A peer-reviewed version of this preprint was published in PeerJ on 23 April 2018.

View the peer-reviewed version (peerj.com/articles/4678), which is the preferred citable publication unless you specifically need to cite this preprint.

Zhang G, Hua Z. 2018. Genome comparison implies the role of Wsm2 in membrane trafficking and protein degradation. PeerJ 6:e4678 https://doi.org/10.7717/peerj.4678

Genome comparison implies the role of Wsm2 in membrane trafficking and protein degradation

Guorong Zhang [1] , **Zhihua Hua** Corresp. [2, 3]

[1] Agricultural Research Center-Hays, Kansas State University, Hays, Kansas, United States

[2] Department of Environmental and Plant Biology, Ohio University, Athens, Ohio, United States

[3] Interdisciplinary Program in Molecular and Cellular Biology, Ohio University, Athens, Ohio, United States

Corresponding Author: Zhihua Hua
Email address: hua@ohio.edu

Wheat streak mosaic virus (WSMV) causes streak mosaic disease in wheat (*Triticum aestivum* L.) and has been an important constraint limiting wheat production in many regions around the world. *Wsm2* is the only resistance gene discovered in wheat genome and has been located in a short genomic region of its chromosome 3B. However, the sequence nature and the biological function of *Wsm2* remain unknown due to the difficulty of genetic manipulation in wheat. In this study, we tested WSMV infectivity among wheat and its two closely related grass species, rice (*Oryza sativa*) and *Brachypodium distachyon*. Based on the phenotypic result and previous genomic studies, we developed a novel bioinformatics pipeline for interpreting a potential biological function of *Wsm2* and its ancestor locus in wheat. In the WSMV resistance tests, we found that rice has a WMSV resistance gene while *Brachypodium* does not, which allowed us to hypothesize the presence of a *Wsm2* ortholog in rice. Our OrthoMCL analysis of protein coding genes on wheat chromosome 3B and its syntenic chromosomes in rice and *Brachypodium* discovered 4,035 OrthoMCL groups as preliminary candidates of *Wsm2* orthologs. Given that *Wsm2* is likely duplicated through an intrachromosomal illegitimate recombination and that *Wsm2* is dominant, we inferred that this new WSMV-resistance gene acquired an activation domain, lost an inhibition domain, or gained high expression compared to its ancestor locus. Through comparison, we identified that 67, 16, and 10 out of 4,035 OrthoMCL orthologous groups contain a rice member with 25% shorter or longer in length, or 10 fold more expression, respectively, than those from wheat and *Brachypodium.* Taken together, we predicted a total of 93 good candidates for a *Wsm2* ancestor locus. All of these 93 candidates are not tightly linked with *Wsm2*, indicative of the role of illegitimate recombination in the birth of *Wsm2*. Further sequence analysis suggests that the protein products of *Wsm2* may combat WSMV disease through a molecular mechanism involving protein degradation and/or membrane trafficking. The 93 putative *Wsm2* ancestor loci

discovered in this study could serve as good candidates for future genetic isolation of the true *Wsm2* locus.

Genome comparison implies the role of Wsm2 in membrane trafficking and protein degradation

Guorong Zhang[1] and Zhihua Hua[2,3]

[1]*Agricultural Research Center-Hays, Kansas State University, Hays, KS 67601, USA;*
[2]*Department of Environmental and Plant Biology, Ohio University, Athens, Ohio 45701, USA;*
[3]*Interdisciplinary Program in Molecular and Cellular Biology, Ohio University, Athens, Ohio 45701, USA;*

Running title: Wsm2 in trafficking and degradation

***Corresponding author:** Zhihua Hua, Department of Environmental and Plant Biology, Ohio University, Athens, Ohio 45701 USA. Tel: +01 740 593 1123, Email: hua@ohio.edu

