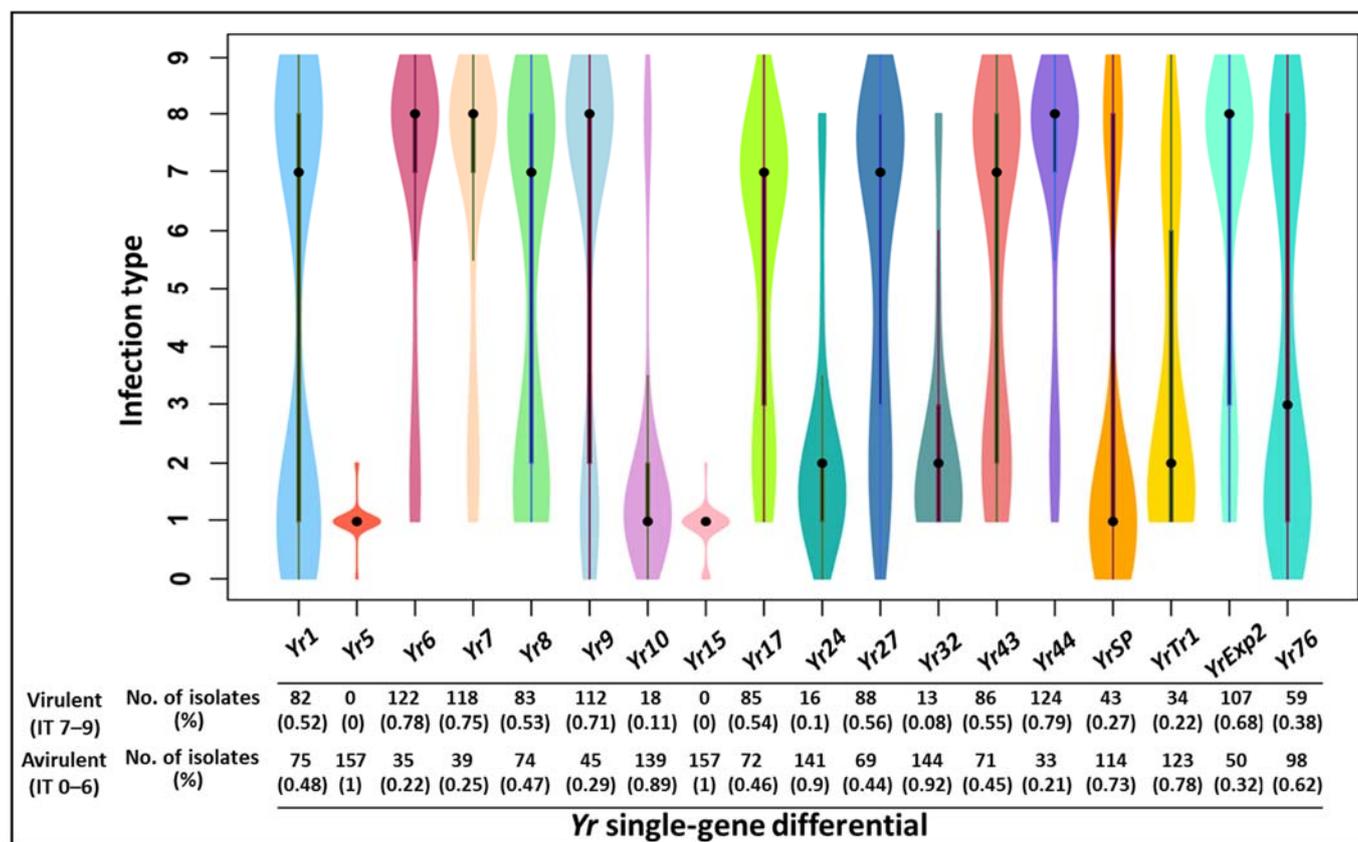


Figure 1. Violin plot showing distributions of infection type (IT) for 157 *Puccinia striiformis* f. sp. *tritici* isolates scored on 18 wheat lines with single *Yr* genes. Solid dots show medians. The numbers of isolates and the frequency values of the two phenotypic classes (avirulent and virulent) are given below the *Yr* genes.

2.3. Population Structure

Principal component analysis (PCA) conducted using the GAPIT program indicated that, using the 209 SP-SNP markers, the 157 *Pst* isolates were optimally separated into three groups (Figure 2A), and PC1 and PC2 explained 19.23% and 7.65%, respectively (Figure 2B). The first two PCs separated the isolates into three molecular groups (MGs). The detailed relationships among the isolates with the country of origin for each of the three MGs are shown in Figure 3. The first group (MG1) was the most diverse, containing the isolates from all countries except Canada. The second group (MG2) was the smallest and was closely related to MG1, containing isolates mainly from Ecuador (80%). Mostly distinct from MG1 and MG2, the third group (MG3) contained isolates from China and countries in North America (the U.S., Canada, and Mexico) and South America (Ecuador). The Efficient Mixed Model Association (EMMA) algorithm was used to establish a kinship matrix and a heat map of values in the kinship matrix showed three groups (Figure S1). These different analyses consistently revealed three molecular groups, and the structures were considered in the following association analysis.

To compare the clusters of the 157 *Pst* isolates in the present study with our previous studies, in which more isolates were genotyped using 14 SSR markers [17,36], another phylogenetic tree with the same 157 *Pst* isolates based on the 14 SSR markers was generated (Figure S2). The clusters in the phylogenetic tree based on the SSR markers (Figure S2) had both similarities and differences with the phylogenetic tree based on the 209 SP-SNP markers (Figure 3). There were three main molecular groups based on both SSR and SP-SNP markers, while the differences were the clustering of the isolates in each group. Most of the isolates in the first group (MG1) based on the SSR markers were also clustered in the MG1 based on the 209 SP-SNP markers. The second group (MG2) based on the SSR markers

contained two subgroups: one subgroup (indicated by the red arrow) contained the isolates that were clustered in the largest group (MG1) based on the 209 SP-SNP markers; the second subgroup (indicated by the black arrow) contained the isolates mainly from Ecuador (9 out of 13 isolates), which was similar to the second group (MG2) based on the 209 SP-SNP markers. The third group (MG3) based on the SSR markers was most distinct from MG1 and MG2 and was similar to the MG3 in the phylogenetic tree based on the 209 SP-SNP markers. Another similarity between the two phylogenetic trees with respect to MG3 was that they both contained most of the Chinese isolates, including 11 of the 12 isolates based on the SSR markers and 8 of the 12 isolates based on the SP-SNP markers. The correlation coefficient of the distance similarity matrices constructed by the two sets of markers was 0.878 ($p < 0.01$). Overall, the high similarity of the clusters based on the SP-SNP and the SSR markers confirmed the population structure and the effectiveness of the SP-SNP markers used in the population genetic analysis.

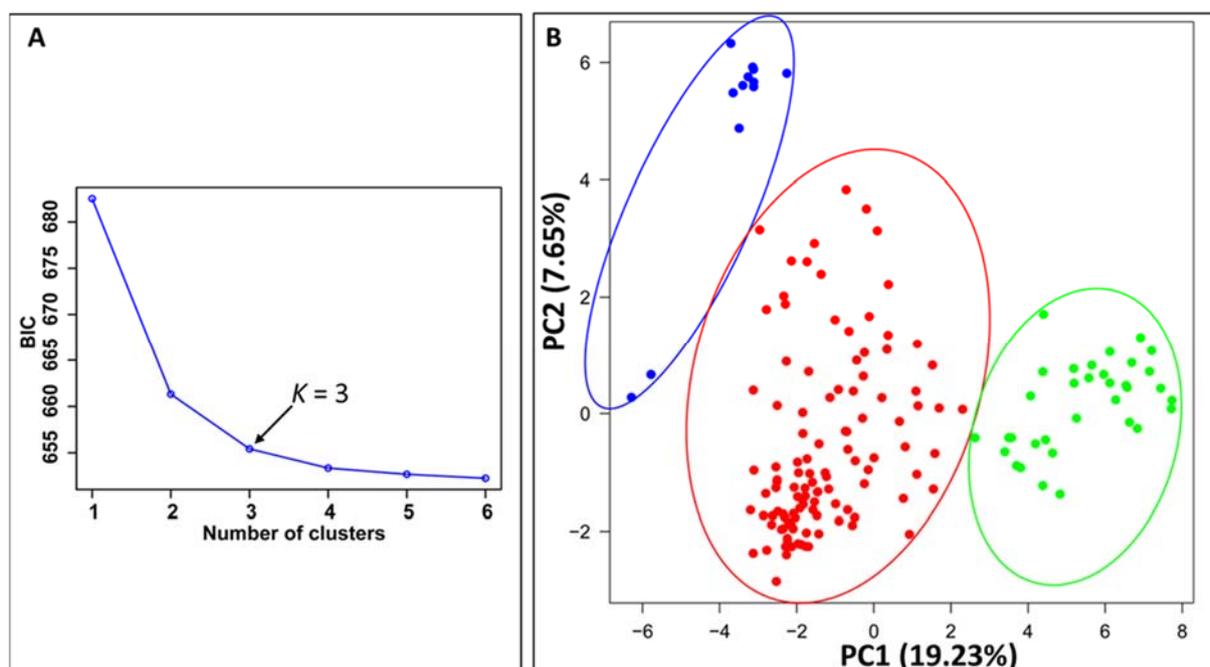


Figure 2. Principal component (PC) analysis of 157 *Puccinia striiformis* f. sp. *tritici* isolates using 209 SP-SNP markers. (A) The optimal K value (indicated by the black arrow) for determining the number of clusters based on the curve of Bayesian information criterion (BIC) values versus the number of clusters assessed with 209 SP-SNP markers. (B) Plot of the second principal component (PC2) against the first principal component (PC1) showing the three molecular groups. Each dot represents an isolate.

