AMERICAN SOCIETY FOR MICROBIOLOGY | mSphere®

RESEARCH ARTICLE
Host-Microbe Biology

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

Chongjing Xia,[a,b,c] Yu Lei,[c,d] Meinan Wang,[c] Wanquan Chen,[b] Xianming Chen[c,e]

[a]Wheat Research Institute, School of Life Sciences and Engineering, Southwest University of Science and Technology, Mianyang, Sichuan, China
[b]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China
[c]Department of Plant Pathology, Washington State University, Pullman, Washington, USA
[d]College of Biological Engineering, Sichuan University of Science and Engineering, Zigong, Sichuan, China
[e]Wheat Health, Genetics, and Quality Research Unit, Agriculture Research Service, U.S. Department of Agriculture, Pullman, Washington, USA

Chongjing Xia and Yu Lei contributed equally to this work. Author order was determined on the basis of seniority.

ABSTRACT *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe (yellow) rust, is an obligate, biotrophic fungus. It was difficult to study the genetics of the pathogen due to the lack of sexual reproduction. The recent discovery of alternate hosts for *P. striiformis* f. sp. *tritici* makes it possible to study inheritance and map genes involved in its interaction with plant hosts. To identify avirulence (*Avr*) genes in *P. striiformis* f. sp. *tritici*, we developed a segregating population by selfing isolate 12-368 on barberry (*Berberis vulgaris*) plants under controlled conditions. The dikaryotic sexual population segregated for avirulent/virulent phenotypes on nine *Yr* single-gene lines. The parental and progeny isolates were whole-genome sequenced at >30× coverage using Illumina HiSeq PE150 technology. A total of 2,637 high-quality markers were discovered by mapping the whole-genome sequencing (WGS) reads to the reference genome of strain 93-210 and used to construct a genetic map, consisting of 41 linkage groups, spanning 7,715.0 centimorgans (cM) and covering 68 Mb of the reference genome. The recombination rate was estimated to be 1.81 ± 2.32 cM/10 kb. Quantitative trait locus analysis mapped six *Avr* gene loci to the genetic map, including an *Avr* cluster harboring four *Avr* genes, *AvYr7*, *AvYr43*, *AvYr44*, and *AvYrExp2*. Aligning the genetic map to the reference genome identified *Avr* candidates and narrowed them to a small genomic region (<200 kb). The discovery of the *Avr* gene cluster is useful for understanding pathogen evolution, and the identification of candidate genes is an important step toward cloning *Avr* genes for studying molecular mechanisms of pathogen-host interactions.

IMPORTANCE Stripe rust is a destructive disease of wheat worldwide. Growing resistant cultivars is the most effective, easy-to-use, economical, and environmentally friendly strategy for the control of the disease. However, *P. striiformis* f. sp. *tritici* can produce new virulent races that may circumvent race-specific resistance. Therefore, understanding the genetic basis of the interactions between wheat genes for resistance and *P. striiformis* f. sp. *tritici* genes for avirulence is useful for improving cultivar resistance for more effective control of the disease. This study developed a high-quality map that facilitates genomic and genetic studies of important traits related to pathogen pathogenicity and adaptation to different environments and crop cultivars carrying different resistance genes. The information on avirulence/virulence genes identified in this study can be used for guiding breeding programs to select

Citation Xia C, Lei Y, Wang M, Chen W, Chen X. 2020. 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 5:e00128-20. https://doi.org/10.1128/mSphere.00128-20.

Editor Aaron P. Mitchell, University of Georgia

Address correspondence to Wanquan Chen, wqchen@ippcaas.cn, or Xianming Chen, xianming.chen@usda.gov.

Received 6 February 2020
Accepted 27 May 2020
Published 17 June 2020

combinations of genes for developing new cultivars with effective resistance to mitigate this devastating disease.

KEYWORDS avirulence, genetics, host-pathogen interaction, *Puccinia striiformis*, QTL mapping, wheat stripe rust, whole-genome sequencing

Rust diseases, caused by fungi in the order *Pucciniales*, are a large threat to food security and impact ecosystems (1). Economically important rusts include leaf rust in coffee (2), soybean rust (3), crown rust in oat (4), flax rust (5), poplar leaf rust (6), and wheat rusts (7). Numerous epidemics of these rusts have been recorded in diverse agriculture systems and in many countries. Extensive efforts have been made to incorporate resistance genes into cultivars to protect plants from attacks of rust pathogens. However, rapidly evolving rust fungi are highly capable of evading plant immunity systems, resulting in ineffective host resistance. To avoid resistance failure and elongate the effectiveness of resistance genes, understanding the molecular mechanisms underlying host-pathogen interactions is essential.

Plant-pathogen interactions were initially studied by Harold Flor and explained by his gene-for-gene concept (8). In this concept, host defense responses are activated by the recognition of a pathogen avirulence (*Avr*) gene by a cognate resistance (*R*) gene in the host. This concept has been supported by the fact that many *Avr* genes from bacteria, oomycetes, and fungi and *R* genes from different plant hosts have been cloned, and direct or indirect interactions between the products of some of the cloned *R* and *Avr* genes have been demonstrated (9). Such pioneering works have considerably increased our understanding of host-pathogen interactions. While *Avr* gene recognition is often referred to as effector-triggered immunity (ETI), pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) is generally thought of as a different type of resistance (10). PTI is activated by the recognition of conserved pathogen PAMPs such as chitin in the fungal cell wall. Pathogen effectors encoded by pathogenicity genes then suppress host PTI, enabling the pathogen to infect and cause disease. During coevolution, host plants gain *R* genes to detect evading pathogens. The recognition of Avr proteins by R proteins leads to ETI. As a coevolutionary arms race, a pathogen avoids the perception of a host R protein by mutating its *Avr* genes or developing new *Avr* genes to overcome or suppress host ETI.

Classical approaches for cloning *Avr* genes and effectors include reverse-genetics and mapping-based positional methods. The reverse-genetics approach has been used to successfully clone many *Avr* genes in different pathosystems, including *Cladosporium fulvum*-tomato, from which the first fungal *Avr* gene, *avr9*, was identified and cloned (11, 12). Even though many effectors from oomycetes and fungi have been cloned using this approach, the reverse-genetics approach for *Avr* identification is not suitable for wheat-*Puccinia* pathosystems because the techniques usually used in reverse genetics, such as transformation, effector delivery systems, and RNA interference, are still not mature for rust fungi (13).

An alternate approach for cloning *Avr* genes in plant pathogens is based on genetic mapping. Briefly, in the genetic mapping-based approach, an *Avr* locus is genetically mapped along with molecular markers; next, the genomic interval between two flanking markers is completely sequenced; and finally, the *Avr* gene is identified, followed by functional validation. The efficiency of this approach has been well demonstrated in cloning avirulence genes in the ascomycete fungus *Leptosphaeria maculans*, the causal agent of stem canker/black leg of oilseed rape and canola (14–18). Recently, this approach has been complemented by comparative genomics approaches, which has significantly accelerated the identification of *Avr* genes in *L. maculans* (19, 20). The mapping-based cloning approach has also been successful in basidiomycetes. In fact, this approach was successfully used to clone *UhAvr1* from the barley smut pathogen, *Ustilago hordei*, the first avirulence gene cloned from basidiomycete fungi (21, 22). Particularly in *Melampsora lini*, the flax rust fungus, which is also in the order *Pucciniales* containing cereal rust pathogens, several *Avr* genes that follow

the gene-for-gene relationship with flax resistance genes have been genetically mapped since the work of Flor and cloned using genome sequencing technology (23, 24). More recently, two *Avr* genes (*AvrSr35* and *AvrSr50*) in *Puccinia graminis* f. sp. *tritici*, the causal agent of stem rust of wheat and barley, were cloned. *AvrSr35* was identified through whole-genome sequencing (WGS) and comparison of chemically mutagenized mutants with a natural isolate (25), while *AvrSr50*, located in a loss-of-heterozygosity region, was identified by analyzing the genome variation and gene expression of spontaneous mutants (26).

Among the rust fungi, *Puccinia striiformis* Westend. f. sp. *tritici* Erikss. causes wheat stripe (yellow) rust and is recognized as one of the most serious plant pathogens threatening global food security (27–30). *P. striiformis* f. sp. *tritici* is a macrocyclic, heteroecious fungus having five spore stages in its complete life cycle. Its urediniospores (n + n) are produced on and can reinfect primary hosts (wheat and grasses), on which teliospores (2n) are usually produced in the late crop season. Teliospores germinate to produce basidiospores (n) after meiosis, and basidiospores infect alternate hosts (*Berberis* spp. and *Mahonia* spp.), on which pycniospores (n) are produced, and fertilize receptive hyphae (n) to produce aeciospores (n + n). Aeciospores infect the primary host and produce urediniospores (31–33). Although economically important, avirulence genes have not been molecularly identified and characterized in *P. striiformis* f. sp. *tritici*. Cloning of *Avr* genes has not been conducted in *P. striiformis* f. sp. *tritici*, and the potential of a genetic mapping-based approach has been impeded, mainly due to the unknown alternate hosts of *P. striiformis* f. sp. *tritici* until the recent discovery of its alternate hosts in the genera *Berberis* and *Mahonia* (31–33). Since then, sexual populations of *P. striiformis* f. sp. *tritici* have been developed, and genetic studies of the inheritance of virulence phenotypes have been conducted (34–36). However, no *Avr* genes could be precisely defined due to the limited number of codominant molecular markers and the fragmented reference genomes. To conquer these limitations, we developed a segregating population through self-fertilizing a *P. striiformis* f. sp. *tritici* isolate on barberry and developed a high-density genetic map consisting of a large number of genome-wide molecular markers by whole-genome sequencing of the progeny population using next-generation sequencing technology. We mapped six *Avr* genes, including four in a gene cluster. Comparison of the high-density map regions with a reference genome of the pathogen enabled us to identify *Avr* candidates in narrow genome regions. The results set the basis for cloning the *Avr* genes for understanding the molecular mechanisms underlying rapid virulence changes in the wheat stripe rust fungus.

RESULTS

Virulence phenotyping. The *P. striiformis* f. sp. *tritici* isolate 12-368 was selected to generate a self-fertilized sexual population based on its capability of producing abundant teliospores and high heterozygosity revealed by molecular markers, representing a different race group from those of our previously established sexual populations. The parental isolate and progeny isolates were kept as urediniospores that have two nuclei and can be asexually reproduced on susceptible wheat plants for a large quantity. The dikaryotic uredinial stage was genetically treated as diploid in the present study and used for virulence phenotyping and genomic DNA sequencing. Based on avirulence/virulence characterization of the set of 18 wheat *Yr* (yellow rust) single-gene differentials, isolate 12-368 was identified as belonging to race PSTv-4, with avirulence to the resistance genes *Yr5*, *Yr7*, *Yr8*, *Yr10*, *Yr15*, *Yr24*, *Yr32*, *Yr43*, *Yr44*, *YrTr1*, and *YrExp2* and virulence to the genes *Yr1*, *Yr6*, *Yr9*, *Yr17*, *Yr27*, *YrSP*, and *Yr76* (37).