Abstract

Wheat streak mosaic virus (WSMV) causes streak mosaic disease in wheat (*Triticum aestivum* L.) and has been an important constraint limiting wheat production in many regions around the world. *Wsm2* is the only resistance gene discovered in wheat genome and has been located in a short genomic region of its chromosome 3B. However, the sequence nature and the biological function of *Wsm2* remain unknown due to the difficulty of genetic manipulation in wheat. In this study, we tested WSMV infectivity among wheat and its two closely related grass species, rice (*Oryza sativa*) and *Brachypodium distachyon*. Based on the phenotypic result and previous genomic studies, we developed a novel bioinformatics pipeline for interpreting a potential biological function of *Wsm2* and its ancestor locus in wheat. In the WSMV resistance tests, we found that rice has a WMSV resistance gene while *Brachypodium* does not, which allowed us to hypothesize the presence of a *Wsm2* ortholog in rice. Our OrthoMCL analysis of protein coding genes on wheat chromosome 3B and its syntenic chromosomes in rice and *Brachypodium* discovered 4,035 OrthoMCL groups as preliminary candidates of *Wsm2* orthologs. Given that *Wsm2* is likely duplicated through an intrachromosomal illegitimate recombination and that *Wsm2* is dominant, we inferred that this new WSMV-resistance gene acquired an activation domain, lost an inhibition domain, or gained high expression compared to its ancestor locus. Through comparison, we identified that 67, 16, and 10 out of 4,035 OrthoMCL orthologous groups contain a rice member with 25% shorter or longer in length, or 10 fold more expression, respectively, than those from wheat and *Brachypodium.* Taken together, we predicted a total of 93 good candidates for a *Wsm2* ancestor locus. All of these 93 candidates are not tightly linked with *Wsm2*, indicative of the role of illegitimate recombination in the birth of *Wsm2*. Further sequence analysis suggests that the protein products of *Wsm2* may combat WSMV disease through a molecular mechanism involving protein degradation and/or membrane trafficking. The 93 putative *Wsm2* ancestor loci discovered in this study could serve as good candidates for future genetic isolation of the true *Wsm2* locus.

Introduction

Wheat streak mosaic virus (WSMV) causes streak mosaic disease in wheat (*Triticum aestivum* L.) and has been reported in many regions around the world (Fahim et al. 2011; Sharp et al. 2002). The WSMV is transmitted by wheat curl mites (WCM: *Aceria tosichella* Keifer) (Navia et al. 2013) and wheat is the preferred host for both WCM and WSMV (Baleya et al. 2001; Murray & Brennan 2009). WSMV-infected wheat plants develop yellow leaf streaks during early infection and the symptom could spread to the entire leaf if the virus is not effectively controlled. Stunted growth is also common in severely infected plants. Price et al. (Price et al. 2010) found that WSMV infection could reduce root development and affect water use efficiency. WSMV is one important constraint limiting wheat production in the Great Plains of United States. The average yield loss in this region was estimated about 2% per year (Appel et al. 2013; Christian & Willis 1993). But, up to 13% reduction in wheat production due to WSMV disease has been reported in Kansas, USA (Sim et al. 1988). In severe cases wheat production could be completely destroyed by WSMV (McNeil et al. 1996).

Unfortunately, effective chemicals are not yet available for controlling WSMV and its WCM vector (Tan et al. 2017). Host resistance is the primary and effective way to suppress WSMV. To date, three WSMV resistance genes, *Wsm1*, *Wsm2*, and *Wsm3*, have been identified. Both *Wsm1* and *Wsm3* were found in a wild relative, *Thinopyrum intermedium* (Host) Barkworth & D.R. Dewey, and they have been introduced into the wheat genome through translocation (Gill et al. 1995; Triebe et al. 1991). However, alien translocation often results in yield penalty due to the incorporation of non-adapted genes. For example, lines introgressed with *Wsm1* showed various yield reductions ranging from 11 to 28% (Sharp et al. 2002), limiting the breeding application of this type of resistant sources. *Wsm2* was discovered in a wheat breeding line CO960293-2 (Haley et al. 2002). Genetic studies have shown that the WSMV resistance in CO960293-2 is controlled by a single dominant allele (*Wsm2*) and it has been genetically mapped on chromosome arm 3BS of the wheat genome (Lu et al. 2011). Recently, our group and others have further located *Wsm2* into 0.4 cM region flanked by *BS00022387_51* and *BS00088683_51* using a dense microarray containing 90,000 single nucleotide polymorphic (SNP) sites (Assanga et al. 2017). However, its sequence nature and biological function still remain elusive.

Wsm2 has been introduced into several wheat cultivars to acquire WSMV resistance without compromising yield, such as “RonL” (Martin et al. 2007), “Snowmass” (Haley et al. 2011), “Clara

CL" (Martin et al. 2014), "Oakley CL" (Zhang et al. 2015a), and "Joe" (Zhang et al. 2016a), demonstrating a great potential in improving WSMV resistance. Through allelic test and *Wsm2*-linked marker analysis in nine wheat lines we have shown that *Wsm2* and/or its tightly linked genes are primarily responsible for the WSMV resistance (Zhang et al. 2015b; Zhang et al. 2016b). Interestingly, the functions of three WSMV-resistant genes, *Wsm1*, *Wsm2*, and *Wsm3,* are all temperature sensitive, which hold their proper WSMV resistant functions up to 20, 18, and 24°C, respectively (Gill et al. 2008; Seifers et al. 2013a; Seifers et al. 2013b; Seifers et al. 1995). A recent study indicated that the replication and movement of WSMV, and the disease symptom development were greatly affected by temperature (Wosula et al. 2017). It remains unclear whether the temperature-dependent effectiveness of WSMV resistance genes is due to the variable pathogenesis of WSMV under different temperature conditions and/or caused by the temperature-mediated gene expression and functional regulation. Only if we isolate these genes would we be able to address these questions more closely. In this work, we applied cross species phenotypic and genomic analyses and discovered that *Wsm2* might encode a protein involved in membrane trafficking and protein degradation.