2.4. SP-SNPs Significantly Associated with Avirulent Genes

The analysis of a mixed linear model with PCA and kinship identified 19 SP-SNP markers significantly associated with 12 avirulence genes (corresponding to 12 *Yr* genes) as having p -values < 0.01 (Table 1). One marker each was found for *AvYr7*, *AvYr10*, *AvYr32*, *AvYr76*, and *AvYrSP*; two markers each for *AvYr6*, *AvYr24*, and *AvYr43*; three markers for *AvYr44*; and four markers each for *AvYr1*, *AvYr9*, and *AvYr27*. The detailed information on the supercontig and position, p -value for the association, MAF, percentage of variation explained (PVE), and alleles of nucleotides for each of the markers is presented in Table 1. Of the 19 SP genes, 8 were predicted to be effectors (Table S3). The QQ and Manhattan plots for each of the marker–avirulence gene associations are shown in Figure 4. The QQ plot shows the deviation of the observed p -values from the null hypothesis that the SP-SNP is not associated with the avirulence gene. Markers along the diagonal line were not associated, while those away from the diagonal line were associated with the avirulence

gene. The Manhattan plot shows the significant p -values above the threshold ($p < 0.01$, $-\log_{10}(p) > 2$) associated with the avirulence gene. Some of the SP-SNPs were significantly associated with two or more avirulence genes; for example, marker SP.SNP.SC.120.10252 was associated with *AvYr10*, *AvYr24*, and *AvYr32* (Table 1, Figure 4). The 19 markers associated with 12 avirulence genes resulted in 27 significant marker–avirulence gene associations. The association of a single marker with multiple avirulence genes was likely due to the correlation of the phenotypic data of the avirulence genes. To test the hypothesis, correlation coefficients of the virulence phenotypes for different avirulence genes were estimated, and the results are presented in Figure 5. The correlation coefficients between most pairs of the avirulence genes were relatively low or moderate. However, high correlations were observed between *AvYr10* and *AvYr32* ($r = 0.88$, $p < 0.001$), between *AvYr24* and *AvYr32* ($r = 0.81$, $p < 0.001$), and between *AvYr10* and *AvYr24* ($r = 0.77$, $p < 0.001$). The PVE values of the 19 SP-SNP markers were low to moderate, ranging from 0.06 to 0.21, but all were significant ($p < 0.01$) (Table 1).

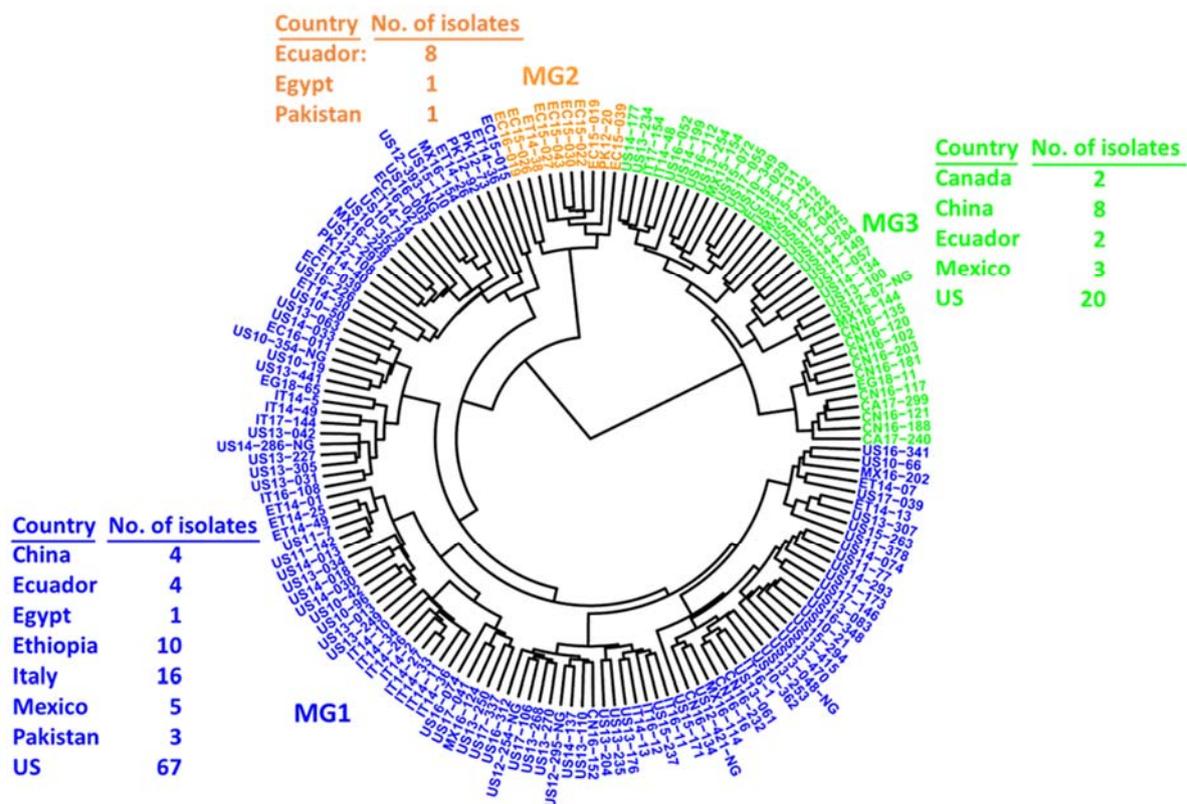


Figure 3. Dendrogram of *Puccinia striiformis* f. sp. *tritici* isolates from nine countries constructed based on dissimilarities assessed with 209 secreted protein gene-based SNP (SP-SNP) markers using hierarchical cluster analysis, showing three molecular groups (MGs) and the isolate numbers from different countries within each MG.

Table 1. SP-SNPs associated with avirulence genes in *Puccinia striiformis* f. sp. *tritici*.

Avirulence Gene ^a	SNP ID	Supercontig ^b	Position in Supercontig ^b	p -Value	MAF ^c	PVE ^d	Allele ^e	Protein ID ^b
<i>AvYr1</i>	SP.SNP.SC.21.152447	21	152447	0.002144	0.38	0.15	C/G	PSTG_04155
	SP.SNP.SC.252.30471	252	30471	0.005105	0.11	0.14	G/T	PSTG_16039
	SP.SNP.SC.23.155650	23	15565	0.008536	0.42	0.13	T/G	PSTG_04466
	SP.SNP.SC.220.72190	220	7219	0.008644	0.45	0.13	A/G	PSTG_15512
<i>AvYr6</i>	SP.SNP.SC.21.152447	21	152447	0.000586	0.38	0.12	C/G	PSTG_04155
	SP.SNP.SC.126.136707	126	136707	0.001405	0.05	0.11	T/C	PSTG_12716

Table 1. Cont.

Avirulence Gene ^a	SNP ID	Supercontig ^b	Position in Supercontig ^b	<i>p</i> -Value	MAF ^c	PVE ^d	Allele ^e	Protein ID ^b
<i>AvYr7</i>	SP.SNP.SC.240.28476	240	28476	0.003675	0.49	0.06	G/A	PSTG_15854
	SP.SNP.SC.126.136707	126	136707	0.001405	0.05	0.11	T/C	PSTG_12716
<i>AvYr9</i>	SP.SNP.SC.21.152447	21	152447	0.000865	0.38	0.14	C/G	PSTG_04155
	SP.SNP.SC.241.57435	241	57435	0.006845	0.49	0.12	T/G	PSTG_15874
	SP.SNP.SC.17.271382	17	271382	0.009011	0.40	0.11	A/T	PSTG_03500
<i>AvYr10</i>	SP.SNP.SC.233.95127	233	95127	0.009733	0.48	0.11	A/G	PSTG_15751
	SP.SNP.SC.120.10252	120	10252	0.004658	0.25	0.07	C/T	PSTG_12413
<i>AvYr24</i>	SP.SNP.SC.120.10252	120	10252	0.000932	0.25	0.11	C/T	PSTG_12413
	SP.SNP.SC.214.8618	214	8618	0.003183	0.24	0.09	T/C	PSTG_15361
<i>AvYr27</i>	SP.SNP.SC.187.104441	187	104441	0.002582	0.10	0.10	A/C	PSTG_14812
	SP.SNP.SC.221.30654	221	30654	0.00544	0.37	0.09	A/T	PSTG_15517
	SP.SNP.SC.12.929747	12	929747	0.007666	0.33	0.09	G/A	PSTG_02640
<i>AvYr32</i>	SP.SNP.SC.21.152447	21	152447	0.008759	0.38	0.09	C/G	PSTG_04155
	SP.SNP.SC.120.10252	120	10252	0.001166	0.25	0.09	C/T	PSTG_12413
<i>AvYr43</i>	SP.SNP.SC.14.299197	14	299197	0.002819	0.08	0.10	C/T	PSTG_02897
	SP.SNP.SC.203.51320	203	5132	0.007604	0.31	0.08	T/G	PSTG_15141
<i>AvYr44</i>	SP.SNP.SC.117.206340	117	20634	0.001316	0.44	0.08	A/G	PSTG_12281
	SP.SNP.SC.221.30654	221	30654	0.00191	0.37	0.07	A/T	PSTG_15517
<i>AvYr76</i>	SP.SNP.SC.14.299197	14	299197	0.003847	0.08	0.06	C/T	PSTG_02897
	SP.SNP.SC.220.75744	220	75744	0.004842	0.46	0.17	A/T	PSTG_15513
<i>AvYrSP</i>	SP.SNP.SC.121.147828	121	147828	0.00136	0.44	0.21	C/A	PSTG_12496