The sexual reproduction of the parental isolate 12-368 through self-fertilization on barberry plants produced 117 progeny isolates (Fig. 1). Besides the above-mentioned 18 *Yr* single-gene lines for differentiating *P. striiformis* f. sp. *tritici* races, the parental and progeny isolates were also tested on 16 additional wheat lines, each with a single different resistance gene (see Materials and Methods). Virulence tests showed that the parental and progeny isolates were all avirulent to 12 *Yr* genes, including *Yr5*, *Yr10*, *Yr15*,

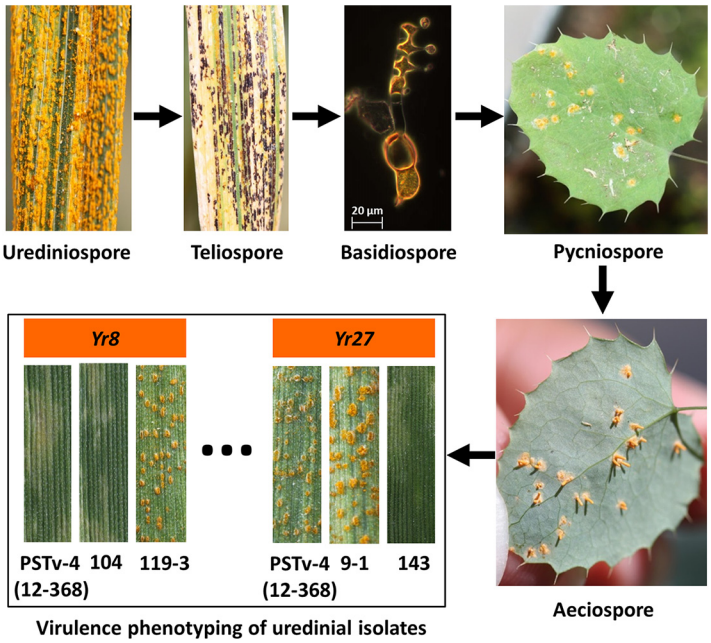

FIG 1 Experimental procedure for sexual population construction in this study.

Yr24, *Yr32*, *YrTr1*, *Yr26*, *YrCV*, *YrTr1*, *Yr45*, *Yr53*, and *Yr64*, indicating that the *Avr* genes corresponding to these *Yr* genes were homozygous in the parental isolate. Similarly, the parental and progeny isolates were all virulent to 13 *Yr* genes, including *Yr1*, *Yr2*, *Yr6*, *Yr9*, *Yr21*, *Yr25*, *Yr28*, *Yr29*, *Yr31*, *Yr76*, *YrA* (i.e., *Yr73* plus *Yr74*), *Yr74* (in the Avocet Susceptible [AvS] cultivar), and *YrSP*, suggesting that the virulence loci corresponding to these *Yr* genes were also homozygous in the parental isolate. Therefore, these avirulence or virulence loci could not be mapped in this study. Detailed infection types (ITs) of parental and progeny isolates are provided in Data Set S1 in the supplemental material.

In contrast, the avirulence/virulence phenotypes of the parental isolate to nine *Yr* genes (*Yr7*, *Yr8*, *Yr17*, *Yr27*, *Yr35*, *Yr41*, *Yr43*, *Yr44*, and *YrExp2*) were segregating in the progeny population (Table 1). Thus, the *Avr* genes corresponding to these resistance genes could be mapped. Since the dikaryotic (two unfused nuclei in a cell) urediniospores are heterozygous at these loci, the parental isolate was considered F_1, and the progeny isolates produced through self-fertilization of the parental isolate were considered F_2. Therefore, the avirulence/virulence phenotypes of progeny isolates should

TABLE 1 Segregation of avirulence/virulence in the progeny isolates derived from selfing parental isolate 12-368 of *Puccinia striiformis* f. sp. *tritici* on wheat *Yr* single-gene lines

Wheat *Yr* gene line	IT[a] of 12-368	No. of progeny isolates		Expected ratio (A/V)	*P*[b]	Avirulence gene(s)
		A	V			
Yr7	3 (A)	86	31	3:1	0.71	*AvYr7*
Yr43	4 (A)	91	36	3:1	0.20	*AvYr43*
Yr44	3 (A)	91	36	3:1	0.20	*AvYr44*
YrExp2	3 (A)	82	35	3:1	0.22	*AvYrExp2*
Yr8	2 (A)	105	12	15:1	0.07	*AvYr8-1*, *AvYr8-2*
Yr27	7 (V)	31	86	1:3	0.71	*avYr27*
Yr17	8 (V)	21	95	1:3	0.09	*avYr17*
Yr41	8 (V)	16	101	3:13	0.16	*AvYr41*, *AvYr41-Inh*
Yr35	7 (V)	48	69	7:9	0.55	*avYr35-1*, *avYr35-2*

[a]IT, infection type based on a scale from 0 to 9, with 0 to 6 being avirulent (A) and 7 to 9 being virulent (V).
[b]*P*, probability of goodness of fit by a χ^2 test.

follow the segregation patterns in an F_2 population. On wheat lines carrying *Yr7*, *Yr8*, *Yr17*, *Yr41*, *Yr43*, *Yr44*, and *YrExp2*, the avirulence/virulence phenotypes of progeny isolates fit the models that would be expected if the avirulence phenotype were dominant, whereas on the wheat lines with *Yr27* or *Yr35*, to which the parental isolate was virulent, phenotypes of the progeny isolates were segregating, suggesting that the virulence phenotypes were dominant (Table 1).

The segregations of the parental avirulent phenotypes on wheat lines with *Yr7*, *Yr43*, *Yr44*, and *YrExp2* fit the 3:1 avirulent/virulent (A/V) ratio, suggesting that each of the avirulent phenotypes of the parental isolate was controlled by a dominant gene. Therefore, the avirulence genes were designated *AvYr7*, *AvYr43*, *AvYr44*, and *AvYrExp2*, respectively. The segregation of avirulence on the wheat line with *Yr8* fit the 15:1 A/V ratio, indicating two dominant avirulence genes, designated *AvYr8-1* and *AvYr8-2*. The virulence phenotypes on wheat lines possessing *Yr17* and *Yr27* segregated at the 1:3 A/V ratio, indicating a recessive avirulence or a dominant virulence gene corresponding to each of the resistance genes. These *P. striiformis* f. sp. *tritici* genes were designated *avYr17* and *avYr27*, respectively. The segregation of virulence to *Yr41* best fit the 3:13 A/V ratio, indicating an epistatic interaction of a dominant inhibitor (*AvYr41-Inh*) over a dominant avirulence gene (*AvYr41*). On the wheat line with *Yr35*, the observed 7:9 A/V ratio indicated two independent recessive genes for avirulence, designated *avYr35-1* and *avYr35-2*.

Genotyping by whole-genome sequencing. To identify molecular markers for genetic mapping, Illumina HiSeq 150-bp paired-end (PE) technology was used to sequence the whole genomes of all 117 progeny isolates as well as the parental isolate. Twenty-six million pairs of reads (7.89 Gb in total) and 1.5 billion pairs of reads (471 Gb in total) were generated from the parental and progeny isolates, respectively. The numbers of filtered reads, percentages of mapped reads, and mapping coverages and qualities for the parental and all progeny isolates are summarized in Table S1. On average, 4.02-Gb sequences were generated for each progeny isolate. The high-quality reads of progeny isolates were mapped to the reference genome of isolate *P. striiformis* f. sp. *tritici* 93-210 (38). On average, 95.72% of reads were mapped to the reference genome. The mapping coverage of the progeny isolates ranged from 24.16× to 50.99×, with a mean coverage of 35.45×. The deep sequencing and high-quality reads enabled us to identify genome-wide variations for genetic mapping.

In total, 2,487 heterozygous single nucleotide polymorphisms (SNPs) and 150 indels in the parental isolate genome were obtained based on their segregation at the 1:2:1 (AA:AB:BB) ratio ($P \geq 0.05$ by a chi-squared test) in the progeny population (Fig. 2A). Contamination analysis suggested that 23 of the progeny isolates had abnormal numbers of crossover events, and these were therefore excluded from subsequent analyses. A total of 2,637 codominant markers were selected from 251 (out of 492) contigs, covering 71.35 Mb (out of 84.62 Mb) of the reference genome (Fig. 2B). The mean distance between two markers in the reference genome was 42.93 kb. Detailed genotypes of parental and progeny isolates are listed in Data Set S2.

Genetic map of *P. striiformis* f. sp. *tritici*. After correcting allele switches and filtering potentially contaminated isolates, a genetic map was generated using the minimum spanning tree algorithm at a *P* value of 1E−10, which contained 2,631 markers in 41 linkage groups (LGs) (LG-1 to LG-41) (Table 2; Data Set S3); the remaining 6 markers could not be linked in the genetic map. The genetic map spanned a total of 7,715.0 centimorgans (cM), with individual LGs of up to 1,011.0 cM (LG-1, with 312 markers). The average genetic distance between markers was 2.94 cM throughout the genome. Comparison between LGs and the reference genome showed that multiple genome contigs could be tagged to one LG. For example, LG-1 included 39 contigs, covering over 10.86 Mb. With this information, we were able to estimate the recombination rate for LGs with at least 10 markers, and the average recombination rate across these LGs was 1.81 ± 2.32 cM/10 kb (Table 3). This rate was higher than those of the flax

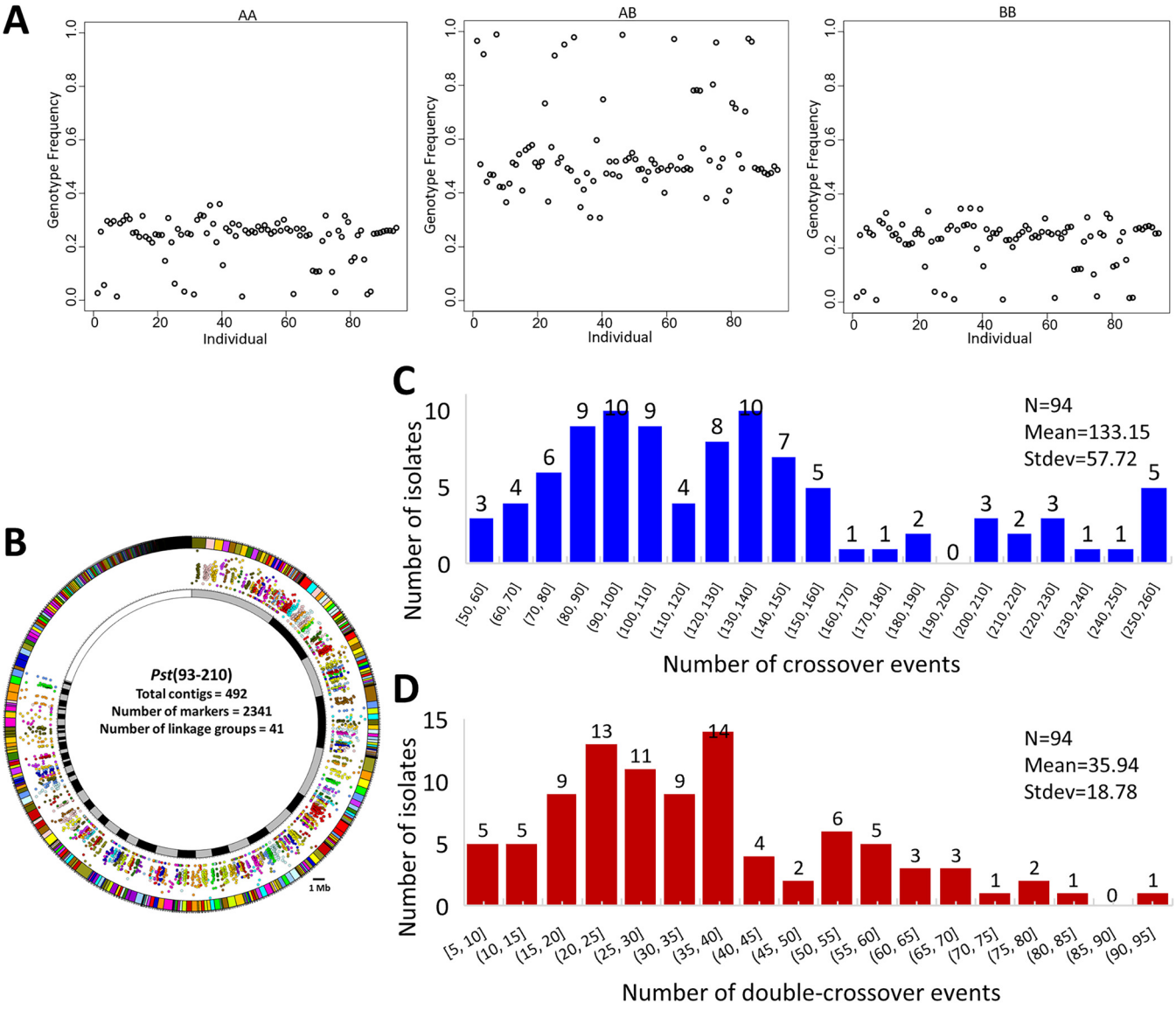

FIG 2 Genotyping by whole-genome sequencing and genetic map construction. (A) Frequencies of homozygous AA, heterozygous (AB), and homozygous BB marker loci. (B) Circos plot of the distribution of markers along the genome and the linkage groups. The outer layer shows the contigs of the reference genome. Each colored bar represents one contig. The middle layer shows the distribution of markers in the reference genome. Each solid dot represents one marker. The inner layer shows the linkage groups of the genetic map. Note that the reference genome is reorganized corresponding to genetic groups, and the empty white bar in the inner layer represents the contigs that cannot be mapped to the linkage groups. *Pst*, *P. striiformis* f. sp. *tritici*. (C) Distribution of different numbers of single-crossover events. (D) Distribution of different numbers of double-crossover events.