Materials and Methods

WSMV infectivity tests on rice and *Brachypodium distachyon*

Rice (*Oryza sativa* ssp. *Japonica*) cultivar Nipponbare, *B. distachyon* (*Brachypodium* hereafter) accession Bd21-3, WSMV-resistant wheat cultivar RonL, and WSMV-susceptible wheat cultivar "Tomahawk" were seeded in rows in two metal flats (21 x 31 cm) filled with a potting mix (Sungro, Vacouver, Canada). Each line had two replications with 12 seeds per replication (row) in each flat. At the two-leaf stage, plants were mechanically inoculated (finger-rubbing) with a WSMV isolate, Sidney 81. Inoculum preparation was done as described in Seifers et al. (Seifers et al. 2006). In brief, infected wheat leaf tissues were grounded at a 1:10 (wt/vol) dilution in 0.02 M potassium phosphate buffer (pH 7) and filtered through cheesecloth. This extract was used as inoculum after adding abrasive (Crystolon flour B, 600 mesh; Norton Co., Worcester, MA) with a concentration of 0.01 g/mL. This method of inoculum preparation and inoculation was used throughout the study. After inoculation, two flats were maintained in different growth chambers (Percival Model PGC-15WC) with one set at 18°C and the other kept at 22°C under a short-day photoperiod condition [12 h fluorescent light (250$\mu Es^{-1}m^{-2}$) and 12 h darkness]. Four weeks after inoculation, indirect enzyme-linked immunosorbent assays (ELISA) were conducted for each plant as described in Seifers et al. (Seifers et al. 2006). Two leaf tissue bulks from non-inoculated Tomahawk plants (healthy check) were included as healthy

controls. The GHV value (Sample ELISA value / Healthy ELISA value) was calculated for each plant. The plant was considered as susceptible if its GHV was greater than 2 (Seifers et al. 2006). The percentage of resistant plants was calculated for each line in each replication. The ANOVA was conducted for the percentage of resistant plants using GLM model with SAS 9.4 and the least significant difference (LSD) at $\alpha = 0.01$ was used to conduct comparison among four genotypes.

A further infectiveness test was conducted on additional 44 *Brachypodium* accessions (Gordon et al. 2017) to examine if there is any variation of WSMV resistance. All 44 accessions plus Bd21-3 were planted in rows in two metal flats (30 x 50 cm) with eight seeds per row. Each accession was planted in one row while an additional row of Bd21-3 was planted as the healthy check. At the two-leaf stage, all plants except for the healthy check were mechanically inoculated with Sidney 81 as described above. After inoculation, both flats were maintained in a growth chamber set at 18°C with the same short-day photoperiod as aforementioned. Four weeks after inoculation, ELISA tests were conducted for each plant including the healthy check plants. The GHV was calculated and used to determine the WSMV susceptibility or resistance. The percentage of susceptible plants (infection rate) was calculated for each accession.

Back assay with inoculum made from infected Bd21-3

Symptomatic leaf tissues from Bd21-3 were bulked and six different dilution rates of inoculums (1:5, 1:10, 1:20, 1:40, 1:80, and 1:160 wt/vol) were made as described above. A new batch of Bd21-3 seedlings were raised at 18°C under a short-day growth condition. Bd21-3 plants were grown in seven rows in one metal flat (21 X 31 cm). At the two-leaf stage, each row was inoculated with a different dilution. The last row was not inoculated and used as the healthy check. Four weeks after inoculation, ELISA tests were conducted on every plant. The GHV of each inoculated plant was calculated and used to determine the WSMV susceptibility or resistance. The percentage of susceptible plants (infection rate) was calculated for each dilution rate.