^a Avirulence genes (*AvYr*) correspond to wheat *Yr* resistance genes. ^b Supercontig, position, and protein ID of the SP-SNP markers according to the reference genome PST-78 in the BROAD Institute Puccinia database (<http://www.broadinstitute.org/>, accessed on 24 November 2021). ^c MAF = minor allele frequency. ^d PVE = phenotypic variance explained by the significantly associated markers. ^e For each SP-SNP marker, the first allele was the major allele and the second allele was the minor allele.

2.5. Accuracy and Sensitivity for Detecting Avirulence/Virulence Genes

The accuracies for the detections of avirulence/virulence ranged from 50.39% in the test of the marker SP.SNP.SC.187.104441 with the avirulence gene *AvYr27* to 94.90% in the test of the marker SP.SNP.SC.120.10252 with the avirulence gene *AvYr32* (Table S4). The higher accuracy indicated that the SP-SNP markers could be used to differentiate more accurately between virulent alleles and avirulent alleles. The sensitivities of all the tests were all higher than 90% except for that for the marker SP.SNP.SC.241.57435 (78%) with the avirulence gene *AvYr9*. A marker with a higher sensitivity should provide a higher rate of correct predictions of virulent or avirulent phenotypes.

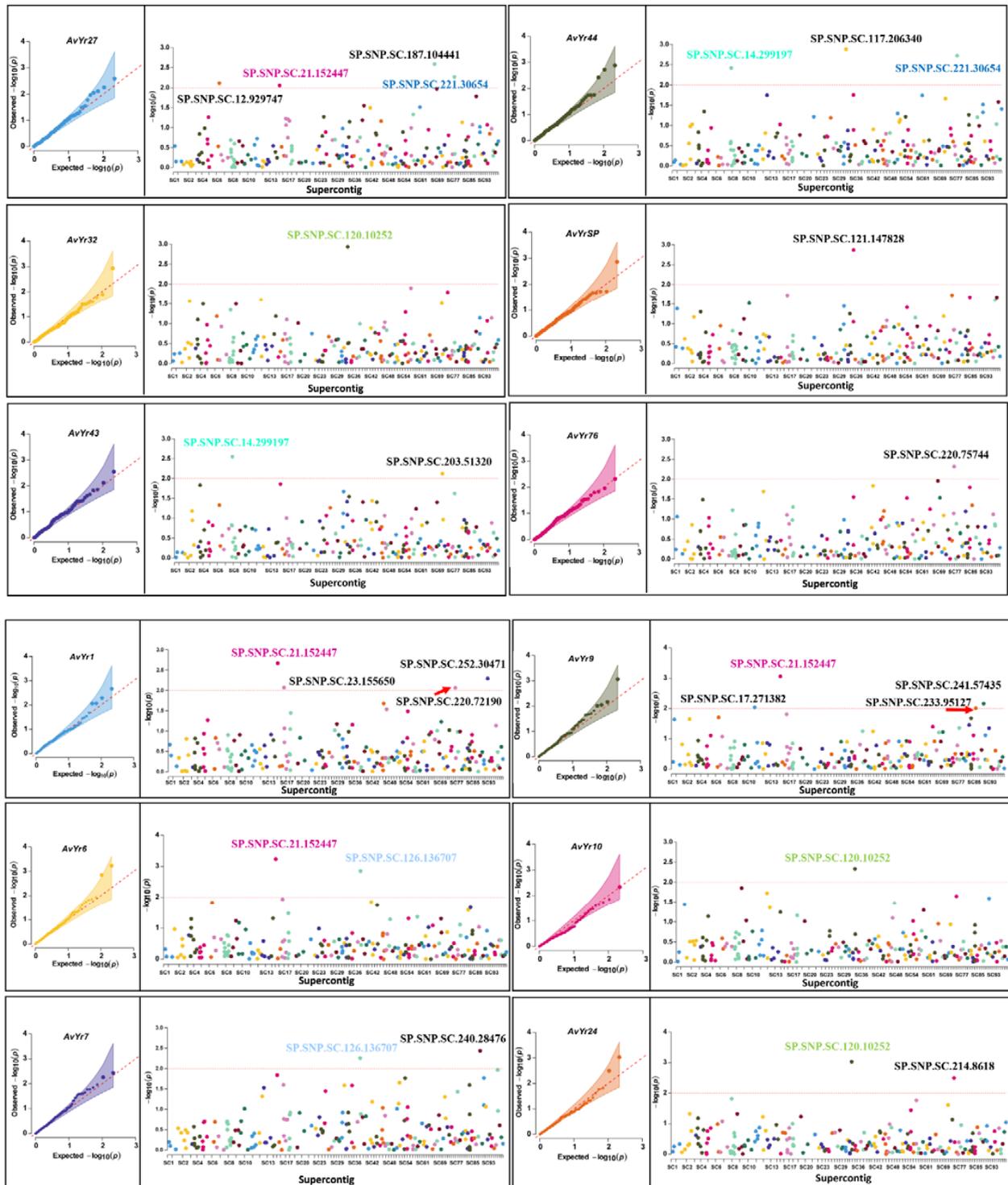


Figure 4. QQ plots and Manhattan plots of SP-SNP markers significantly associated with 12 avirulence (*Av*) genes. In the QQ plots, the X-axis represents the genomic position of SP-SNPs in the supercontigs of the PST-78 reference genome, and along the Y-axis are the $-\log_{10}$ transformed significance p -values. The red dashed lines represent the Bonferroni-corrected threshold $-\log_{10}(p)$ of 2.0. In the Manhattan plots, each dot represents a SP-SNP locus, and its genomic position is referred to the supercontig of the PST-78 reference genome. The SP-SNPs of same color are in the same supercontig.