rust fungus *M. lini* (1.18 cM/10 kb) and *Zymoseptoria tritici* (1.25 cM/10 kb) and much higher than that of *Fusarium graminearum* (0.3 to 05 cM/10 kb) (Table 3).

To investigate potential genome features that might contribute to the relatively large *P. striiformis* f. sp. *tritici* genetic map and high recombination rate, we estimated the numbers of single- and double-crossover events in each isolate, with averages of 133.15 and 35.94, respectively (Fig. 2C and D). We also calculated the genome coverage by CpG islands in *P. striiformis* f. sp. *tritici* and compared this value with those of a few other plant-pathogenic fungi. The CpG islands in the *P. striiformis* f. sp. *tritici* reference isolate covered 7.26% of the genome (Table 3). The CpG island coverage was higher than those of *M. lini* (5.09%), *Fusarium graminearum* (2.87%), and *Zymoseptoria tritici* (1.64%) but slightly lower than that of the pine fusiform rust fungus *Cronartium quercuum* f. sp. *fusiforme* (7.37%) (Table 3). The average recombination rate of the *P. striiformis* f. sp. *tritici* CpG islands was estimated to be 1.81 cM/10 kb, the same as the

TABLE 2 General features of the constructed linkage map[a]

Chr	No. of markers	Length (cM)	Avg genetic distance between markers (cM) ± SD	No. of contigs	Contig length (bp)	Mean recombination rate (cM/10 kb) ± SD
LG-1	312	1,011.0	3.25 ± 3.04	39	10,869,479	0.66 ± 3.07
LG-2	288	762.7	2.65 ± 1.99	27	6,567,233	1.94 ± 2.21
LG-3	232	696.5	3.01 ± 2.23	21	4,370,414	1.66 ± 2.43
LG-4	216	655.5	3.04 ± 2.54	24	6,208,200	1.72 ± 2.37
LG-5	183	436.7	2.39 ± 2.36	18	5,300,867	1.35 ± 1.61
LG-6	190	413.4	2.18 ± 1.66	13	1,825,696	1.48 ± 2.18
LG-7	124	387.0	3.14 ± 2.75	13	2,688,337	1.80 ± 2.20
LG-8	132	343.6	2.62 ± 1.63	9	2,509,186	1.84 ± 2.71
LG-9	101	293.7	2.93 ± 2.40	11	2,209,512	1.99 ± 2.07
LG-10	93	292.1	3.17 ± 2.29	5	943,798	1.95 ± 2.01
LG-11	67	268.9	4.07 ± 3.58	7	1,838,758	2.11 ± 2.01
LG-12	76	248.3	3.31 ± 2.55	10	2,280,812	1.61 ± 1.78
LG-13	72	232.5	3.27 ± 3.00	11	2,224,764	1.94 ± 2.02
LG-14	51	179.0	3.58 ± 3.65	8	1,834,719	2.11 ± 2.25
LG-15	67	167.7	2.54 ± 1.55	5	1,431,574	1.77 ± 2.15
LG-16	48	153.8	3.27 ± 1.74	7	1,390,043	2.17 ± 2.78
LG-17	49	152.1	3.16 ± 2.29	3	549,149	2.01 ± 2.11
LG-18	58	147.4	2.58 ± 1.46	7	1,158,936	1.42 ± 1.70
LG-19	30	135.0	4.65 ± 3.47	9	833,085	3.19 ± 3.37
LG-20	28	81.9	3.03 ± 1.28	2	425,950	1.05 ± 1.52
LG-21	27	81.4	3.13 ± 2.35	4	923,042	1.92 ± 1.86
LG-22	13	62.0	5.14 ± 3.48	3	768,977	1.46 ± 1.06
LG-23	15	54.1	3.86 ± 3.43	1	602,119	1.60 ± 1.14
LG-24	19	50.9	2.82 ± 1.51	1	512,194	1.84 ± 2.48
LG-25	19	48.4	2.68 ± 1.52	5	866,427	1.99 ± 1.49
LG-26	26	43.9	1.75 ± 1.79	3	602,608	0.76 ± 1.45
LG-27	12	43.3	3.94 ± 2.64	1	287,392	1.94 ± 0.93
LG-28	9	38.3	4.78 ± 3.31	2	838,633	NC
LG-29	14	32.2	1.16 ± 1.46	2	476,402	3.55 ± 2.49
LG-30	10	31.0	3.44 ± 3.26	2	947,870	2.58 ± 2.44
LG-31	7	26.9	4.48 ± 1.88	1	295,440	NC
LG-32	5	25.4	6.36 ± 2.67	1	336,304	NC
LG-33	3	20.9	10.46 ± 3.39	1	536,190	NC
LG-34	7	19.4	3.23 ± 0.68	1	295,889	NC
LG-35	6	18.7	3.73 ± 1.19	1	261,047	NC
LG-36	6	17.3	3.45 ± 1.55	2	397,004	NC
LG-37	5	14.4	3.59 ± 0.74	1	815,364	NC
LG-38	3	13.5	6.74 ± 0.32	2	630,032	NC
LG-39	4	9.5	3.16 ± 1.46	2	554,726	NC
LG-40	2	4.1		1	181,331	NC
LG-41	2	0.5		1	230,054	NC
Overall	2,631	7,715.0	2.94 ± 2.39	251	68,819,557	1.81 ± 2.32

[a]Chr, chromosome; NC, not calculated because the linkage has <10 markers.

average recombination rate of the overall genome mentioned above. Due to the lack of data for the other fungi, we were unable to compare the CpG island recombination rates of *P. striiformis* f. sp. *tritici* with those of these fungi.

QTL mapping for avirulence loci. The genetic map and infection type (IT) data of the segregating avirulence/virulence phenotypes of the progeny population were used for quantitative trait locus (QTL) mapping. Six avirulence genes (*AvYr7*, *AvYr43*, *AvYr44*, *AvYrExp2*, *AvYr8-1*, and *avYr27*) were mapped to three LGs (Table 4). *AvYr8-1* was mapped to LG-19, at the 6.82- to 31.09-cM region flanked by markers C085_283131 and C182_5504. *avYr27* was mapped to LG-4, between the 446.65- and 528.51-cM positions flanked by markers C162_20837 and C086_188415, respectively (Data Set S3). The QTL confidence interval of *AvYr8-1* covered three contigs, contig 1.085 (from kb 200 to kb 283), contig 1.137 (from kb 12 to kb 29), and contig 1.182 (from kb 5 to kb 14) in the reference genome of isolate 93-210. The *avYr27* interval covered two contigs, contig 1.086 (from kb 10 to kb 188) and contig 1.162 (from kb 1 to kb 164). Four avirulence genes, *AvYr44*, *AvYr7*, *AvYr43*, and *AvYrExp2*, were mapped to the same LG region

TABLE 3 Genome-wide CpG islands in selected plant pathogens

Parameter	Value for pathogen				
	C. quercuum f. sp. *fusiforme*[a]	*F. graminearum*[b]	*M. lini*[c]	*P. striiformis* f. sp. *tritici*[d]	*Z. tritici*[e]
Genome length (bp)	76,567,842	37,946,458	189,516,653	84,531,325	39,686,251
Mean recombination rate (cM/10 kb) ± SD (reference)	Unknown (51)[g]	0.3–0.5 (52)[g]	1.18 (24)[g]	1.81 ± 2.32	1.25 (53)[g]
Total no. of CpGs	1,310,360	1,846,827	4,449,449	2,003,593	2,796,481
No. of CpG dinucleotides in CpG islands (%)	483,899 (36.9)	128,731 (6.97)	765,979 (17.2)	587,056 (29.30)	105,599 (3.77)
No. of predicted CpG islands	29,754	6,260	40,815	29,671	5,454
Island coverage (%)[f]	7.37	2.87	5.09	7.26	1.64
Island length (bp)					
Avg ± SD	189.68 ± 130.65	174.07 ± 108.73	236.65 ± 164.82	207.02 ± 135.58	120.05 ± 77.74
Min	6	8	6	6	8
Max	1,490	918	3,223	1,733	697
Avg island GC% ± SD	54.62 ± 7.25	60.58 ± 6.92	54.85 ± 8.98	57.51 ± 2.08	66.64 ± 8.07
Avg CpG O/E ratio ± SD	1.37 ± 0.35	1.53 ± 0.31	1.31 ± 0.36	1.37 ± 0.08	1.73 ± 0.30

[a]The reference genome was from G11 (https://genome.jgi.doe.gov/portal/Croqu1/download/Croqu1_AssemblyScaffolds.fasta.gz).
[b]The reference genome was from isolate RRES (GenBank accession no. HG970335).
[c]The reference genome was from isolate CH5 (https://genome.jgi.doe.gov/portal/Melli1/download/Melli1_AssemblyScaffolds.fasta.gz).
[d]The reference genome was from isolate PST93-210.
[e]The reference genome was from isolate IPO323 (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/219/625/GCF_000219625.1_MYCGR_v2.0/GCF_000219625.1_MYCGR_v2.0_genomic.fna.gz).
[f]Percentage of the genome covered by CpG islands.
[g]Reference from which the previously estimated recombination rates were retrieved.

between the 29.06- and 57.49-cM positions flanked by markers C022_56722 and C022_180222 in LG-22, respectively (Fig. 3A; Data Set S3).

Genomic location of the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster. All markers from similar QTL regions for *AvYr44*, *AvYrYr7*, *AvYr43*, and *AvYrExp2* were located in a single contig, namely, contig 1.002, in the reference genome of isolate 93-210. The leftmost marker was at bp 6812, and the rightmost marker was at bp 180222, indicating that the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster was most likely within the region of the

TABLE 4 Quantitative trait loci for *Avr* genes identified in the selfing population of isolate 12-368 of *Puccinia striiformis* f. sp. *tritici*

QTL	Genetic map							Physical map[b]	
	Linkage group	Flanking markers	Peak position (cM)	Interval (cM)	LOD	*P* value	PVE[a]	Contig	Interval (kb)
AvYr8-1	LG-19	C085_283131, C182_5504	19.00	6.82–31.09	6.73	<0.0001	30.85	1.085 1.137 1.182	200–283 12–29 5–14
avYr27	LG-4	C162_20837, C086_188415	484	446.65–528.51	6.16	<0.0001	28.65	1.086 1.162	10–188 1–164
AvYr44	LG-22	C022_56722, C022_180222	47.00	29.06–57.49	12.07	<0.0001	48.40	1.022	0–200
AvYr7	LG-22	C022_56722, C022_180222	46.00	29.06–57.49	10.3	<0.0001	43.14	1.022	0–200
AvYr43	LG-22	C022_56722, C022_180222	43.00	29.06–57.49	10.91	<0.0001	44.98	1.022	0–200
AvYrExp2	LG-22	C022_56722, C022_180222	47.00	29.06–57.49	9.36	<0.0001	40.13	1.022	0–200

[a]PVE, percentage of variance explained by the QTL, calculated as $1 - 10^{-\frac{2}{n}\mathrm{LOD}}$, where n is the total number of individuals.
[b]Based on the reference genome of isolate 93-210 (38).