Orthology relationship analysis

To define orthologous genes in *Brachypodium*, rice, and wheat that likely encode Wsm2, the protein sequences of previously annotated genes on *Brachypodium* chromosome 2, rice chromosome 1, and wheat chromosome 3B were retrieved from each genome project (Choulet et al. 2014; International Brachypodium 2010; Kawahara et al. 2013). The sequences were

combined and analyzed for orthology relationships based on their similarities using OrthoMCL (Li et al., 2003). Briefly, an all-against-all BLASTp search (Altschul et al. 1990) was performed to find sequence similarity between each pair of sequences. The resulting sequence similarity matrix was subjected to a Markov Cluster Algorithm (MCL) clustering analysis to define orthologous groups among three species. An inflation value of 1.5 was identified to be the best to yield all putative OrthoMCL groups among three species analyzed.

Expression analysis

We used the number of expression sequence tags (EST) to represent the relative expression level of an orthologous gene in *Brachypodium*, rice, and wheat. To identify the number of ESTs of each orthologous member, the EST sequences of each species were downloaded from the EST database at Genbank (https://www.ncbi.nlm.nih.gov/nucest). The coding sequence of an orthologous gene from each species was used as a query for BLASTn search against its EST database (Altschul et al. 1990). An EST was considered to reflect a true expression of a query gene if 1) it had >95% identity to the query coding sequence, 2) the aligned sequences cover at least 75% of the EST or the query sequence, and 3) at least 50 nucleotides of the EST was included in the alignment (Hua et al. 2011). To compare expression of orthologous genes across three species, the absolute EST value of each gene was normalized by the total EST number of the corresponding species.

Functional prediction of a putative *Wsm2* candidate

The protein sequence of a putative *Wsm2* candidate gene was used as a query to search against the Pfam-A protein-protein interaction database (https://pfam.xfam.org, Version 31) by HMMER3, an accelerated profile hidden Markov model (profile HMM) search tool (Eddy 2011). The presence of a predicted Pfam-A protein-protein interaction domain (e-value cutoff ≤1) was used to categorize the putative biological function of a candidate.

Identification of the physical position of a *Wsm2* ancestor locus

The physical position (coordinate) of a putative *Wsm2* ancestor locus was retrieved based on the Generic Feature Format (GFF3) file from the wheat chromosome 3B genome project (Choulet et al. 2014). The distribution of putative *Wsm2* ancestor loci were visualized by plotting each locus on chromosome 3B.

Results

Absence of WSMV resistance genes in wild species

To examine whether a WSMV-resistant gene could be generated through spontaneous natural mutations in a wild species, we asked whether *Brachypodium*, a strictly self-pollinated species, is resistant to WSMV. Since all three up-to-date identified *Wsm* loci are temperature sensitive, we carried out the WSMV infectiveness tests on *Brachypodium* at two different temperatures, 18°C and 22°C, at which the *Wsm2*-containing wheat cultivar RonL shows WSMV resistant and susceptible phenotypes, respectively (Seifers et al. 2013b).

In the initial screen (Table 1), the susceptible wheat check cultivar Tomahawk was severely infected at both temperature conditions, indicating the effectiveness of the WSMV isolate Sidney 81 in this study. As a positive control, over 90% of the RonL plants were resistant to WSMV at 18°C while all individuals displayed a susceptible symptom at 22°C, confirming the temperature-sensitive phenotype of *Wsm2* in RonL. Among 35 *Brachypododium* Bd21-3 plants examined (17 plants at 18°C and 18 plants at 22°C), 80% (14 plants at each temperature) were susceptible to WSMV based on ELISA tests (GHV > 2, which indicates WSMV susceptibility (Seifers et al. 2006)). Statistically, no significant difference ($p < 0.01$, ANOVA test) was observed between WSMV-susceptible Tomahawk and Bd21-3 at both temperatures (Table 1), suggesting that Bd21-3 does not express any WSMV resistance gene. To confirm the susceptibility of Bd21-3 to WSMV, we performed a back assay, which used extracts from infected Bd21-3 plants as pathogen sources. To further understand the dynamic infection of WSMV, the original extract was diluted in a series of concentrations and used to inoculate Bd21-3 seedlings. The overall infection rates ranged from 50 to 100% (Table 3). The first five dilutions were very infective and have infected most of inoculated plants (83.3~100%). Therefore, our data indicate that Bd21-3 is WSMV susceptible.

Since new sources of temperature-sensitive resistance to WSMV have been identified from a large collection of wheat accessions (Seifers et al. 2013a), we then asked whether any *Brachypodium* natural variants could possess a WSMV-resistant gene. In total, 44 *Brachypodium* accessions with extensive genetic variations (Gordon et al. 2017) were selected for WSMV-infectiveness analysis (Table 2). All accessions together with Bd21-3 were grown under the same condition (see Materials and Methods) and inoculated with the same WSMV isolate Sidney 81 under 18°C as described above. Four weeks after inoculation, ELISA tests were conducted to examine the susceptibility/resistance of individuals. Unexpectedly, all the accessions had a 100% infection rate and all of them displayed a greater ELISA value than the

uninoculated Bd21-3 control and large GHVs ranging from 56.3 to 112.1 (Table 2), suggesting that *Brachypodium* might not contain any genetic sources for WSMV resistance and that a WSMV resistance gene is not likely attributed to spontaneous natural mutations.