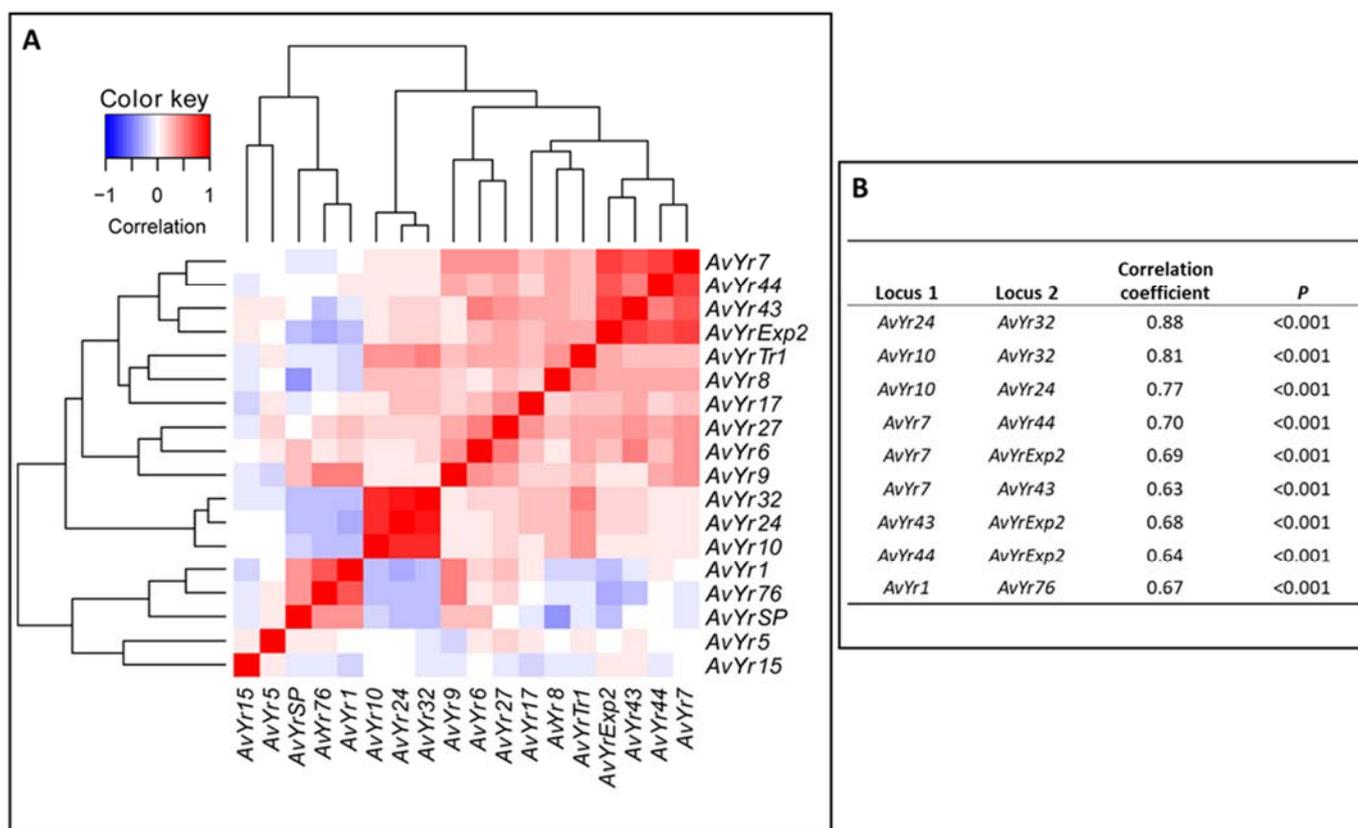


Figure 5. Correlation coefficients between 18 avirulence/virulence loci of *Puccinia striiformis* f. sp. *tritici* (A). The avirulence genes were symbolized as *AvYr1*, *AvYr5*, and so on, corresponding to their resistance *Yr* genes *Yr1*, *Yr5*, and so on. The correlations with coefficient values > 0.60 ($p < 0.001$) are listed in (B).

3. Discussion

The results of the present study confirmed that association analysis can be a powerful tool for identifying molecular markers associated with *Pst* avirulence genes, especially when the markers are developed from polymorphic SP-SNPs, as first reported by Xia et al. (2016). Different from the previous study that utilized non-selected isolates from two years and only from the U.S., the present study used 157 isolates selected from nine countries over an eight-year period. These isolates were identified as 126 races using 18 *Yr* single-gene lines, thus representing diverse avirulence/virulence profiles. The 157 isolates were also previously identified as 157 multi-locus genotypes (MLGs) using 14 simple sequence repeat (SSR) markers, representing major molecular groups [17,36]. The highly diverse isolates should be more suitable for association analysis for identifying markers associated with avirulence genes. From the 209 SP-SNP markers used to genotype the 157 isolates, we found 27 significant ($p < 0.01$) marker–trait associations involving 12 *AvYr* genes and 19 SP-SNP loci (Table 1). The number of avirulence genes with associated markers is relatively high compared with previous association analysis studies [23,35,37] and this can be attributed to the large number of SP-SNPs and the selected isolates that have relatively balanced virulence/avirulence ratios. Similar to previous studies, we also found that some markers were associated with different avirulence genes, suggesting that these avirulence genes are located in a gene cluster [23,24,35]. For example, the SP-SNP marker SP.SNP.SC.120.10252 was significantly associated with *AvYr10*, *AvYr24*, and *AvYr32*, which is supported by the phenotypic correlations between these avirulence genes [38,39].

In the present study, we identified SP-SNP markers for 12 of the 18 avirulence genes corresponding to the 18 *Yr* single genes in the differentials but did not find markers for the remaining 6 genes. As none of the *Pst* isolates were virulent to either *Yr5* or *Yr15*, it

was not possible to identify markers for their corresponding avirulence genes. However, it was possible to find markers for *AvYr8*, *AvYr17*, *AvYrTr1*, and *AvYrExp2*, as the 157 isolates showed relatively balanced virulence/avirulence ratios (Figure 1). Failing to identify markers for these genes may be due to the limitation of the 209 SP-SNPs used in the present study, which do not cover the entire genome. Secondly, only SP-SNPs showing co-dominant polymorphisms among the 14 whole-genome sequenced isolates used in the previous study [37] were used to design primers for SP-SNP markers. Avirulence/virulence alleles that have presence/absence or indel polymorphisms likely escaped from the test with these markers.

Among the 12 avirulence genes with associated SP-SNP markers, 7 had two or more markers in different supercontigs. The supercontigs were assigned based on the PST-78 reference genome, the best available at that time [37]. Some of the supercontigs may be linked but some of them may be far away. Further annotation of these SP-SNPs is needed to determine their genomic and molecular relationships. The genomic relationships of the markers associated with the same avirulence gene could be improved by annotation using the high-quality genomes recently established [40–42]. Nevertheless, different supercontig markers for a single avirulence gene may indicate that the avirulence phenotype is controlled by different genes in different genomic regions. Two-gene controlled virulence was demonstrated by genetic analysis of sexually produced *Pst* populations [23,24].

In an association study, individuals should be distinct [43]. However, as a predominantly asexually reproduced fungal pathogen, population structures of *Pst* have been reported in previous studies [17,36,44,45]. In a structured population, individuals with differences in allele frequencies between sub-populations due to ancestries which are unrelated to the trait of interest can cause false-positive associations in association studies [43]. To address this problem, PC analysis, which can effectively determine population structure, was used in the present study prior to the association analysis. However, PC analysis only accounts for fixed effects of genetic ancestry and does not account for relatedness between individuals. Therefore, the mixed-model approach, which utilizes both fixed effects (candidate SNPs and fixed covariates) and random effects (the genotypic covariance matrix) involving kinship and cryptic relatedness, was used in the further association analysis. In addition, false-positive associations might also be caused by statistical fluctuations governed by chance in multiple-hypothesis testing [46]. To control statistical fluctuations, various statistical approaches have been proposed, including Bonferroni correction and estimation of the false detection rate (FDR), which are two common correction methods [47,48]. In the present study, Bonferroni correction was applied for the suggestive threshold p -value ($p = 1/Ne$), where Ne represents the effective number of SNP markers [49]; therefore, $p = 1/Ne = 1/209 = 0.0047 \approx 0.01$ was chosen ($-\log_{10}(0.01) = 2$) as the threshold for significance in the association analysis. Besides these two approaches, replicating genotype–phenotype associations in larger and independent populations is another way of establishing the credibility of genome-wide association studies (GWAS) [50]. Therefore, in future studies, we will genotype a large number of *Pst* isolates with the SP-SNPs identified in the present study to confirm the associations and also genotype the isolates using additional SP-SNPs to identify more markers.

By using the SP-SNP markers, the *Pst* isolates in the present study were clustered into molecular groups similar to those in our previous studies, in which more isolates were genotyped using the set of 14 SSR markers [17,36]. For example, isolates EC15-039, EC15-019, EC15-022, EC15-26, EC15-27, EC15-30, and EC16-019, which were from Ecuador, were separated into a molecular group different from the groups of other isolates in the present study, consistent with the results of our previous study. However, the SP-SNP markers significantly associated with avirulence genes can provide information for virulence evolution and may be used to tag virulence genes [33,35,37]. Every year, we establish more than 300 isolates from stripe rust collections throughout the U.S. and phenotype them for virulence/avirulence profiles [14,15,39]. Genotyping the isolates with the SP-SNPs may confirm the avirulence-associated SP-SNPs identified in the present study

and previous studies [33,35,37], which may lead to the identification and cloning of the avirulence genes.

All SP-SNP markers used in the present study were developed based on the SP genes characterized by Xia et al. (2017) through analyzing the whole-genome sequences of 14 *Pst* isolates [37]. All SP genes were annotated for identification of polymorphic SNPs and predicted for effectors using the EffectorP program. In the present study, among the 19 SP-SNP markers significantly associated with *Avr* genes, 8 were predicted as *Pst* effectors with high confidence (Table 1), indicating their possible involvement in interactions with host plants [37]. However, most of the SP genes were predicted to encode hypothetical proteins (Table S3), while only a few were annotated as being involved in different biological processes. None of them is known to be involved in pathogenicity. One SP gene, PSTG_15874, which was significantly correlated with avirulence to *Yr9*, has a high homology with a phosphatidylinositol/phosphatidylglycerol transfer protein (PG/PI-TP) gene. The PG/PI-TP proteins belong to the ML (MD-2-related lipid-recognition) domain family and have been shown to bind phosphatidylglycerol and phosphatidylinositol, but the biological significance of these proteins is still not clear [51].

For association analysis, high resolution mapping depends on the number of markers as well as on linkage disequilibrium (LD) decay [46,48,52]. LD is the nonrandom association of alleles at different loci that plays a central role in association analysis [48]. Therefore, for establishing the credibility of the association analysis, the LD decay of the *Pst* genome also should be estimated [35]. Xia et al. (2020) has developed a high-quality map for *Pst* comprising 41 lineage groups, which will enable us to identify *Avr* candidates in narrow genome regions and study their functions [31]. In the present study, even though we identified more SP-SNP markers than the previous study [35], we were still unable to estimate LD decay because of unknown physical distances between them. Moreover, due to budgetary and time limitations, we only tried the primers of the first 390 of the more than 900 SP genes and used 209 successful markers in the present study. For future studies, we will use the remaining SP genes. Alternately, we can genome-sequence and RNA-sequence the isolates used in the present study to identify avirulence genes involved in interactions with the *Yr* genes.