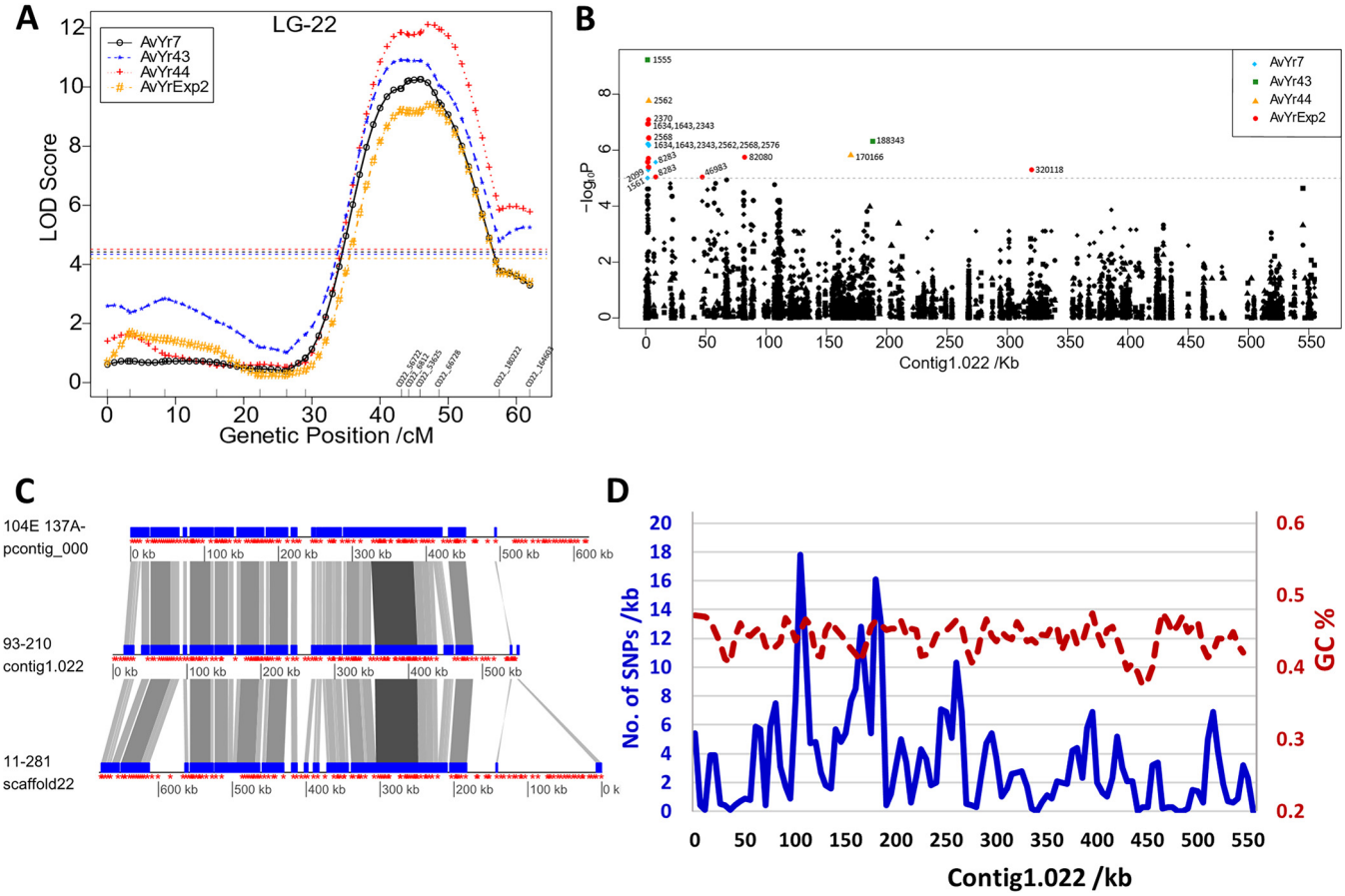

FIG 3 The *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster. (A) QTL mapping identified a region carrying the *AvYr44*, *AvYr7*, *AvYr43*, and *AvYrExp2* genes. (B) Manhattan plot of the genome-wide association study (GWAS) performed on contig 1.022. (C) Syntenic analysis of contig 1.022 from the reference genome of isolate 93-210 with homologous contigs from the genomes of isolates 104E 137A− and 11-281. Note that the first 10 kb of contig 1.022, which harbors most of the significant GWAS signals, were absent in 104E 137A− and 11-281. Blue blocks are syntenic regions. Red stars are genes. The darker the linking lines between two genomes, the higher degree of similarity between the genomes. (D) SNP density and GC content of contig 1.022.

first 200 kb of contig 1.002 in the reference genome. Our previous annotation of this region of the reference genome (39) identified 47 protein-coding genes, 4 of which have signal peptides in the N terminus (Table 5).

To validate the QTL mapping results and to identify putative casual variations, we performed a genome-wide association study (GWAS) for each of the *AvYr44*, *AvYr7*, *AvYr43*, and *AvYrExp2* genes. Instead of using only markers in nontransposable elements (non-TEs) and without significant distortion for the 1:2:1 ratio in the QTL analysis, all markers and all 117 progeny isolates were used for the GWAS. In this way, 609 markers from contig 1.022 were obtained and used in the GWAS. Seventeen out of the 18 significantly associated SNPs were within the first 200 kb on contig 1.022. Surprisingly, a major GWAS signal peak was within the first 5-kb region (Fig. 3B). This peak matched the gene *PSTG_03388*, and several of the SNPs within the gene were nonsynonymous (Table 6). Thus, both QTL analyses and GWASs located the *AvYr7-AvYr43-AvYrExp2* locus to the first 200-kb region of contig 1.022.

We calculated the SNP density and GC content in contig 1.022. The region harboring the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster had 49 SNPs within the first 5 kb and only 4 between kb 5 and kb 20. The GC content was 47% within the first 20 kb, relatively in balance with the AT content along the contig (Fig. 3C). We further investigated whether AT-rich genomic regions exist in the *P. striiformis* f. sp. *tritici* genome and whether contig 1.022 is located in an AT-rich region. The unimodal distributions of GC content indicated that there were no distinct AT-rich segments in either the whole genome or

TABLE 5 Candidate genes of the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster in the confidence interval defined by QTL analysis

93-210 gene name	Functional annotation	Secreted	Mature length (amino acids)	% cysteine
PSTG_03388	Hypothetical protein	No	291	1.03
PSTG_03389	Enoyl-(acyl carrier protein) reductase	No	351	1.14
PSTG_03390	ABC transporter	No	1,053	0.66
PSTG_03391	Hypothetical protein	No	102	1.96
PSTG_03392	Hypothetical protein	No	814	1.11
PSTG_03393	Hypothetical protein	No	199	0.50
PSTG_03394	Hypothetical protein	Yes	228	0.88
PSTG_03395	Noncatalytic module family expansin	No	351	0.57
PSTG_03396	Hypothetical protein	No	545	0.37
PSTG_03397	Hypothetical protein	No	458	1.09
PSTG_03398	Hypothetical protein	No	489	0.82
PSTG_03399	Hypothetical protein	Yes	143	1.40
PSTG_03400	Threonine dehydratase I	No	486	1.03
PSTG_03401	Hypothetical protein	No	394	0.25
PSTG_03402	Hypothetical protein	No	400	0.75
PSTG_03403	Hypothetical protein	No	414	0.97
PSTG_03404	Hypothetical protein	No	154	0.65
PSTG_03405	Ca^{2+}/calmodulin-dependent protein kinase	No	452	0.66
PSTG_03406	zap1 metalloregulator involved in zinc-responsive transcriptional regulation	No	628	1.27
PSTG_03407	Ubiquitin-specific protease 7	No	1,116	0.45
PSTG_03408	Hypothetical protein	Yes	130	2.31
PSTG_03409	Hypothetical protein	No	556	0.72
PSTG_03410	RING-H2 finger ATL54-like	No	243	0.82
PSTG_03411	P-loop-containing nucleoside triphosphate hydrolase	No	189	1.59
PSTG_03412	Per1-like; involved in manganese homeostasis	No	458	0.87
PSTG_03413	Inositol monophosphatase	No	315	0.95
PSTG_03414	Subunit of cytochrome *bd* ubiquinol oxidase	No	128	0.00
PSTG_03415	Hypothetical protein	Yes	338	0.30
PSTG_03416	Hypothetical protein	No	148	1.35
PSTG_03417	Hypothetical protein	No	173	0.58
PSTG_03418	Hypothetical protein	No	135	0.00
PSTG_03419	Hypothetical protein	No	189	0.53
PSTG_03420	3-Deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthetase I family	No	435	1.38
PSTG_03421	Hypothetical protein	No	678	1.18
PSTG_03422	tRNA			
PSTG_03423	Hypothetical protein	No	696	1.01
PSTG_03424	Hypothetical protein	No	420	1.43
PSTG_03425	Hypothetical protein	No	380	1.05
PSTG_03426	Hypothetical protein	No	142	1.41
PSTG_03427	Hypothetical protein	No	193	0.52
PSTG_03428	Hypothetical protein	No	301	0.66
PSTG_03429	Hypothetical protein	No	409	1.22
PSTG_03430	Hypothetical protein	No	23	0.00
PSTG_03431	Hypothetical protein	No	269	0.74
PSTG_03432	Hypothetical protein	No	90	0.00
PSTG_03433	Hypothetical protein	No	241	0.41
PSTG_03434	RNA polymerase I-specific transcription initiation factor rrn11	No	270	1.48

contig 1.022 (Fig. S1). Therefore, we conclude that the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster is not located in an AT-rich region. Next, we attempted to investigate the genomic environment of the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster. Considering that this cluster resides in one of the contig terminal regions, we tried to extend contig 1.022 by aligning it to three other well-assembled genomes, those of isolates 104E 137A− (40), 11-281 (41), and 93TX-2, in addition to the reference genome (93-210) (Fig. S2). Surprisingly, the homologous contigs of these genomes terminated around the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster, except for the genomes of the parental and reference isolates (Fig. 3D). In fact, the first 10-kb region of contig 1.022 in the reference genome of isolate 93-210 was mostly or partially absent in the 104E 137A−, 93TX-2, and 11-281 genomes (Fig. S2). Besides the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster, the remaining regions were highly conserved among these three genomes. Most SNPs associated with the *Avr* cluster were found within the first 2,600 bp of the contig in the parental isolate

TABLE 6 SNPs within the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster significantly associated with avirulent (*Avr*) and virulent (*avr*) phenotypes