Presence of a WSMV resistance gene in rice

Since *Brachypodium* is a naturally self-pollinated wild species (International Brachypodium 2010) and wheat is a crop, we next asked whether another crop species, rice (*Oryza sativa*), could contain a WSMV resistance gene. Both wheat and rice have been domesticated for over 10,000 years (Meyer et al. 2012) and it is known that domestication has significantly changed genome arrangement of crops from their wild relatives by fixing elite agronomic traits that benefits agricultural production (Chantret et al. 2005). We chose Nipponbare as a test rice cultivar because of the availability of its well-annotated genome (Kawahara et al. 2013). We performed WSMV infectiveness tests on Nipponbare together with *Brachypodium* and wheat lines at both 18 and 22°C (see Materials and Methods). Interestingly, among 20 Nipponbare seedlings examined at 18°C, all of them had a similar resistant percentage as RonL (Table 1), indicating the presence of a WSMV-resistant gene in rice genome. Surprisingly, all rice plants remained resistance to WSMV at 22°C while RonL did not. Thus, rice might have a different resistance gene or allele than *Wsm2*.

Identifying candidates of *Wsm2* orthologs in *Brachypodium*, rice and wheat

The missing of an effective WSMV resistant gene in 45 natural populations of *Brachypodium* suggests that WSMV resistant genes are not likely generated through single nucleotide polymorphic or short insertion/deletion mutations, which often arise from random natural mutations. Comparative genomic analysis has revealed high inter- and intrachromosomal gene duplication rates in the wheat genome, particularly in chromosome 3B (Choulet et al. 2014; Dubcovsky & Dvorak 2007). This high recombination rate might contribute to the birth of a WSMV resistance gene, especially *Wsm2*, which was gained through a three-way cross hybridization of susceptible parental lines CO850034, PI222668, and TAM107 (Seifers et al. 2006). It is likely that exon shuffling through DNA recombination in the process of breeding gave rise to a new function of an ancestor *Wsm2* locus for WSMV resistance. The discovery of WSMV resistance in Nipponbare inbreed line and the syntenic relationship between rice chromosome 1 and wheat chromosome 3B implied that Nipponbare might encode a *Wsm2* homologous, which was gained through exon shuffling on chromosome 1 during the breeding process. Given the dominant function of *Wsm2* allele (Lu et al. 2011), exon shuffling resulted in

the ancestor *Wsm2* locus to 1) lose an ancestral inhibition domain, 2) acquire an activation domain, or 3) increase expression. Since *Wsm2* was produced only through four generations of segregation, intrachromosomal recombination is more likely to happen than interchromosomal recombination to give the birth of *Wsm2* because the frequency of DNA paring between two separate chromosomes is lower than that within a chromosome.

The large genome size of wheat has limited its genetic manipulation. In order to isolate *Wsm2*, we developed a bioinformatics pipeline to predict the candidates of a *Wsm2* ancestor locus and its orthologs in *Brachypodium* and rice (Figure 1). Since wheat chromosome 3B is syntenic to *Brachypodium* chromosome 2 and rice chromosome 1, amino acid sequences of protein coding genes annotated on these three chromosomes were retrieved from each genome project. In total, 5,070, 7,074, and 7,264 protein sequences were obtained from *Brachypodium* (Bd21-3) (International Brachypodium 2010), rice (Nipponbare) (Kawahara et al. 2013), and wheat 3B (Choulet et al. 2014) genomes, respectively. These sequences were then combined for an OrthoMCL analysis (Li et al. 2003) and 4,035 OrthoMCL groups were resolved as preliminary candidates of *Wsm2* orthologs (Figure 1, File S1). Surprisingly, we did not find any potential rice orthologs of *Wsm2* described in a previous study (Tan et al. 2017), although our list did include all *Brachypodium* genes from the same work, indicating that the previous orthology analysis could be problematic.

Prediction of a *Wsm2* ancestor locus

Wsm2 ancestor locus might acquire a new function for WSMV resistance through deletion of a repression domain, acquisition of an activation domain, or upregulation of expression (Figure 1). Therefore, we reasoned that a putative *Wsm2* ortholog in rice would be 25% shorter or longer in length, or 10 fold higher in expression than its orthologs in *Brachypodium* and wheat. Based on these criteria, we first compared the protein sequence length differences between a rice ortholog and the other members from wheat and *Brachypodium* within the same orthologous group. In total, we found that 67 and 16 out of 4,035 OrthoMCL orthologous groups contain a rice sequence that are 25% shorter or longer, respectively, than those from wheat and *Brachypodium* (Table S1). Therefore, the wheat members from these two groups represent good candidates of a *Wsm2* ancestor locus.