The accuracy and sensitivity of these SP-SNP markers for detecting the associated avirulence/virulence genes should provide more information for the subsequent use of these markers in monitoring individual virulence factors and race changes in the pathogen population. The higher the accuracy and sensitivity of the markers, the more useful the markers are for predicting the virulence phenotype. For example, high accuracies were found in the tests of the same marker SP.SNP.SC.120.10252 associated with the avirulence genes *AvYr10* (92.86%), *AvYr24* (92.86%), and *AvYr32* (94.90%), as the majority of isolates are avirulent to the corresponding *Yr* genes. However, caution should be taken when using the markers. Further testing of the markers for the opposite alleles for the virulence phenotypes needs more isolates virulent to these *Yr* genes, which may take years to accumulate as virulence frequencies for these genes have been low in the past several decades [13–16,39].

The present study aimed to identify additional SP-SNP markers associated with avirulence genes from the sequences of SP genes. Virulence-associated markers can be used to monitor virulence and race changes in the pathogen population. However, the Ion Protein sequencing approach used in the present study for identifying SNPs is not efficient for routine tests for monitoring virulence changes. To overcome this drawback, SNP markers can be converted to Kompetitive Allele Specific PCR (KASP) markers, which should be used for efficiently monitoring *Pst* populations. The SP-SNP markers identified in the present study, together with the other nearly 100 SP-SNP markers associated with *Pst* virulence in previous studies [27,31,35,37,44], will be converted to KASP markers for further confirmation of the association of the SP-SNP markers with their avirulence genes and also for establishing a set of virulence-related markers for monitoring changes in both virulence and molecular genotypes in the *Pst* population. Compared to SNP genotyping, KASP markers are relatively cheap and easy to use, as demonstrated in our program for

stripe rust resistance genes in wheat [53–57]. The present study provides genomic resources for further marker development and research on host–pathogen interactions, as well as pathogen population dynamics.

4. Materials and Methods

4.1. Isolate Selection

A total of 157 *Pst* isolates used in the present study were selected based on races and MLGs from *Pst* collections from the U.S. and eight other countries assembled in the period 2010–2020 [13,17,36]. These isolates had different MLGs and represented 126 races (Table 2).

Table 2. Numbers of the *Puccinia striiformis* f. sp. *tritici* isolates and races from nine countries in 2010–2018 used in this study.

Country	No. of Isolates	Year	Races ^a
Canada	2	2017	PSTv-37, PSTv-14
China	12	2016	PSTv-225, PSTv-229, PSTv-230, PSTv-231, PSTv-250, PSTv-259, PSTv-267, PSTv-270, PSTv-274, PSTv-277, PSTv-278, PSTv-280
Ecuador	13	2015/2016	PSTv-20, PSTv-106, PSTv-221, PSTv-285, PSTv-286, PSTv-287, PSTv-289, PSTv-294, PSTv-298, PSTv-303, PSTv-305, PSTv-306, PSTv-327
Egypt	2	2018	PSTv-120, PSTv-15
Ethiopia	11	2014	PSTv-41, PSTv-47, PSTv-76, PSTv-105, PSTv-106, PSTv-107, PSTv-110, PSTv-116
Italy	18	2014/2016/2017	PSTv-121, PSTv-125, PSTv-127, PSTv-129, PSTv-130, PSTv-131, PSTv-132, PSTv-133, PSTv-134, PSTv-135, PSTv-136, PSTv-137, PSTv-192, PSTv-232, PSTv-295, PSTv-317, PSTv-320
Mexico	8	2015/2016	PSTv-53, PSTv-78, PSTv-109, PSTv-198, PSTv-252, PSTv-292, PSTv-296, PSTv-307
Pakistan	4	2012	PSTv-11, New, PSTv-37, New
USA	87	2010/2011/2012/ 2013/2014/2015/ 2016/2017	PSTv-1, PSTv-2, PSTv-3, PSTv-4, PSTv-6, PSTv-7, PSTv-8, PSTv-11, PSTv-14, PSTv-15, PSTv-16, PSTv-17, PSTv-18, PSTv-19, PSTv-20, PSTv-22, PSTv-23, PSTv-24, PSTv-25, PSTv-27, PSTv-28, PSTv-29, PSTv-31, PSTv-32, PSTv-33, PSTv-34, PSTv-35, PSTv-37, PSTv-39, PSTv-40, PSTv-41, PSTv-42, PSTv-43, PSTv-44, PSTv-45, PSTv-46, PSTv-47, PSTv-48, PSTv-52, PSTv-53, PSTv-64, PSTv-65, PSTv-67, PSTv-71, PSTv-72, PSTv-73, PSTv-74, PSTv-75, PSTv-76, PSTv-77, PSTv-78, PSTv-79, PSTv-101, PSTv-109, PSTv-120, PSTv-121, PSTv-122, PSTv-123, PSTv-144, PSTv-175, PSTv-198, PSTv-201, PSTv-214, PSTv-221, PSTv-239, PSTv-284, PSTv-293, PSTv-321, PSTv-322

^a The races were identified using the 18 *Yr* single-gene lines as differentials [13–15,39].

4.2. Virulence Data

The *Pst* isolates were tested for their avirulence/virulence patterns using the 18 *Yr* single-gene lines, and the virulence data have been reported previously [13–15,39]. The avirulence or virulence of isolates on a particular *Yr* resistance gene line was represented by an infection type (IT), which was scored using a 0-to-9 scale, with 0 as the most avirulent and 9 as the most virulent [58]. For the association analysis in the present study, the phenotype of an isolate for a particular *Yr* gene was defined as avirulent when IT was 0–6 and virulent when IT was 7–9 [14]. To reduce phenotypic variation within avirulent and virulent classes, isolates with ITs 0–2 for avirulent reactions and with ITs 7–9 for virulent reactions were selected, and thus intermediate ITs (3–6) were mostly avoided (Figure 1, Table S1). The 18 resistance genes were *Yr1*, *Yr5*, *Yr6*, *Yr7*, *Yr8*, *Yr9*, *Yr10*, *Yr15*, *Yr17*, *Yr24*, *Yr27*, *Yr32*, *Yr43*, *Yr44*, *YrExp2*, *YrSP*, *YrTr1*, and *Yr76*. Their corresponding avirulence genes were

symbolized as *AvYr1*, *AvYr5*, and so on. Generally, the selected 157 isolates represented a relatively balanced virulent–avirulent profile for the majority of the *Yr* genes.

4.3. DNA Extraction

DNA extraction from urediniospores of the *Pst* isolates was described in our previous study [36]. The concentration of the DNA stock solution was determined using a ND-1000 spectrophotometer (Bio-Rad, Hercules, CA, USA), and the quality was checked in a 0.8% agarose gel. A work solution of 0.5 ng μL^{-1} was made from the stock solution by adding sterile deionized water for use as a DNA template in polymerase chain reaction (PCR).

4.4. Development of SP-SNP Markers

The genomic sequences of *Pst*-specific SP genes containing SNPs among 14 whole-genome sequenced *Pst* isolates [37] identified using the IGV software (<https://igv.org/app>, accessed on 27 September 2020) were used to develop SP-SNP markers. The SP-SNP primers were designed using the Sequenom MassArray Assay Design 4.0 software (Sequenom, San Diego, CA, USA). The primers were modified by adding barcodes and specific sequences compatible with the Ion Torrent Proton System (LifeTechnologies, Carlsbad, CA, USA). The locus-specific forward primers for the first round of PCR were tailed with an M13-derived sequence (GATGTAACGACGGCCAGTG) at the 5'-end to enable the addition of barcoded adapters during the second round of PCR. The Ion truncated P1 adapter sequence (CCTCTCTATGGCAGTCGGTGAT) was concatenated to the 5'-end of the locus-specific reverse primers (Table S5). For the second round of PCR, the forward fusion primer consisted of, from 5' to 3', the standard Ion A adapter sequence (CCATCTCATCCCTGCGTGTCTCCGACTCAG), a unique barcode with 10–12 nucleotides, followed by the M13 tail sequence (Table S5). A combination of different barcodes with the M13 tail proved the flexibility required to multiplex the same set of markers in different samples. The reverse primer for the second round of PCR was the Ion truncated P1 adapter sequence.

4.5. Isolate Genotyping

Prior to sequencing, library construction, sample purification, size selection, and quantification were conducted using the standard procedures [59–61]. The library construction included two steps of PCR. In the first step of PCR, each reaction (10 μL) contained 1.0 μL of McLab 10 \times Taq PCR buffer (McLab, San Francisco, CA, USA), 0.45 μL of 25 mM MgCl_2 , 1 μL 5 mM dNTP, 1 μL of 125 nM primer pool, 0.2 μL 5 U/ μL McLab HoTaq polymerase (McLab), 2 μL of 0.5 ng/ μL DNA (total 1 ng), and 4.35 μL sterile ddH₂O. The amplification cycles and conditions were 94 °C for 1 min for initial denaturation; 35 cycles of 94 °C for 20 s, 56 °C for 2 min, and 68 °C for 30 s; and 3 min of final extension at 72 °C. The first-step PCR products were diluted at 1:1 with ddH₂O into a 96-well plate for the second step of PCR. In the second step of PCR, each reaction (6 μL) contained 0.5 μL of McLab 10 \times Taq PCR buffer, 0.2 μL of 25 mM MgCl_2 , 0.025 μL 100 mM dNTP, 0.2 μL P1 reverse primer, 0.2 μL 5 U/ μL McLab HoTaq polymerase, 2 μL of diluted PCR products from the first-step PCR, 0.875 μL sterile ddH₂O, and 2 μL of 2 μM barcoded adapters 289–384 and 385–480. The amplification cycles and conditions were 94 °C for 1 min for initial denaturation; 15 cycles of 94 °C for 15 s, 60 °C for 30 s, and 72 °C for 1 min; and 3 min of final extension at 72 °C.