Avr gene	SNP	*P*_FDR_adj	Genotype(s)		Annotation
			Avr	*avr*	
AvYr44	C022_2562	3.30E−05	GA	AA	Intergenic
	C022_170166	0.00144	TC	TT	Intergenic
AvYr7	C022_1634	0.00018	GA	GG	PSTG_03388, nonsynonymous, Pro/Leu
	C022_1643	0.00018	TC	CC	PSTG_03388, nonsynonymous, Glu/Gly
	C022_2343	0.00018	TC	CC	Intergenic
	C022_2562	0.00018	GA	AA	Intergenic
	C022_2568	0.00018	GT	TT	Intergenic
	C022_2576	0.00018	AC	AA	Intergenic
	C022_2584	0.00018	GA	AA	Intergenic
	C022_8283	0.00064	GC, CC	GG	PSTG_03390, nonsynonymous, Asp/His
	C022_2099	0.00109	GT	TT	PSTG_03388, nonsynonymous, Pro/Gln
	C022_1555	0.00172	TA	AA	PSTG_03388, synonymous
	C022_1561	0.00172	GA	AA	PSTG_03388, synonymous
AvYr43	C022_1555	1.12E−10	TA	AA	PSTG_03388, synonymous
	C022_188343	0.00045	AT	AA	Intergenic
	C079_150209	0.00226	GA	AA	PSTG_03421, intron
AvYrExp2	C022_1634	5.47E−05	GA	GG	PSTG_03388, nonsynonymous, Pro/Leu
	C022_1643	5.47E−05	TC	CC	PSTG_03388, nonsynonymous, Glu/Gly
	C022_2343	5.47E−05	TC	CC	Intergenic
	C022_2370	5.54E−05	CT	TT	Intergenic
	C022_2562	8.70E−05	GA	AA	Intergenic
	C022_2568	8.70E−05	GT	TT	Intergenic
	C022_2576	8.70E−05	AC	CC	Intergenic
	C022_2584	8.70E−05	GA	AA	Intergenic
	C022_2282	0.00034	AG	GG	Intergenic
	C022_2294	0.00034	TC	CC	Intergenic
	C022_82080	0.00034	AG, GG	AA	PSTG_03406, synonymous
	C022_1531	0.00038	GT	TT	PSTG_03388, synonymous
	C022_1890	0.00038	CG	GG	PSTG_03388, nonsynonymous, Glu/Gln
	C022_2099	0.00051	GT	TT	PSTG_03388, nonsynonymous, Pro/Gln
	C022_2676	0.00051	AC, CC	AA	Intergenic
	C022_320118	0.00059	TT, T–		Intergenic
	C022_8283	0.00096	GC, CC	GG	PSTG_03390, nonsynonymous, Asp/His
	C022_46983	0.00096	TA, AA	TT	PSTG_03396, synonymous

and within the first 2,700 bp of the contig in the reference genome. The first seven SNPs were within the first exon identified in the contig. In summary, the results suggested that the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster resides in a genetically complex region attached to a highly conserved genomic region.

DISCUSSION

In the present study, we used Illumina sequencing technology to construct a high-density genetic map for mapping *Avr* genes in the wheat stripe rust fungus. We generated a *P. striiformis* f. sp. *tritici* sexual population for genetic mapping by self-fertilizing *P. striiformis* f. sp. *tritici* isolate 12-368 (race PSTv-4). Whole-genome deep sequencing of progeny isolates generated 2,637 high-quality codominant molecular markers, which enabled us to construct a high-density genetic map for *P. striiformis* f. sp. *tritici* comprising 41 LGs. QTL analysis mapped six *Avr* genes in three LG regions. Moreover, an avirulence gene cluster carrying four *Avr* genes was identified and located at a single contig in the *P. striiformis* f. sp. *tritici* reference genome. Aligning the genetic map to the reference genome enabled us to further locate the *Avr* candidates at a small genomic region (<200 kb). This study provides the resources for functional cloning of *Avr* genes and a better understanding of the genomic basis of the rapid evolution of virulence in the wheat-*P. striiformis* f. sp. *tritici* pathosystem.

Isolate-dependent inheritance of avirulence/virulence in *P. striiformis* f. sp. *tritici*. The segregation patterns of phenotypes in the progeny population suggest a

complex inheritance of avirulence/virulence in *P. striiformis* f. sp. *tritici*, which is consistent with previous observations (34–36). First, different phenotypes of avirulence/virulence to individual *Yr* genes could be controlled by one or two genes in a single isolate. In isolate 12-368 used in the present study, the avirulence phenotypes to *Yr7*, *Yr43*, *Yr44*, and *YrExp2* were each controlled by a single dominant gene, whereas the avirulence phenotype to *Yr8* was controlled by two dominant genes. Different numbers of genes controlling avirulence/virulence phenotypes in a single isolate have also been observed in other *P. striiformis* f. sp. *tritici* selfing populations (34–36) as well as in other fungi (42). One possible explanation is that there may be an unidentified resistance gene(s) in the wheat *Yr* gene lines used in the genetic studies of *P. striiformis* f. sp. *tritici*, even though many of these wheat lines are nearly isogenic in the Avocet Susceptible (AvS) background. AvS was reported to have *Yr74*, which is complementary to *Yr73*, to provide resistance to some *P. striiformis* f. sp. *tritici* races in Australia (43). It is likely that some of the *Yr* near-isogenic lines also have *Yr74*. Yuan et al. (36) reported that the virulence phenotype in the parental isolate (08-220, race PSTv-11) was heterozygous and segregated in its selfed progeny population (36). In contrast, the parental isolate and the progeny isolates in the present study were all virulent on AvS (*Yr74*). Therefore, the presence of *Yr74* in the AvS background of many *Yr* single-gene lines did not influence the segregation ratios of the wheat lines with an AvS background. Thus, it is possible for two avirulence genes to interact with a single *Yr* gene. Second, the particular avirulence/virulence phenotype to a single *Yr* gene could be controlled by one gene in one isolate but by two or more genes in another isolate. Third, the complex inheritance in *P. striiformis* f. sp. *tritici* is also suggested by the fact that the avirulence/virulence phenotype to a single *Yr* gene could be dominant in one isolate but recessive in another. For example, the avirulence phenotype on *Yr7* was determined by a single dominant avirulence gene in isolate 12-368 in the present study and in isolate GS-2013 (34) but was determined by a single dominant virulence gene in isolates CY32 (35) and PSTv-11 (36). A possible explanation for this is the underestimation of the interaction between a dominant avirulence gene and a dominant inhibitor gene by misclassification of the segregation of the 3:13 ratio as 1:3, at least in some cases. Taken together, the inheritance of avirulence/virulence and the interactions between avirulence/virulence-controlling genes in *P. striiformis* f. sp. *tritici* are isolate dependent.

The complex interactions between fungal *Avr* genes and their corresponding *R* genes have been described in many plant pathosystems (9). The isolate-dependent nature of avirulence/virulence inheritance observed in *P. striiformis* f. sp. *tritici* has also been reported in other rust fungi (44–46). Until now, no universal genetic models have been available to explain such complex inheritance. Therefore, further identification and comparison of genes and genomic regions between different isolates showing different avirulence/virulence inheritance are needed to test the association of complex inheritance and the plasticity of the genomic environment.

Avirulence gene cluster in *P. striiformis* f. sp. *tritici*. The QTL analysis and the GWAS mapped four *Avr* genes in a small genomic region (Table 4 and Fig. 3A and B), indicating a cluster of *Avr* genes in *P. striiformis* f. sp. *tritici* isolate 12-368. The existence of *Avr/Vir* clusters has also been revealed from different *P. striiformis* f. sp. *tritici* mapping populations in previous studies (34–36). However, due to the limited number of molecular markers (34, 35) or the lack of codominant markers (36), either previous studies were not able to identify cosegregating markers or the flanking markers were too far away from the cluster to precisely define the genomic intervals of the cluster. In contrast, QTL mapping with the highly improved reference genome enabled us to map an *Avr* cluster, *AvYr44-AvYr7-AvYr43-AvYrExp2*, in LG-22 to a single contig of the reference genome. *Avr* and effector genes usually reside in plastic genomic regions, and gene clusters are often located in such regions (47). Such *Avr* gene clusters are not uncommon in cereal rust fungi. For example, the *Avr* genes in the flax rust fungus *M. lini* were genetically mapped to four small regions, and the genes within each region were tightly linked and inherited as a unit (5). Several *Avr* genes have been cloned from

these clusters. Interestingly, single *Avr* genes (e.g., *AvrL567* and *AvrM14*) in *M. lini* controlling avirulence to several *R* genes have been reported (23, 24). Such a single *Avr* gene (allele) recognized by multiple *R* genes has also been reported in the pathosystem of *Leptosphaeria maculans*-oilseed rape (48). Syntenic analysis revealed that the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster resides in one of the contig termini, which was difficult to assemble in different isolates (Fig. 3D; Fig. S1). It will be interesting to determine whether this *Avr* cluster is located in a subtelomere-like region, similar to *AvrPita* from *Pyricularia oryzae* (49) and *AvrStb6* from *Z. tritici* (50), or a region adjacent to a repetitive region of the genome. Further work using linked- or long-read sequencing, e.g., bacterial artificial chromosome sequencing, PacBio technology, or nanopore sequencing, is needed to dissect the cluster of *AvYr44-AvYr7-AvYr43-AvYrExp2* loci for cloning the *Avr* genes and for providing a genomic basis for the rapid avirulence changes in *P. striiformis* f. sp. *tritici*. The genes identified in the first 20 kb of the contig can be studied for expression to determine if they are functionally associated with avirulence.

Large genetic map with a high recombination rate of *P. striiformis* f. sp. *tritici*. Generally, our genetic map was consistent with the physical map (Data Set S3). For example, the markers of each large contig (>200 kb in length) were always located nearby in the genetic map. However, we noticed two types of inconsistencies. First, some of the markers from different contigs were interwoven in the genetic map, especially the markers from short contigs. We speculate that these might be due to either the highly repetitive nature of the *P. striiformis* f. sp. *tritici* genome or the high heterozygosity between two marker loci. Second, even though markers from the same contig were located nearby in the genetic map, their locations were not linearly correlated. This inconsistency might be caused by structural variations (e.g., genome rearrangement) between the parental and reference isolates or by possible genotyping errors from intrinsic sequencing bias and errors. An improved and haplotype-solved reference genome, especially from the parental isolate, as well as long-read sequencing with high accuracy (e.g., PacBio HiFi), is needed to solve such inconsistencies.

We took the advantage of our previously assembled high-continuity reference genome for comparison between the genetic map and the physical map (38, 39). This enabled us to discover new genetic features from a sexual *P. striiformis* f. sp. *tritici* population. Our study showed that *P. striiformis* f. sp. *tritici* has a large genetic map with a total genetic distance of 7,715.0 cM, which is comparable to other rust fungi. In the pine fusiform rust fungus (*C. quercuum* f. sp. *fusiforme*), a genetic map of 3,006 cM was constructed using 421 (including 208 randomly amplified polymorphic DNA [RAPD], 34 simple sequence repeat [SSR], and 184 amplified fragment length polymorphism [AFLP]) markers (51). Similarly, Anderson et al. (24) generated a genetic map of 5,860 cM using 13,412 restriction site-associated DNA sequence (RADseq) markers in the flax rust fungus (*M. lini*). One of the factors contributing to the slightly larger map in the present study might be the genotyping platform used since markers from whole-genome sequencing could potentially cover the whole genome. In the present study, 68.81 Mb, out of the 84.53-Mb (81.4%) genome, were covered by the markers from whole-genome sequencing. In the *M. lini* study, a slightly lower coverage (68.9%) of the genome was anchored to the genetic map using RADseq markers (24). We also noticed that *P. striiformis* f. sp. *tritici* has a slightly higher recombination rate (1.81 cM/10 kb) (Table 2) than other fungal pathogens, e.g., 0.3 to 0.5 cM/10 kb in *F. graminearum* (52), 1.25 cM/10 kb in *Z. tritici* (53), and 1.18 cM/10 kb in *M. lini* (24). We speculate that the higher recombination rate in *P. striiformis* f. sp. *tritici* is associated with the high number of detected single- and double-crossover events. Compared with the average of 114.6 crossovers per F_2 individual in *M. lini*, *P. striiformis* f. sp. *tritici* has a relatively high number of crossovers per individual, at 133.15 (Fig. 2C). In addition to the high number of crossover events, we also noticed that rust fungi with high recombination rates also have high percentages of CpG islands in their genomes. It has been proposed that the depletion of nucleosome occupancy in particular functional features such as CpG islands increases the accessibility of the recombination machinery (54). We found that the percentages of genome coverage by CpG islands in the basidiomycete rust fungi

are higher than those in the ascomycete fungi in our comparison. We noticed that *Z. tritici* was an exception in its slightly higher recombination rate than that of *M. lini* but much lower CpG island coverage (Table 3). Because only a few plant-pathogenic fungi have genome-wide recombination rates available, it is not possible to explicitly determine the role of CpG islands in the variation of recombination rates across different plant pathogens. Further studies are needed to test this hypothesis when more plant pathogens have their genome-wide recombination rates available. It is also useful to further investigate the genomic features of recombination hot spot regions besides the CpG islands to explain the recombination rate variation across the genome and among different plant pathogens.