The gain of WSMV resistant function in *Wsm2* could be also attributed to a gene dosage-dependent response. One simple way to increase *Wsm2* transcripts is through fusion of an

ancestor *Wsm2* to a strong promoter via recombination. To identify the possibility of this process, we counted the number of ESTs for each member in each of 4,035 OrthoMCL orthologous groups (File S1). If rice expresses a *Wsm2* ortholog responsible for its WSMV resistance, which is gene dosage-dependent, we reason that the expression of this ortholog would have a significant higher expression than its orthologs in wheat and *Brachypodium*. To find these orthologs, we compared their expression across three species based on normalized EST values (see Materials and Methods). In total, 17,661, 37,590, and 69,162 ESTs were retrieved from the genomes of *Brachypodium*, rice, and wheat, respectively. Through BLASTn search (Altschul et al. 1990), we identified 10 rice genes that have 10 fold more normalized ESTs than their corresponding orthologous members in *Brachypodium* and wheat (Table S1). Taken together, we predicted a total of 93 good candidates for a *Wsm2* ancestor locus.

The *Wsm2* ancestor locus is not likely linked with *Wsm2*

To identify the linkage relationship of a *Wsm2* ancestor locus with *Wsm2*, we retrieved the physical positions (coordinates) of all 93 candidates and plotted them on chromosome 3B (Figure 2). As a control, the positions of eight SNP markers tightly linked with *Wsm2* were also identified (Assanga et al. 2017) (File S2). Although none of our candidate genes are tightly linked with *Wsm2*, we identified 10 genes that reside in the R1 and R3 distal regions (Figure 2), two regions with high recombination rates on wheat chromosome 3B (Choulet et al. 2014). With respect of the overall high recombination rate of chromosome 3B, we cannot rule out the possibility of the remaining 83 candidates to be a *Wsm2* ancestor locus. Certainly, the closer to the centromere region a candidate gene is, the less likely it could be a *Wsm2* ancestor locus.

Putative function of a *Wsm2* candidate gene

Since all 93 candidates could be a *Wsm2* ancestor locus, we further analyzed the functional domains in each protein sequence using HMMER3 (http://hmmer.org) to search against Pfam-A protein-protein interaction database (https://pfam.xfam.org, Version 31). Based on the broad function of each domain, we classified the putative functions of 93 candidates into seven categories, including glycosylation and membrane trafficking, protein ubiquitylation and degradation, transcription factor, chaperone, exonuclease, epigenetic regulation, and unknown (Table S2). Among these, we found that 11 and 17 candidates likely play a role in protein degradation (including ubiquitylation) and membrane trafficking (including glycosylation), which are 5.5 and 8.5 fold more than the third large known functional category (transcription factor),

respectively. Therefore, proteins encoded by *Wsm2* may combat WSMV disease through a molecular mechanism involving protein degradation and/or membrane trafficking.

Discussion

Conditions that may influence the result of WSMV infectivity test

In the WSMV infection tests, it is easy to determine the viral infectivity in wheat through visible disease symptoms (streaks or mosaic) on leaves. However, such symptoms are not easily observed in *Brachypodium* because of its small-sized leaves. Therefore, we used ELISA tests to assist evaluation. In the initial testing, few small *Brachypodium* plants (1~3 plants in each replication) were not infected due to inoculation challenges on narrow and skinny leaves. In the later test of 45 *Brachypodium* accessions, fertilizer was applied to stimulate robust and healthy plant growth, which allowed effective finger-rubbing inoculation on leaves. Not surprisingly, 100% infection rate was detected on all 45 accessions, including Bd21-3 that was used in the initial test. Therefore, healthy and large leaf area is important for evaluating WSMV infectivity in grass species.

Seed purity or temperature fluctuations could also impact the infectivity result due to temperature sensitivity of *Wsm2*. For example, in the initial testing, RonL did not show 100% resistance to WSMV at 18^{o}C (Table 1), which could be explained by the problems of either seed purity or temperature fluctuations of the growth condition. This is not uncommon in WSMV-infectivity test. A similar result was observed for *Wsm2*-containing wheat line CO960293 in previous studies (Lu et al. 2011, Seifers et al. 2013b). In addition, weak plants may cause WSMV infectivity/growth. For example, in the course of our WSMV infectivity tests, we detected GHV greater than 2 (2.1 and 6.5, Table 1, Rep I) in two small rice plants, which indicates WSMV susceptibility (Seifers et al. 2006). However, in other replications, all rice plants examined remained healthy and displayed 100% resistance to WSMV (Table 1). Collectively, our data suggest that replicates and number of individuals are important to give a comprehensive evaluation of WSMV infectivity tests.