The second-step PCR products were cleaned using a QIAquick PCR Purification kit (Qiagen, Hilden, Germany). The size selection of the libraries was first performed on a 4% E-Gel SizeSelect Gel (Life Technologies, Carlsbad, CA, USA) and then on a 2% E-Gel1 SizeSelect™ Gel (LifeTechnologies) to select PCR fragments from 140 bp to 250 bp. Purified libraries were quantified using a Qubit1 dsDNA HS assay kit (LifeTechnologies), diluted to the appropriate concentration as recommended by LifeTechnologies (minimum 80 pmol/L) and a size distribution between 185 bp to 260 bp. The diluted sample libraries were prepared for sequencing using an Ion Torrent Proton NGS platform (Life Technologies).

4.6. Data Analyses

The IT distribution and density of isolates on the 18 *Yr* single-gene differentials were determined using the *vioplot* package in the R program 4.1.1. The genotypic data from an initial set of 390 SP-SNP markers were subjected for quality control by excluding the markers with >50% missing data and a minor allele frequency (MAF) less than 0.05. The filtered set of 209 SP-SNP markers was finally retained, and the remaining missing data were imputed using the Genome Association and Prediction Integrated Tool (GAPIT) in the R program 4.1.1.

Prior to the association analysis, principal component analysis (PCA) was performed to determine the optimal genetic clusters within GAPIT [62,63] to reduce false positives caused by population structures. However, PCA only accounts for fixed effects of genetic ancestry and does not account for relatedness between individuals. Therefore, a mixed-model approach [62], which used both fixed effects (candidate SNPs and fixed covariates) and random effects (the genotypic covariance matrix) involving kinship and cryptic relatedness, was used in the association analysis. The VanRaden method [64] was used to estimate the relationships among isolates by computing a kinship matrix.

To further confirm the clusters of genetically related isolates based on both the 209 SP-SNP markers used in the present study and also to compare the clusters with the 154 isolates using the 14 SSR markers in the previous study [17,36], a hierarchical cluster analysis was conducted using the dissimilarity values and the “ward.D2” method with the “hclust” function in the R stats 4.1.1 program [65].

Associations between SP-SNPs and the 18 virulence/avirulence traits of the *Pst* isolates were analyzed in the program GAPIT, using the commands “myGAPIT <- GAPIT (Y = myY, G = myG, PCA.total = “”, kinship.cluster = c(“ward.D2”), kinship.group = c(“Mean”), model = “MLM”, multiple_analysis = TRUE, NJtree.group = “”, and Geno.View.output = FALSE, file.output = T)”. The markers were identified as significant if the *p*-value was equal to or less than 0.01 and $-\log_{10}(p)$ equal to or greater than 2 after Bonferroni correction [66]. The Manhattan plots were drawn using the ‘CMplot’ package in the R program 4.1.1. The correlation relationships among the 18 virulence phenotypic traits were estimated using the ‘heatmap2’ package in the R program 4.1.1.

To determine how reliable the SP-SNP markers identified from the association analysis for detecting the avirulence/virulence phenotypes were, accuracies and sensitivities coupled with their 95% confidence intervals were calculated [67]. For this analysis, accuracy was measured as the proportion of correct predictions among all the predictions and sensitivity was measured as the proportion of predictions correctly identified by the test, these values showing how good the test was for detecting the avirulence or virulence phenotypes. If isolates had the avirulence genotype and the marker allele for avirulence or vice versa, the predictions of the phenotypes of the isolates by the marker were considered correct predictions (CPs). If the isolates had unmatched marker alleles and phenotypes, the marker predictions of phenotypes were considered incorrect predictions (ICPs). Accuracy = (Number of correct predictions)/(Number of all predictions). Sensitivity = (Number of correct predictions that are correctly identified by the test)/(Number of all correct predictions).

5. Conclusions

In this study, we identified 19 SP-SNP markers significantly associated with 12 avirulence genes: *AvYr1*, *AvYr6*, *AvYr7*, *AvYr9*, *AvYr10*, *AvYr24*, *AvYr27*, *AvYr32*, *AvYr43*, *AvYr44*, *AvYrSP*, and *AvYr76*. Some SP-SNP markers were significantly associated with two or more avirulence genes, suggesting that these avirulence loci are located in an avirulence gene cluster, consistent with our previous studies. The present study further confirmed that association analysis in combination with SP-SNP markers is powerful when it comes to identifying markers for avirulence genes. The virulence-related SP-SNP markers in the present study, as well as other SP-SNP markers identified in our previous study, should be useful in developing a set of virulence-related markers for monitoring changes in virulence and genotypes in *Pst* populations. The genomic resources developed in the present

study can be used for further research on host–pathogen interactions as well as pathogen population dynamics.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms23084114/s1>.

Author Contributions: Conceptualization, X.C. and Q.B.; methodology, Q.B., C.X., M.W. and D.R.S.; software, Q.B. and D.R.S.; validation, Q.B., M.W. and X.C.; formal analysis, Q.B.; investigation, Q.B. and M.W.; resources, X.C. and D.R.S.; data curation, Q.B.; writing—original draft preparation, Q.B.; writing—review and editing, X.C.; visualization, Q.B.; supervision, X.C.; project administration, X.C.; funding acquisition, X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the U.S. Department of Agriculture, Agricultural Research Service (Project No. 2090-22000-018-00D), Washington Grain Commission (Project No. 13C-3061-3144), and Washington State University, Department of Plant Pathology, College of Agricultural, Human, and Natural Resource Sciences, Agricultural Research Center, HATCH Project Number WNP00461.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the Supplementary Materials.

Acknowledgments: The authors would like to thank Tobin Peever, Scot Hulbert, and Weidong Chen for their suggestions, Travis Ruff for technical assistance, and the reviewers for their comments which helped to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Disclaimer: Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. USDA is an equal opportunity provider and employer.

Abbreviations

SP	Secreted protein
SNP	Single-nucleotide polymorphism
<i>Pst</i>	<i>Puccinia striiformis</i> f. sp. <i>tritici</i>
Yr	Yellow rust
QTL	Quantitative trait loci
EMS	Ethyl methanesulfonate
MLG	Multi-locus genotype
SSR	Simple sequence repeat
IT	Infection type
PCR	Polymerase Chain Reaction
CA	California
USA	United States of America
MAF	Minor allele frequency
GAPIT	Genome Association and Prediction Integrated Tool
P	Probability value
PCA	Principal component analysis
PC	Principal component
MG	Molecular groups
EMMA	Efficient Mixed Model Association
PVE	Percentage of variation explained
FDR	False detection rate
GWAS	Genome-wide association study
LD	Linkage disequilibrium
KASP	Kompetitive Allele Specific PCR