In summary, our high-density genetic map reveals a generally high recombination rate of *P. striiformis* f. sp. *tritici*. More studies are needed to investigate the contribution of the high number of sexual recombination events to *P. striiformis* f. sp. *tritici* genome architecture and the rapid evolution of virulence. Moreover, our high-quality genetic map with dense markers will provide a valuable resource for anchoring genomic contigs to chromosomes.

In conclusion, using molecular markers generated through whole-genome sequencing of a self-fertilized population, we generated a high-density genetic map for *P. striiformis* f. sp. *tritici* comprising 41 lineage groups. Moreover, the SNP and indel markers are attached to the sequences of fragments that have been mapped to the reference genome, which allows direct comparison with data from future similar studies. The high-density genetic map will be valuable to further anchor fragmented contigs to chromosomes. Furthermore, the avirulence gene cluster of *AvYr44-AvYr7-AvYr43-AvYrExp2* was identified from QTL mapping and located at a short genome region through genome comparison. Further studies on the detailed genomic environment of this *Avr* cluster and cloning of these genes will shed light on the genomic basis of the rapid virulence changes of this destructive pathogen.

MATERIALS AND METHODS

Isolate selection, urediniospore multiplication, and teliospore production. *P. striiformis* f. sp. *tritici* isolate 12-368 (race PSTv-4), collected from Washington State in the United States in 2012, was selected to generate a segregating population based on its avirulence/virulence profile and its abilities to produce telia and infect barberry plants. The purification, urediniospore multiplication, and teliospore production of the isolate were conducted according to a previously described procedure (36). Briefly, the pure isolate was obtained from a single uredinium. Urediniospores were multiplied on seedlings of the winter wheat cultivar Nugaines (37). The spring wheat cultivar Avocet Sensitive (AvS) was inoculated at the flag-leaf stage and grown under controlled conditions to produce telia.

Developing a sexual population. A sexual population was developed by self-fertilization of isolate 12-368 according to procedures described previously (36, 55) (Fig. 1). Briefly, wheat leaves bearing telia were prepared by removing the epidermal layers and then placed on moist filter paper in petri dishes. Teliospore germination and basidiospore production were checked under a microscope periodically. Suspensions of germinated teliospores and produced basidiospores were used to inoculate 10-day-old barberry leaves. After incubation in a dew chamber at 10°C in the dark with 100% relative humidity for 36 to 48 h, the inoculated barberry plants were transferred to a growth chamber for pycnial production. For self-fertilization, the nectar containing pycniospores from one pycnium was transferred to another pycnium. To avoid duplicated fertilization of a pycnium, the transfer of nectar was conducted in one direction such that the nectar from the first pycnium was transferred to the second pycnium but the nectar from the second pycnium was never retransferred back to the first pycnium or another pycnium. As expected, in this way, an aecium was produced only in the opposite leaf surface of the second pycnium but not in the first pycnium. Usually, a cluster of aecial cups was formed after one pycnium was fertilized. When an aecium was mature, only a single aecial cup was excised with a sterile razor blade, the aeciospores from the aecial cup were used to inoculate seedlings of wheat cultivar Nugaines, and the leftover cups within one aecium were stored in liquid nitrogen for backup. About 13 to 15 days after aeciospore inoculation, a single uredinium from an inoculated Nugaines leaf was isolated, and the urediniospores from the uredinium were used to inoculate Nugaines seedlings again to multiply enough urediniospores, which were considered one progeny isolate produced from infection by a single aeciospore. The increased urediniospores of both parental isolate 12-368 and progeny isolates were used for virulence testing and DNA extraction.

Virulence phenotyping and genetic analysis. A total of 34 wheat genotypes, each carrying a single *Yr* resistance gene, were used to obtain avirulence/virulence phenotypes of isolate 12-368 and the progeny isolates. These 34 *Yr* genes were *Yr1*, *Yr5*, *Yr6*, *Yr7*, *Yr8*, *Yr9*, *Yr10*, *Yr15*, *Yr17*, *Yr24*, *Yr27*, *Yr32*, *Yr43*, *Yr44*, *YrSP*, *YrTr1*, *YrExp2*, *Yr76*, *Yr2*, *Yr21*, *Yr25*, *Yr26*, *Yr28*, *Yr29*, *Yr31*, *Yr35*, *YrCV*, *YrTr1*, *YrCN19*, *YrA*, *YrAvS*, *Yr45*, *Yr53*, and *Yr64*. Five to seven seeds of each wheat line were planted for each isolate test. About

18 days after inoculation, the IT data of an isolate on each plant were recorded based on a scale of 0 to 9 (56). As genetically pure seeds of each *Yr* single-gene line were specially produced and each *P. striiformis* f. sp. *tritici* isolate was obtained initially from a single uredinium and carefully multiplied, the up to seven plants of each line in a single isolate test mostly had identical ITs. In cases where different ITs were observed, the identical ITs of most plants or at least three plants were used. An isolate was considered avirulent to a specific *Yr* gene when the IT was 0 (inoculated leaves showing no visible symptoms), 1 (showing necrotic or chlorotic flecks), 2 (showing necrotic or chlorotic blotches without sporulation), or 3 to 6 (showing necrotic or chlorotic blotches with trace to moderate sporulation) or virulent when the IT was 7 to 9 (showing abundant sporulation with or without necrosis or chlorosis) (37). Since urediniospores of isolate 12-368 are dikaryotic, typical for *P. striiformis* f. sp. *tritici* and other rust fungi, we assumed that the parental isolate is heterokaryotic or "heterozygous" for many loci and therefore treated it as an F_1 generation in the present study. The progeny isolates generated by selfing isolate 12-368 were considered the F_2 generation. Therefore, the segregation of virulence phenotypes and molecular markers was expected to follow the segregation ratios of an F_2 population. The chi-squared test was used to determine whether the observed segregation of A/V phenotypes on a particular wheat line fit a theoretical genetic ratio.

DNA extraction and whole-genome sequencing. Genomic DNA was extracted from urediniospores using the cetyltrimethylammonium bromide (CTAB) method as previously described (57). The quality and quantity of the extracted DNA were checked using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific) and 0.8% agarose (Thermo Fisher Scientific) gel electrophoresis. DNA libraries were constructed and whole-genome sequencing was performed using Illumina HiSeq PE150 technology (Novogene Co. Ltd., Sacramento, CA).

Genomic variation calling. For each isolate, Illumina sequencing reads were checked and trimmed using Trimmomatic (version 0.36) (58). Parameters were set as LEADING:10 TRAILING:10 SLIDINGWINDOW:4: 15 MINLEN:50 in the paired-end (PE) model. After trimming, reads with lengths of less than 50 bp or that were not paired were excluded from subsequent analyses. Potential sequencing errors in trimmed reads were corrected using Lighter software (59), with the parameter set as –K 21 90000000. The high-quality and error-corrected Illumina reads were used for genomic variation calling. The previously assembled genome of U.S. isolate *P. striiformis* f. sp. *tritici* 93-210 was used as a reference genome (38) since it has a well-assembled genome (38). Also collected from the U.S. Pacific Northwest, the reference isolate was identified as belonging to race PSTv-20 and was avirulent to all 18 wheat *Yr* single-gene lines, except *Yr17*, used to differentiate *P. striiformis* f. sp. *tritici* races in the United States (37).

Genome-wide markers were identified according to a previously proposed framework (60). Briefly, the Burrows-Wheeler alignment (BWA) tool version 0.7.15 (61) was used to index the sequence reads to the reference genome, and the mem algorithm with default parameters was used to map filtered paired-end Illumina reads to the indexed reference genome. Next, Genome Analysis Toolkit (GATK) version 3.3 (62) was used to identify genomic variations. RealignerTargetCreator software was used to define interval targets for local realignment, and IndelRealigner was used to perform indel realignment. Two rounds of genomic variation calling were performed using GATK HaplotypeCaller with default parameters, as described previously (38, 63). VCFtools software (version 0.1.13) was used to manipulate SNPs and indels stored in vcf format. SNPs and indels of high confidence were filtered using VCFtools, with parameters set as –min-alleles 2 –max-alleles 2 –minQ 1000 –min-meanDP 30 –max-meanDP 60 –max-missing 1 (meaning that only biallelic SNPs and indels with a minimum quality of 1,000, a minimum mean depth of 30, a maximum mean depth of 60, and no missing data were used). By setting –min-meanDP 30 –max-meanDP 60, the variations in repetitive regions could be partially filtered since our sequencing depth ranged from 30× to 60×. To further remove potentially problematic variations in repetitive regions, we filtered SNPs and indels from transposable element (TE) regions as previously defined (38, 39). By doing this, genome-wide variations from only non-TE regions were kept for subsequent analyses. Moreover, duplicate markers (that is, markers with identical genotypes across all isolates) were identified and excluded from the construction of maps. To further reduce the computational burden, only one marker was selected within a 1,000-bp sequence region. The markers were named as contig no._position. For example, C001_1000 is the marker located in contig 1.001 at bp 1000 in the reference genome of *P. striiformis* f. sp. *tritici* 93-210 (38).

Genetic map construction and QTL mapping. Before map construction, we first checked the segregation ratios of markers. In general, markers with *P* values from chi-squared tests of $\geq$0.05 for the 1 (homozygous AA):2 (heterozygous AB):1 (homozygous BB) ratio expected for a single locus were used for map construction. In total, 2,637 high-quality genome-wide markers and 94 progeny isolates were used for map construction. Genetic map construction was performed using qtl version 1.41-6 (64, 65) and ASMap version 1.0-2 (66) in the R package. The qtl program was used to analyze the genotypes. Even though the progeny population generated from selfing the isolate 12-368 could be considered an F_2 population, the linkage phases were unknown due to the lack of paternal and maternal isolates. Therefore, some markers might have switched alleles. This issue was solved by checking logarithm of the odds (LOD) scores against the estimated recombination fractions for all marker pairs using the qtl program. The markers were considered to have alleles switched when they were tightly associated with other markers but had recombination fractions of $>$0.5, which were corrected according to a procedure described by Broman et al. (64). After correcting allele switches, the ASMap package using the minimum spanning tree algorithm was used to infer linkage groups and optimally order markers, with parameters set as dist.fun="kosambi," bychr=FALSE, p.value=1e-10, anchor=TRUE, noMap.dis=20, and detectBadData=TRUE. The genotyping errors and rates were calculated using the qtl package, and the number of observed crossovers per individual was estimated using the ASMap

package. The genetic map was reconstructed until no genotype errors were detected. In addition, we excluded isolates that were potentially contaminated if their large proportion of alternative alleles was absent from the parental isolate. Analysis with such potentially contaminated isolates showed that these isolates had >3-times-larger numbers of crossover events than other isolates. The number of single- and double-crossover events per individual was calculated using the countXO function in the R/qtl package (64, 65).