The contribution of crop domestication in the birth of *Wsm2*

Brachypodium is evolutionarily close to wheat. However, the former is present naturally and is strictly self-pollinated wild species (Asplund et al. 2000; International Brachypodium 2010) and the later is a crop species that has been domesticated for ~12,000 years (Asplund et al. 2000; Meyer et al. 2012). The domestication process has significantly increased genome

rearrangement and produced contrastive genome architecture of a domestic species comparing to its wild-type progenitor genome (Yue et al. 2017). Both the lack and the presence of a WSMV resistance gene in *Brachypodium* and rice, respectively, suggest that a WSMV-resistant trait is more likely a product of crop domestication. Consistent with this hypothesis, various wild grass species have been shown to be susceptible to WSMV due to the lack of genetic sources (Ito et al. 2012).

Wsm2 is not likely linked with its ancestor locus

Previous functional prediction suggested that a *Wsm2* ancestor locus is linked to *Wsm2* (Tan et al. 2017). However, our broad orthologous group analysis did not identify any rice genes described in Tan et al. (Tan et al. 2017) that share an ortholog in wheat (File S1). In addition, all *Brachypodium* genes discovered in their work have either a similar length or a comparable expression level as the orthologous members in wheat (Table S1), further lowering the possibility of close genetic linkage between *Wsm2* and its ancestor locus. Since our cross species analysis of WSMV resistance suggested that *Wsm2* likely gained the pathogen resistant function through intrachromosomal recombination (Tables 1 to 3), we inferred that a *Wsm2* ancestor is not necessarily linked with *Wsm2*.

Through comparative genome analysis, a previous study has discovered that the wheat *Hardness* (*Ha*) locus was a rejoining product of DNA fragments separated from two different loci via illegitimate recombination (Chantret et al. 2005). Such recombination events could occur in any genomic region, which is not necessary related to transposon-mediate DNA insertion/deletion (Gregory 2004; Kirik et al. 2000). The identification of wide distribution of short conserved sequence motifs at rearrangement breakpoints suggested that illegitimate recombination between unlinked genomic regions is a major evolutionary driving force in wheat domestication (Chantret et al. 2005). Therefore, the unlinkage of our 93 candidates of *Wsm2* ancestor locus with *Wsm2* indicates that the *Wms2* ancestor locus is not necessary within the *Wsm2* locus and that the birth of *Wms2* is likely attributed to illegitimate recombination on chromosome 3B.

The role of protein degradation and membrane trafficking in pathogen defense

The discovery of many *Wsm2* ancestor candidates expressing a domain involved in protein degradation and membrane trafficking is intriguing. Recent genetic, genomic, and proteomic studies have highlighted the role of these two biochemical mechanisms in plant pathogen

defense at various stages, including perception, response, and defense (Duplan & Rivas 2014; Furlan et al. 2012; Furniss & Spoel 2015; Li et al. 2014; Marino et al. 2012). Through genome annotation, we have discovered that the ubiquitin-26S proteasome system (UPS) is extremely large in the wheat genome, in part due to its polyploidy nature (unpublished result). For example, we discovered that wheat genome encodes the largest family of ubiquitin and ubiquitin-like genes in 50 plant genomes (Hua et al. 2018), further indicating the importance of the UPS in regulating wheat development and growth. In addition, the role of protein ubiquitylation is also intimately connected with membrane trafficking in cells (Clague & Urbe 2017). Therefore, our study implies a putative role of *Wsm2* in ubiquitylation and/or membrane trafficking-mediated protein degradation.

Conclusions

In this study, we compared WSMV resistance among three closely related grass species (Tables 1-3) and developed a novel bioinformatics pipeline for predicting potential candidates of a *Wsm2* ancestor locus (Figure 1). Given that *Wsm2* is likely duplicated through an intrachromosomal illegitimate recombination and the dominant phenotype of *Wsm2*, we inferred that this new WSMV-resistant gene acquired an activation domain, lose an inhibition domain, or gain high expression compared to its ancestor locus (Table S1). The resulting 93 putative *Wsm2* ancestor loci could serve as good candidates for future genetic isolation of the true *Wsm2* locus. We may design new polymerase chain reaction (PCR) primers based on the nucleotide sequences of each candidate to examine the presence of an additional copy that is linked to the *Wsm2* locus. The finding of such a copy could serve a starting point to clone the full-length of a putative *Wsm2* gene via thermal asymmetric interlaced PCR (Liu et al. 1995). This may provide an efficient way to isolate and characterize the molecular function of *Wsm2*.