References

1. Chen, X.M. Epidemiology and control of stripe rust [*Puccinia striiformis* f. sp. *tritici*] on wheat. *Can. J. Plant. Pathol.* **2005**, *27*, 314–337. [[CrossRef](#)]
2. Chen, X.M. Pathogens which threaten food security: *Puccinia striiformis*, the wheat stripe rust pathogen. *Food Secur.* **2020**, *12*, 239–251. [[CrossRef](#)]
3. Chen, X.M.; Kang, Z.S. History of research, symptoms, taxonomy of the pathogen, host range, distribution, and impact of striperust. In *Stripe Rust*; Chen, X.M., Kang, Z.S., Eds.; Springer: Dordrecht, The Netherlands, 2017; pp. 1–33.
4. Stubbs, R.W. *Stripe Rust: The Cereal Rusts II: Disease, Distribution, Epidemiology and Control*; Roelfs, A.P., Bushnell, W.R., Eds.; Academic Press, Inc.: New York, NY, USA, 1985; pp. 61–101.
5. Wellings, C.R. Global status of stripe rust: A review of historical and current threats. *Euphytica* **2011**, *179*, 129–141. [[CrossRef](#)]
6. Ali, S.; Gladieux, P.; Leconte, M.; Gautier, A.; Justesen, A.F.; Hovmøller, M.S.; Enjalbert, J.; de Vallavieille-Pope, C. Origin, migration routes and worldwide population genetic structure of the wheat yellow rust pathogen *Puccinia striiformis* f.sp. *tritici*. *PLoS Pathog.* **2014**, *10*, e1003903. [[CrossRef](#)] [[PubMed](#)]
7. Ali, S.; Rodriguez-Algaba, J.; Thach, T.; Sørensen, C.K.; Hansen, J.G.; Lassen, P.; Nazari, K.; Hodson, D.P.; Justesen, A.F.; Hovmøller, M.S. Yellow rust epidemics worldwide were caused by pathogen races from divergent genetic lineages. *Front. Plant Sci.* **2017**, *8*, 1057. [[CrossRef](#)] [[PubMed](#)]
8. Boshoff, W.; Pretorius, Z.; Van Niekerk, B. Establishment, distribution, and pathogenicity of *Puccinia striiformis* f. sp. *tritici* in South Africa. *Plant Dis.* **2002**, *86*, 485–492. [[CrossRef](#)]
9. Hovmøller, M.S.; Yahyaoui, A.H.; Milus, E.A.; Justesen, A.F. Rapid global spread of two aggressive strains of a wheat rust fungus. *Mol. Ecol.* **2008**, *17*, 3818–3826. [[CrossRef](#)]
10. Hovmøller, M.; Walter, S.; Bayles, R.; Hubbard, A.; Flath, K.; Sommerfeldt, N.; Leconte, M.; Czembor, P.; Rodriguez-Algaba, J.; Thach, T. Replacement of the European wheat yellow rust population by new races from the centre of diversity in the near-Himalayan region. *Plant Pathol.* **2016**, *65*, 402–411. [[CrossRef](#)]
11. Hovmøller, M.; Justesen, A.; Brown, J. Clonality and long-distance migration of *Puccinia striiformis* f. sp. *tritici* in north-west Europe. *Plant Pathol.* **2002**, *51*, 24–32. [[CrossRef](#)]
12. Milus, E.A.; Kristensen, K.; Hovmøller, M.S. Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f. sp. *tritici* causing stripe rust of wheat. *Phytopathology* **2009**, *99*, 89–94. [[CrossRef](#)]
13. Chen, X.M.; Wang, M.N.; Wan, A.M.; Bai, Q.; Li, M.J.; López, P.F.; Maccaferri, M.; Mastrangelo, A.M.; Barnes, C.W.; Cruz, D.F.C.; et al. Virulence characterization of *Puccinia striiformis* f. sp. *tritici* collections from six countries in 2013 to 2020. *Can. J. Plant. Pathol.* **2021**, *43*, S308–S322.
14. Wan, A.M.; Chen, X.M. Virulence Characterization of *Puccinia striiformis* f. sp. *tritici* Using a new set of Yr single-gene line differentials in the United States in 2010. *Plant Dis.* **2014**, *98*, 1534–1542. [[CrossRef](#)] [[PubMed](#)]
15. Wan, A.M.; Chen, X.M.; Yuen, J. Races of *Puccinia striiformis* f. sp. *tritici* in the United States in 2011 and 2012 and Comparison with Races in 2010. *Plant Dis.* **2016**, *100*, 966–975. [[CrossRef](#)] [[PubMed](#)]
16. Liu, T.L.; Wan, A.M.; Liu, D.C.; Chen, X.M. Changes of races and virulence genes in *Puccinia striiformis* f. sp. *tritici*, the wheat stripe rust pathogen, in the United States from 1968 to 2009. *Plant Dis.* **2017**, *101*, 1522–1532. [[CrossRef](#)]
17. Bai, Q.; Wan, A.M.; Wang, M.N.; See, D.R.; Chen, X.M. Molecular characterization of wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) collections from nine countries. *Int. J. Mol. Sci.* **2021**, *22*, 9457. [[CrossRef](#)]
18. Flor, H.H. Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **1971**, *9*, 275–296. [[CrossRef](#)]
19. Catanzariti, A.M.; Dodds, P.N.; Lawrence, G.J.; Ayliffe, M.A.; Ellis, J.G. Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* **2006**, *18*, 243–256. [[CrossRef](#)]
20. Dodds, P.N.; Lawrence, G.J.; Catanzariti, A.M.; Teh, T.; Wang, C.I.; Ayliffe, M.A.; Kobe, B.; Ellis, J.G. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8888–8893. [[CrossRef](#)]
21. Dodds, P.N.; Lawrence, G.J.; Catanzariti, A.M.; Ayliffe, M.A.; Ellis, J.G. The *Melampsora lini AvrL567* avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* **2004**, *16*, 755–768. [[CrossRef](#)]
22. Pernaci, M.; De Mita, S.; Andrieux, A.; Pétrowski, J.; Halkett, F.; Duplessis, S.; Frey, P. Genome-wide patterns of segregation and linkage disequilibrium: The construction of a linkage genetic map of the poplar rust fungus *Melampsora larici-populina*. *Front. Plant Sci.* **2014**, *5*, 454. [[CrossRef](#)]
23. Xia, C.J.; Lei, Y.; Wang, M.N.; Chen, W.Q.; Chen, X.M. An avirulence gene cluster in the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) identified through genetic mapping and whole-genome sequencing of a sexual population. *MSphere* **2020**, *5*, e00128-20. [[CrossRef](#)] [[PubMed](#)]
24. Yuan, C.Y.; Wang, M.N.; Skinner, D.Z.; See, D.R.; Xia, C.J.; Guo, X.H.; Chen, X.M. Inheritance of virulence, construction of a linkage map, and mapping dominant virulence genes in *Puccinia striiformis* f. sp. *tritici* through characterization of a sexual population with genotyping-by-sequencing. *Phytopathology* **2018**, *108*, 133–141. [[CrossRef](#)] [[PubMed](#)]
25. Kamoun, S. Groovy times: Filamentous pathogen effectors revealed. *Curr. Opin. Plant Biol.* **2007**, *10*, 358–365. [[CrossRef](#)] [[PubMed](#)]
26. Petre, B.; Joly, D.L.; Duplessis, S. Effector proteins of rust fungi. *Front. Plant Sci.* **2014**, *5*, 416. [[CrossRef](#)] [[PubMed](#)]