After the genetic map was constructed, we explored genome-wide features that might contribute to the high recombination rate of *P. striiformis* f. sp. *tritici*. The AT-rich regions were determined using OcculterCut v1.1 with default parameters (67). Since recombination often increases at the CpG islands in mammalian genomes (68), we also analyzed the distribution of CpG islands in *P. striiformis* f. sp. *tritici* and several other plant pathogens with genetic maps available. CpG islands were identified using CpGcluster v2.0 software (69), with a distance threshold of 75 and a *P* value of 1E−3.

The genetic map generated with ASMap was used for QTL mapping of *Avr* genes, and QTL mapping was performed using the qtl package. First, the QTL genotype probabilities given the available marker genotypes were calculated using the calc.genoprob function, with parameters set as step=1, error.prob=0.02, map.function="kosambi," and stepwidth="fixed." The recorded ITs were used as phenotypes for QTL mapping. The QTL mapping approach was used, instead of single-gene locus-like markers, to map *Avr* genes because it was not always possible to classify isolates into homozygous avirulent, homozygous virulent, and heterozygous as for the genotypic data. Since the ITs did not follow a normal distribution, the nonparametric interval mapping method was selected by setting model="np" in the scanone function. Instead of using an arbitrary LOD threshold for all phenotypes, we calculated 5% LOD thresholds via a permutation test for each phenotype by setting n.perm=1000. To determine the QTL confidence intervals, the Bayesian credible intervals were calculated using the bayesint function. QTL mapping was performed separately for the segregating phenotype data on each wheat *Yr* single-gene line.

Association analysis and genomic environment of the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster. To validate our QTL mapping results and to further identify potential *Avr*-associated variations, we used all the isolates from the sexual population for association analysis. Here, we focused on *AvYr44*, *AvYr7*, *AvYr43*, and *AvYrExp2* since (i) each of the avirulence phenotypes of these genes was controlled by a single gene and (ii) these four genes were mapped to a cluster in the genetic map and located in a single contig (contig 1.022) (see Results) in the reference genome. Genomic variations in contig 1.022 with <20% missing data were retrieved, and all the sequence variations were used for the GWAS, including those in repetitive regions and those with *P* values by chi-squared tests of less than 0.05. The FarmCPU method implemented in GAPIT v3 software was used to perform association analysis with three principal components as covariates (70). The associations between genomic variations and virulence phenotypes were considered significant if the *P* value was <1E−05 and the false discovery rate-adjusted *P* value (*P*_FDR_adj) was 0.01. Significantly associated variations were annotated based on our previous 93-210 reference genome (38).

To investigate the genomic environment of the *AvYr44-AvYr7-AvYr43-AvYrExp2* cluster, we compared the homologous regions of this cluster among three well-assembled *P. striiformis* f. sp. *tritici* genomes, namely, isolates 93-210 (38), 11-281 (41), and 104E 137A− (40). First, contig 1.022 of 93-210 was subjected to a BLAST search against the 11-281 and 104E 137A− genome sequences to detect potential homologous contigs. Next, the three homologous contigs were aligned using progressiveMauve software (71), and the generated alignment in a ".backbone" file was visualized using genoPlotR (72).

Data availability. All data sets generated for this study are included in the figures, tables, and supplemental material of the article. The complete set of sequence data was deposited in the National Center for Biotechnology Information (NCBI) database under accession no. PRJNA599033 and in the SRA under accession no. SRP239501. Further inquiries can be directed to the corresponding authors.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, RTF file, 1 MB.

FIG S2, RTF file, 0.4 MB.

TABLE S1, DOCX file, 0.03 MB.

DATA SET S1, XLSX file, 0.03 MB.

DATA SET S2, XLSX file, 1 MB.

DATA SET S3, XLSX file, 0.2 MB.

ACKNOWLEDGMENTS

This work was financially supported by the U.S. Department of Agriculture Agricultural Research Service (project no. 2090-22000-018-00D) and the Washington Grain Commission (projects 13C-3061-5682 and 13C-3061-3144) to X. Chen and the National Natural Science Foundation of China (grant no. 31901833) and the Ph.D. Foundation of Southwest University of Science and Technology (no. 19zx7116) to C. Xia. This study was conducted in the facilities of the U.S. Department of Agriculture Agricultural

Research Service and Washington State University (WSU), Pullman, WA (Department of Plant Pathology, College of Agricultural, Human, and Natural Resource Sciences, Agricultural Research Center, project no. WNP00663, Washington State University, Pullman, WA). We highly appreciate China Scholarship Council scholarships to C. Xia and Y. Lei.

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 U.S. Department of Agriculture. The USDA is an equal-opportunity provider and employer.

We highly appreciate technical support from the WSU Information Technology staff. We also acknowledge the Kamiak high-performance computing clusters at WSU for providing computing resources.

REFERENCES

1. Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL, Gurr SJ. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186–194. https://doi.org/10.1038/nature10947.
2. Talhinhas P, Batista D, Diniz I, Vieira A, Silva DN, Loureiro A, Tavares S, Pereira AP, Azinheira HG, Guerra-Guimarães L, Várzea V, Silva MDC. 2017. The coffee leaf rust pathogen *Hemileia vastatrix*: one and a half centuries around the tropics. Mol Plant Pathol 18:1039–1051. https://doi.org/10.1111/mpp.12512.
3. Goellner K, Loehrer M, Langenbach C, Conrath U, Koch E, Schaffrath U. 2010. *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. Mol Plant Pathol 11:169–177. https://doi.org/10.1111/j.1364-3703.2009.00589.x.
4. Nazareno ES, Li F, Smith M, Park RF, Kianian SF, Figueroa M. 2018. *Puccinia coronata* f. sp. *avenae*: a threat to global oat production. Mol Plant Pathol 19:1047–1060. https://doi.org/10.1111/mpp.12608.
5. Lawrence GJ, Dodds PN, Ellis JG. 2007. Rust of flax and linseed caused by *Melampsora lini*. Mol Plant Pathol 8:349–364. https://doi.org/10.1111/j.1364-3703.2007.00405.x.
6. Duplessis S, Major I, Martin F, Séguin A. 2009. Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. Crit Rev Plant Sci 28:309–334. https://doi.org/10.1080/07352680903241063.
7. Dean R, van Kan JA, Pretorius ZA, Hammond-Kosack KE, di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD. 2012. The top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol 13:414–430. https://doi.org/10.1111/j.1364-3703.2011.00783.x.
8. Flor HH. 1971. Current status of the gene-for-gene concept. Annu Rev Phytopathol 9:275–296. https://doi.org/10.1146/annurev.py.09.090171.001423.
9. Petit-Houdenot Y, Fudal I. 2017. Complex interactions between fungal avirulence genes and their corresponding plant resistance genes and consequences for disease resistance management. Front Plant Sci 8:1072. https://doi.org/10.3389/fpls.2017.01072.
10. Jones JDG, Dangl JL. 2006. The plant immune system. Nature 444:323–329. https://doi.org/10.1038/nature05286.
11. van Kan JAL, van den Ackerveken GFJM, de Wit PJGM. 1991. Cloning and characterization of cDNA of avirulence gene *avr9* of the fungal pathogen *Cladosporium fulvum*, causal agent of tomato leaf mold. Mol Plant Microbe Interact 4:52–59. https://doi.org/10.1094/mpmi-4-052.
12. van den Ackerveken GFJM, Van Kan JAL, de Wit PJGM. 1992. Molecular analysis of the avirulence gene *avr9* of the fungal tomato pathogen *Cladosporium fulvum* fully supports the gene-for-gene hypothesis. Plant J 2:359–366. https://doi.org/10.1046/j.1365-313X.1992.t01-34-00999.x.
13. Yin CT, Hulbert S. 2011. Prospects for functional analysis of effectors from cereal rust fungi. Euphytica 179:57–67. https://doi.org/10.1007/s10681-010-0285-x.
14. Attard A, Gout L, Gourgues M, Kühn M-L, Schmit J, Laroche S, Ansan-Melayah D, Billault A, Cattolico L, Balesdent M-H, Rouxel T. 2002. Analysis of molecular markers genetically linked to the *Leptosphaeria maculans* avirulence gene *AvrLm1* in field populations indicates a highly conserved event leading to virulence on Rlm1 genotypes. Mol Plant Microbe Interact 15:672–682. https://doi.org/10.1094/MPMI.2002.15.7.672.
15. Gout L, Fudal I, Kuhn ML, Blaise F, Eckert M, Cattolico L, Balesdent MH, Rouxel T. 2006. Lost in the middle of nowhere: the *AvrLm1* avirulence gene of the dothideomycete *Leptosphaeria maculans*. Mol Microbiol 60:67–80. https://doi.org/10.1111/j.1365-2958.2006.05076.x.
16. Fudal I, Ross S, Gout L, Blaise F, Kuhn ML, Eckert MR, Cattolico L, Bernard-Samain S, Balesdent MH, Rouxel T. 2007. Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* genome: map-based cloning of *AvrLm6*. Mol Plant Microbe Interact 20:459–470. https://doi.org/10.1094/MPMI-20-4-0459.
17. Parlange P, Daverdin G, Fudal I, Kuhn ML, Balesdent MH, Blaise F, Grezes-Besset B, Rouxel T. 2009. *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change. Mol Microbiol 71:851–863. https://doi.org/10.1111/j.1365-2958.2008.06547.x.
18. Balesdent M-H, Fudal I, Ollivier B, Bally P, Grandaubert J, Eber F, Chèvre A-M, Leflon M, Rouxel T. 2013. The dispensable chromosome of *Leptosphaeria maculans* shelters an effector gene conferring avirulence towards *Brassica rapa*. New Phytol 198:887–898. https://doi.org/10.1111/nph.12178.
19. Ghanbarnia K, Fudal I, Larkan NJ, Links MG, Balesdent MH, Profotova B, Fernando WGD, Rouxel T, Borhan MH. 2015. Rapid identification of the *Leptosphaeria maculans* avirulence gene *AvrLm2*, using an intraspecific comparative genomics approach. Mol Plant Pathol 16:699–709. https://doi.org/10.1111/mpp.12228.
20. Plissonneau C, Daverdin G, Ollivier B, Blaise F, Degrave A, Fudal I, Rouxel T, Balesdent MH. 2016. A game of hide and seek between avirulence genes *AvrLm4-7* and *AvrLm3* in *Leptosphaeria maculans*. New Phytol 209:1613–1624. https://doi.org/10.1111/nph.13736.
21. Linning R, Lin D, Lee N, Abdennadher M, Gaudet D, Thomas P, Mills D, Kronstad JW, Bakkeren G. 2004. Marker-based cloning of the region containing the *UhAvr1* avirulence gene from the basidiomycete barley pathogen *Ustilago hordei*. Genetics 166:99–111. https://doi.org/10.1534/genetics.166.1.99.
22. Ali S, Laurie JD, Linning R, Cervantes-Chávez JA, Gaudet D, Bakkeren G. 2014. An immunity-triggering effector from the barley smut fungus *Ustilago hordei* resides in an Ustilaginaceae-specific cluster bearing signs of transposable element-assisted evolution. PLoS Pathog 10:e1004223. https://doi.org/10.1371/journal.ppat.1004223.
23. Dodds PN, Lawrence GJ, Catanzariti A-M, Ayliffe MA, Ellis JG. 2004. The *Melampsora lini AvrL567* avirulence genes are expressed in haustoria and their products are recognized inside plant cells. Plant Cell 16:755–768. https://doi.org/10.1105/tpc.020040.
24. Anderson C, Khan MA, Catanzariti AM, Jack CA, Nemri A, Lawrence GJ, Upadhyaya NM, Hardham AR, Ellis JG, Dodds PN, Jones DA. 2016. Genome analysis and avirulence gene cloning using a high-density RADseq linkage map of the flax rust fungus, *Melampsora lini*. BMC Genomics 17:667. https://doi.org/10.1186/s12864-016-3011-9.
25. Salcedo A, Rutter W, Wang S, Akhunova A, Bolus S, Chao S, Anderson N, De Soto MF, Rouse M, Szabo L, Bowden RL, Dubcovsky J, Akhunov E. 2017. Variation in the *AvrSr35* gene determines *Sr35* resistance against wheat stem rust race Ug99. Science 358:1604–1606. https://doi.org/10.1126/science.aao7294.
26. Chen J, Upadhyaya NM, Ortiz D, Sperschneider J, Li F, Bouton C, Breen S, Dong C, Xu B, Zhang X, Mago R, Newell K, Xia X, Bernoux M, Taylor JM, Steffenson B, Jin Y, Zhang P, Kanyuka K, Figueroa M, Ellis JG, Park RF, Dodds PN. 2017. Loss of *AvrSr50* by somatic exchange in stem rust leads to virulence for *Sr50* resistance in wheat. Science 358:1607–1610. https://doi.org/10.1126/science.aao4810.