27. Li, Y.X.; Xia, C.J.; Wang, M.N.; Yin, C.T.; Chen, X.M. Whole-genome sequencing of *Puccinia striiformis* f. sp. *tritici* mutant isolates identifies avirulence gene candidates. *BMC Genom.* **2020**, *21*, 454. [[CrossRef](#)] [[PubMed](#)]
28. Li, Y.X.; Wang, M.N.; See, D.R.; Chen, X.M. Ethyl-methanesulfonate mutagenesis generated diverse isolates of *Puccinia striiformis* f. sp. *tritici*, the wheat stripe rust pathogen. *World J. Microbiol. Biotechnol.* **2019**, *35*, 28. [[CrossRef](#)] [[PubMed](#)]
29. Dagvadorj, B.; Ozketen, A.C.; Andac, A.; Duggan, C.; Bozkurt, T.O.; Akkaya, M.S. A *Puccinia striiformis* f. sp. *tritici* secreted protein activates plant immunity at the cell surface. *Sci. Rep.* **2017**, *7*, 1141. [[CrossRef](#)]
30. Zhao, M.X.; Wang, J.F.; Ji, S.; Chen, Z.J.; Xu, J.H.; Tang, C.L.; Chen, S.T.; Kang, Z.S.; Wang, X.J. Candidate effector Pst_8713 impairs the plant immunity and contributes to virulence of *Puccinia striiformis* f. sp. *tritici*. *Front. Plant Sci.* **2018**, *9*, 1294. [[CrossRef](#)]
31. Yang, Q.; Huai, B.Y.; Lu, Y.X.; Cai, K.Y.; Guo, J.; Zhu, X.G.; Kang, Z.S.; Guo, J. A stripe rust effector Pst18363 targets and stabilises TaNUDX23 that promotes stripe rust disease. *New Phytol.* **2020**, *225*, 880–895. [[CrossRef](#)]
32. Liu, B.; Chen, X.M.; Kang, Z.S. Gene sequencing reveals heterokaryotic variations and evolutionary mechanisms in *Puccinia striiformis*. *Open J. Genet.* **2012**, *1*, 253925.
33. Xia, C.J.; Wan, A.M.; Wang, M.N.; Jiwan, D.A.; See, D.R.; Chen, X.M. Secreted protein gene derived-single nucleotide polymorphisms (SP-SNPs) reveal population diversity and differentiation of *Puccinia striiformis* f. sp. *tritici* in the United States. *Fungal Biol.* **2016**, *120*, 729–744. [[CrossRef](#)] [[PubMed](#)]
34. Lei, Y.; Wang, M.N.; Wan, A.M.; Xia, C.J.; See, D.R.; Zhang, M.; Chen, X.M. Virulence and molecular characterization of experimental isolates of the stripe rust pathogen (*Puccinia striiformis*) indicate somatic recombination. *Phytopathology* **2017**, *107*, 329–344. [[CrossRef](#)] [[PubMed](#)]
35. Xia, C.J.; Wang, M.N.; Wan, A.M.; Jiwan, D.A.; See, D.R.; Chen, X.M. Association analysis of SP-SNPs and avirulence genes in *Puccinia striiformis* f. sp. *tritici*, the wheat stripe rust pathogen. *Am. J. Plant Sci.* **2016**, *7*, 126. [[CrossRef](#)]
36. Bai, Q.; Wan, A.M.; Wang, M.N.; See, D.R.; Chen, X.M. Population diversity, dynamics, and differentiation of wheat stripe rust pathogen *Puccinia striiformis* f. sp. *tritici* from 2010 to 2017 and comparison with 1968 to 2009 in the United States. *Front. Microbiol.* **2021**, *12*, 696835. [[CrossRef](#)] [[PubMed](#)]
37. Xia, C.J.; Wang, M.N.; Cornejo, O.E.; Jiwan, D.A.; See, D.R.; Chen, X.M. Secretome characterization and correlation analysis reveal putative pathogenicity mechanisms and identify candidate avirulence genes in the wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici*. *Front. Microbiol.* **2017**, *8*, 2394. [[CrossRef](#)]
38. Wan, A.M.; Wang, X.J.; Kang, Z.S.; Chen, X.M. Variability of the stripe rust pathogen. In *Stripe Rust*; Springer: New York, NY, USA, 2017; pp. 35–154.
39. Wang, M.N.; Wan, A.M.; Chen, X.M. Race characterization of *Puccinia striiformis* f. sp. *tritici* in the United States from 2013 to 2017. *Plant Dis.* **2022**. [[CrossRef](#)]
40. Li, Y.X.; Xia, C.J.; Wang, M.N.; Yin, C.T.; Chen, X.M. Genome sequence resource of a *Puccinia striiformis* isolate infecting wheatgrass. *Phytopathology* **2019**, *109*, 1509–1512. [[CrossRef](#)]
41. Schwessinger, B.; Sperschneider, J.; Cuddy, W.S.; Garnica, D.P.; Miller, M.E.; Taylor, J.M.; Dodds, P.N.; Figueroa, M.; Park, R.F.; Rathjen, J.P. A near-complete haplotype-phased genome of the dikaryotic wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici* reveals high interhaplotype diversity. *mBio* **2018**, *9*, e02275-17. [[CrossRef](#)]
42. Xia, C.J.; Wang, M.N.; Yin, C.T.; Cornejo, O.E.; Hulbert, S.H.; Chen, X.M. Genome sequence resources for the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) and the barley stripe rust pathogen (*Puccinia striiformis* f. sp. *hordei*). *Mol. Plant-Microbe Interact.* **2018**, *31*, 1117–1120. [[CrossRef](#)]
43. Balding, D.J. A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* **2006**, *7*, 781–791. [[CrossRef](#)]
44. Liu, T.L.; Bai, Q.; Wang, M.N.; Li, Y.X.; Wan, A.M.; See, D.R.; Xia, C.J.; Chen, X.M. Genotyping *Puccinia striiformis* f. sp. *tritici* Isolates with SSR and SP-SNP Markers Reveals Dynamics of the Wheat Stripe Rust Pathogen in the United States from 1968 to 2009 and Identifies Avirulence Associated Markers. *Phytopathology* **2021**, *111*, 1828–1839. [[CrossRef](#)] [[PubMed](#)]
45. Sharma-Poudyal, D.; Bai, Q.; Wan, A.M.; Wang, M.N.; See, D.; Chen, X.M. Molecular characterization of international collections of the wheat stripe rust pathogen *Puccinia striiformis* f. sp. *tritici* reveals high diversity and intercontinental migration. *Phytopathology* **2020**, *110*, 933–942. [[CrossRef](#)] [[PubMed](#)]
46. Hirschhorn, J.N.; Daly, M.J. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* **2005**, *6*, 95–108. [[CrossRef](#)] [[PubMed](#)]
47. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [[CrossRef](#)]
48. Flint-Garcia, S.A.; Thornsberry, J.M.; Buckler IV, E.S. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* **2003**, *54*, 357–374. [[CrossRef](#)]
49. Li, M.X.; Yeung, J.M.; Cherny, S.S.; Sham, P.C. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum. Genet.* **2012**, *131*, 747–756. [[CrossRef](#)]
50. Chanock, S.J.; Manolio, T.; Boehnke, M.; Boerwinkle, E.; Hunter, D.J.; Thomas, G.; Hirschhorn, J.N.; Abecasis, G.R.; Altshuler, D.; Bailey-Wilson, J.E.; et al. Replicating genotype-phenotype associations. *Nature* **2007**, *447*, 655–660.
51. Record, E.; Moukha, S.; Asther, M. Characterization and expression of the cDNA encoding a new kind of phospholipid transfer protein, the phosphatidylglycerol/phosphatidylinositol transfer protein from *Aspergillus oryzae*: Evidence of a putative membrane targeted phospholipid transfer protein in fungi. *Biochim. Biophys. Acta Gene Regul. Mech.* **1999**, *1444*, 276–282.

52. Wang, W.Y.; Barratt, B.J.; Clayton, D.G.; Todd, J.A. Genome-wide association studies: Theoretical and practical concerns. *Nat. Rev. Genet.* **2005**, *6*, 109–118. [[CrossRef](#)]
53. Mu, J.M.; Liu, L.; Liu, Y.; Wang, M.N.; See, D.R.; Han, D.; Chen, X.M. Genome-wide association study and gene specific markers identified 51 genes or QTL for resistance to stripe rust in US winter wheat cultivars and breeding lines. *Front. Plant Sci.* **2020**, *11*, 998. [[CrossRef](#)]
54. Liu, L.; Yuan, C.Y.; Wang, M.N.; See, D.R.; Chen, X.M. Mapping Quantitative Trait Loci for high-temperature adult-plant resistance to stripe rust in spring wheat PI 197734 using a doubled haploid population and genotyping by multiplexed sequencing. *Front. Plant Sci.* **2020**, *11*, 596962. [[CrossRef](#)] [[PubMed](#)]
55. Liu, L.; Yuan, C.Y.; Wang, M.N.; See, D.R.; Zemetra, R.; Chen, X.M. QTL analysis of durable stripe rust resistance in the North American winter wheat cultivar Skiles. *Theor. Appl. Genet.* **2019**, *132*, 1677–1691. [[CrossRef](#)] [[PubMed](#)]
56. Liu, L.; Wang, M.N.; Zhang, Z.W.; See, D.R.; Chen, X.M. Identification of stripe rust resistance loci in US spring wheat cultivars and breeding lines using genome-wide association mapping and *Yr* gene markers. *Plant Dis.* **2020**, *104*, 2181–2192. [[CrossRef](#)] [[PubMed](#)]
57. Liu, L.; Wang, M.N.; Feng, J.Y.; See, D.R.; Chao, S.; Chen, X.M. Combination of all-stage and high-temperature adult-plant resistance QTL confers high-level, durable resistance to stripe rust in winter wheat cultivar Madsen. *Theor. Appl. Genet.* **2018**, *131*, 1835–1849. [[CrossRef](#)]
58. Line, R.F.; Qayoum, A. *Virulence, Aggressiveness, Evolution, and Distribution of Races of Puccinia striiformis (the Cause of Stripe Rust of Wheat) in North America, 1968–1987*[Line, Roland F., and Abdul Qayoum]; U.S. Department of Agriculture, Agricultural Research Service and National Technical Information Service: New York, NY, USA, 1992.
59. Bernardo, A.; Wang, S.; St. Amand, P.; Bai, G. Using next generation sequencing for multiplexed trait-linked markers in wheat. *PLoS ONE* **2015**, *10*, e0143890. [[CrossRef](#)]
60. Ruff, T.M.; Marston, E.J.; Eagle, J.D.; Sthapit, S.R.; Hooker, M.A.; Skinner, D.Z.; See, D.R. Genotyping by multiplexed sequencing (GMS): A customizable platform for genomic selection. *PLoS ONE* **2020**, *15*, e0229207. [[CrossRef](#)]
61. Liu, S. Genotyping by Multiplexing Amplicon Sequencing. 2015. Available online: <http://schnablelab.plant-genomics.iastate.edu/docs/resources/protocols/pdf/GBMAS-20150706.pdf> (accessed on 26 May 2020).
62. Tang, Y.; Liu, X.L.; Wang, J.B.; Li, M.; Wang, Q.S.; Tian, F.; Su, Z.B.; Pan, Y.; Liu, D.; Lipka, A.E. GAPIT version 2: An enhanced integrated tool for genomic association and prediction. *Plant Genome* **2016**, *9*. [[CrossRef](#)]
63. Wang, J.B.; Zhang, Z.W. GAPIT Version 3: Boosting power and accuracy for genomic association and prediction. *Genom. Proteom. Bioinform.* **2021**, *19*, 629–640. [[CrossRef](#)]
64. Lipka, A.E.; Tian, F.; Wang, Q.S.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.A.; Buckler, E.S.; Zhang, Z.W. GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **2012**, *28*, 2397–2399. [[CrossRef](#)]
65. Murtagh, F.; Legendre, P. Ward’s hierarchical agglomerative clustering method: Which algorithms implement Ward’s criterion? *J. Classif.* **2014**, *31*, 274–295. [[CrossRef](#)]
66. Bland, J.M.; Altman, D.G. Multiple significance tests: The Bonferroni method. *BMJ* **1995**, *310*, 170. [[CrossRef](#)] [[PubMed](#)]
67. Zhu, W.; Zeng, N.; Wang, N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. In *NESUG Proceedings; Health Care and Life Sciences*: Baltimore, MD, USA, 2010; Volume 19, p. 67.