27. Chen XM. 2005. Epidemiology and control of stripe rust [Puccinia striiformis f. sp. tritici] on wheat. Can J Plant Pathol 27:314–337. https://doi.org/10.1080/07060660509507230.
28. Wellings CR. 2011. Global status of stripe rust: a review of historical and current threats. Euphytica 179:129–141. https://doi.org/10.1007/s10681-011-0360-y.
29. Chen W, Wellings C, Chen X, Kang Z, Liu T. 2014. Wheat stripe rust (yellow) rust caused by *Puccinia striiformis* f. sp. *tritici*. Mol Plant Pathol 15:433–446. https://doi.org/10.1111/mpp.12116.
30. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. 2019. The global burden of pathogens and pests on major food crops. Nat Ecol Evol 3:430–439. https://doi.org/10.1038/s41559-018-0793-y.
31. Jin Y, Szabo L, Carson M. 2010. Century-old mystery of *Puccinia striiformis* life history solved with the identification of *Berberis* as an alternate host. Phytopathology 100:432–435. https://doi.org/10.1094/PHYTO-100-5-0432.
32. Wang MN, Chen XM. 2013. First report of Oregon grape (*Mahonia aquifolium*) as an alternate host for the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) under artificial inoculation. Plant Dis 97:839. https://doi.org/10.1094/PDIS-09-12-0864-PDN.
33. Zhao J, Wang M, Chen X, Kang Z. 2016. Role of alternate hosts in epidemiology and pathogen variation of cereal rusts. Annu Rev Phytopathol 54:207–228. https://doi.org/10.1146/annurev-phyto-080615-095851.
34. Tian Y, Zhan G, Chen X, Tungruentragoon A, Lu X, Zhao J, Huang L, Kang Z. 2016. Virulence and simple sequence repeat marker segregation in a *Puccinia striiformis* f. sp. *tritici* population produced by selfing a Chinese isolate on *Berberis shensiana*. Phytopathology 106:185–191. https://doi.org/10.1094/PHYTO-07-15-0162-R.
35. Wang L, Zheng D, Zuo S, Chen X, Zhuang H, Huang L, Kang Z, Zhao J. 2018. Inheritance and linkage of virulence genes in Chinese predominant race CYR32 of the wheat stripe rust pathogen *Puccinia striiformis* f. sp. *tritici*. Front Plant Sci 9:120. https://doi.org/10.3389/fpls.2018.00120.
36. Yuan C, Wang M, Skinner DZ, See DR, Xia C, Guo X, Chen X. 2018. 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 108:133–141. https://doi.org/10.1094/PHYTO-04-17-0139-R.
37. Wan A, Chen X. 2014. 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 98:1534–1542. https://doi.org/10.1094/PDIS-01-14-0071-RE.
38. Xia C, Wang M, Yin C, Cornejo OE, Hulbert SH, Chen X. 2018. 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 31:1117–1120. https://doi.org/10.1094/MPMI-04-18-0107-A.
39. Xia C, Wang M, Yin C, Cornejo OE, Hulbert SH, Chen X. 2018. Genomic insights into host adaptation between the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) and the barley stripe rust pathogen (*Puccinia striiformis* f. sp. *hordei*). BMC Genomics 19:664. https://doi.org/10.1186/s12864-018-5041-y.
40. Schwessinger B, Sperschneider J, Cuddy WS, Garnica DP, Miller ME, Taylor JM, Dodds PN, Figueroa M, Park RF, Rathjen JP. 2018. A near-complete haplotype-phased genome of the dikaryotic wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici* reveals high interhaplotype diversity. mBio 9:e02275-17. https://doi.org/10.1128/mBio.02275-17.
41. Li Y, Xia C, Wang M, Yin C, Chen X. 2019. Genome sequence resource of a *Puccinia striiformis* isolate infecting wheatgrass. Phytopathology 109:1509–1512. https://doi.org/10.1094/PHYTO-02-19-0054-A.
42. Bourras S, McNally KE, Müller MC, Wicker T, Keller B. 2016. Avirulence genes in cereal powdery mildews: the gene-for-gene hypothesis 2.0. Front Plant Sci 7:241. https://doi.org/10.3389/fpls.2016.00241.
43. Dracatos PM, Zhang P, Park RF, McIntosh RA, Wellings CR. 2016. Complementary resistance genes in wheat selection 'Avocet R' confer resistance to stripe rust. Theor Appl Genet 129:65–75. https://doi.org/10.1007/s00122-015-2609-7.
44. Statler GD. 2000. Inheritance of virulence of *Puccinia triticina* culture X47, the F1 of the cross 71-112 × 70-1. Can J Plant Pathol 22:276–279. https://doi.org/10.1080/07060660009500475.
45. Zambino PJ, Kubelik AR, Szabo LJ. 2000. Gene action and linkage of avirulence genes to DNA markers in the rust fungus *Puccinia graminis*. Phytopathology 90:819–826. https://doi.org/10.1094/PHYTO.2000.90.8.819.
46. Webb CA, Fellers JP. 2006. Cereal rust fungi genomics and the pursuit of virulence and avirulence factors. FEMS Microbiol Lett 264:1–7. https://doi.org/10.1111/j.1574-6968.2006.00400.x.
47. Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol 10:417–430. https://doi.org/10.1038/nrmicro2790.
48. Rouxel T, Balesdent M. 2017. Life, death and rebirth of avirulence effectors in a fungal pathogen of *Brassica* crops, *Leptosphaeria maculans*. New Phytol 214:526–532. https://doi.org/10.1111/nph.14411.
49. Orbach MJ, Farrall L, Sweigard JA, Chumley FG, Valent B. 2000. A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. Plant Cell 12:2019–2032. https://doi.org/10.1105/tpc.12.11.2019.
50. Zhong Z, Marcel TC, Hartmann FE, Ma X, Plissonneau C, Zala M, Ducasse A, Confais J, Compain J, Lapalu N, Amselem J, McDonald BA, Croll D, Palma-Guerrero J. 2017. A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the *Stb6* resistance gene. New Phytol 214:619–631. https://doi.org/10.1111/nph.14434.
51. Kubisiak TL, Anderson CL, Amerson HV, Smith JA, Davis JM, Nelson CD. 2011. A genomic map enriched for markers linked to *Avr1* in *Cronartium quercuum* f. sp. *fusiforme*. Fungal Genet Biol 48:266–274. https://doi.org/10.1016/j.fgb.2010.09.008.
52. Laurent B, Palaiokostas C, Spataro C, Moinard M, Zehraoui E, Houston RD, Foulongne-Oriol M. 2018. High-resolution mapping of the recombination landscape of the phytopathogen *Fusarium graminearum* suggests two-speed genome evolution. Mol Plant Pathol 19:341–354. https://doi.org/10.1111/mpp.12524.
53. Lendenmann MH, Croll D, Stewart EL, McDonald BA. 2014. Quantitative trait locus mapping of melanization in the plant pathogenic fungus *Zymoseptoria tritici*. G3 (Bethesda) 4:2519–2533. https://doi.org/10.1534/g3.114.015289.
54. de Castro E, Soriano I, Marín L, Serrano R, Quintales L, Antequera F. 2012. Nucleosomal organization of replication origins and meiotic recombination hotspots in fission yeast. EMBO J 31:124–137. https://doi.org/10.1038/emboj.2011.350.
55. Wang MN, Chen XM. 2015. Barberry does not function as an alternate host for *Puccinia striiformis* f. sp. *tritici* in the U. S. Pacific Northwest due to teliospore degradation and barberry phenology. Plant Dis 99:1500–1506. https://doi.org/10.1094/PDIS-12-14-1280-RE.
56. Line RF, Qayoum A. 1992. Virulence, aggressiveness, evolution, and distribution of races of Puccinia striiformis (the cause of stripe rust of wheat) in North America, 1968-87. Technical bulletin number 1788. US Department of Agriculture Agricultural Research Service, Pullman, WA.
57. Xia C, Wan A, Wang M, Jiwan DA, See DR, Chen X. 2016. 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 120:729–744. https://doi.org/10.1016/j.funbio.2016.02.007.
58. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
59. Song L, Florea L, Langmead B. 2014. Lighter: fast and memory-efficient sequencing error correction without counting. Genome Biol 15:509. https://doi.org/10.1186/s13059-014-0509-9.
60. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498. https://doi.org/10.1038/ng.806.
61. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698.
62. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. 2013. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinformatics 43:11.10.1–11.10.33. https://doi.org/10.1002/0471250953.bi1110s43.
63. Xia C, Wang M, Cornejo OE, Jiwan DA, See DR, Chen X. 2017. 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 8:2394. https://doi.org/10.3389/fmicb.2017.02394.
64. Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890. https://doi.org/10.1093/bioinformatics/btg112.

65. Broman KW. 2010. Genetic map construction with R/qtl. Technical report #214. University of Wisconsin—Madison Department of Biostatistics and Medical Informatics, Madison, WI.
66. Wu Y, Bhat P, Close TJ, Lonardi S. 2008. Efficient and accurate construction of genetic linkage maps from minimum spanning tree of a graph. PLoS Genet 4:e1000212. https://doi.org/10.1371/journal.pgen.1000212.
67. Testa AC, Oliver RP, Hane JK. 2016. OcculterCut: a comprehensive survey of AT-rich regions in fungal genomes. Genome Biol Evol 8:2044–2064. https://doi.org/10.1093/gbe/evw121.
68. Smeds L, Mugal CF, Qvarnström A, Ellegren H. 2016. High resolution mapping of crossover and non-crossover recombination events by whole-genome re-sequencing of an avian pedigree. PLoS Genet 12: e1006044. https://doi.org/10.1371/journal.pgen.1006044.
69. Hackenberg M, Previti C, Luque-Escamilla PL, Carpena P, Martínez-Aroza J, Oliver JL. 2006. CpGcluster: a distance-based algorithm for CpG-island detection. BMC Bioinformatics 7:446. https://doi.org/10.1186/1471-2105-7-446.
70. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. PLoS Genet 12:e1005767. https://doi.org/10.1371/journal.pgen.1005767.
71. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.
72. Guy L, Kultima JR, Andersson SGE. 2010. genoPlotR: comparative gene and genome visualization in R. Bioinformatics 26:2334–2335. https://doi.org/10.1093/bioinformatics/btq413.