Acknowledgments

Contribution number 18-373-J from the Kansas Agricultural Experiment Station. This work was supported by the USDA National Institute of Food and Agriculture competitive grant (2017-67007-25939 to G.Z.) and Hatch grant (1001453 to G.Z.), and Ohio University Start-Up grant (SU1007172 to Z.H.). We thank Daniel P. Woods (University of Wisconsin-Madison) for providing the seeds of 45 *Brachypodium* accessions and Jeff Ackerman (Kansas State University) for valuable assistance in virus testing.

Supplemental Information

433 **File S1.** List of 4,035 orthologous groups encoded in *Brachypodium* chromosome 2, rice
434 chromosome 1, and wheat Chromosom 3B
435 **File S2.** Nucleotide sequences of 8 SNP markers tightly linked with *Wsm2*
436 **Table S1.** Sequence and expression comparison of 93 *Wsm2* ancestor candidates with
437 *Brachypodium* and rice orthologous members
438 **Table S2.** Putative biochemical functions of 93 *Wsm2* ancestor candidates
439 **Table S3.** Experimental raw data for WSMV infectivity tests on Nipponbare and Bd21-3
440 **Table S4.** Experimental raw data for WSMV infectivity tests on 45 *Brachypodium* natural
441 variants.
442 **Table S5.** Experimental raw data for back assay on Bd21-3
443
444

Figure Legends

Figure 1. A diagram showing the analysis procedures and summary of results in identifying *Wsm2* ancestor candidates.

Figure 2. Physical relationship of 93 candidates of *Wsm2* ancestor loci with *Wsm2* locus on wheat chromosome 3B. A) Positions of 16 candidates that may acquire an activation domain to become *Wsm2*. B) Locations of 67 candidates that may form *Wsm2* by deletion of an inhibition domain. C) Distribution of 10 candidates that may be changed as *Wsm2* through fusion with a strong promoter in the *Wsm2* locus. D) The place of the Wsm2 locus on chromosome 3B that is represented by its tightly linked 8 SNP markers. E) Schematic representation of the structure of chromosome 3B adopted from Choulet et al. (Choulet et al. 2014). Circle dots: 93 candidates of *Wsm2* ancestor loci. Black diamonds: 8 SNP markers that are tightly linked with *Wsm2*. Red diamonds: beginning and end of chromosome 3B. R1, R2, and R3: three regions of chromosome 3B with different recombination rates. C: centromeric/ pericentromeric region. Red/yellow shaded regions: two distal regions (R1 and R3) with high recombination rates. +/-: Watson/Crick DNA strands of chromosome 3B.

References

Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403-410. 10.1016/S0022-2836(05)80360-2

Appel JA, De Wolf E, Bockus WW, and Bowden RL. 2013. Preliminary 2013 Kansas wheat disease loss estimates. *Kansas Cooperative Plant Disease Survey Report*.

Asplund L, Hagenblad J, and Leino MW. 2000. Re-evaluating the history of the wheat domestication gene NAM-B1 using historical plant material. *J Archaeological Science* 37:2303-2307.

Assanga S, Zhang G, Tanc CT, Ruddc JC, Ibrahima A, Xue Q, Chao S, Fuentealbac MP, and Liu S. 2017. Saturated genetic mapping of wheat streak mosaic virus resistance gene Wsm2 in wheat. *Crop Sci* 57:332-339. 10.2135/cropsci2016.04.0233

Baleya GJ, Talbertb LE, Martinb JM, Youngb MJ, Habernichtb DK, Kushnakc GD, Bergb JE, Lanningb SP, and Bruckner PL. 2001. Agronomic and end-use qualities of wheat streak mosaic virus resistant spring wheat. *Crop Sci* 41:1779-1784. 10.2135/cropsci2001.1779

Chantret N, Salse J, Sabot F, Rahman S, Bellec A, Laubin B, Dubois I, Dossat C, Sourdille P, Joudrier P, Gautier MF, Cattolico L, Beckert M, Aubourg S, Weissenbach J, Caboche M, Bernard M, Leroy P, and Chalhoub B. 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (Triticum and Aegilops). *Plant Cell* 17:1033-1045. 10.1105/tpc.104.029181

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Le Gouis J, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury JM, Mayer K, Berges H, Quesneville H, Wincker P, and Feuillet C. 2014. Structural and functional partitioning of bread wheat chromosome 3B. *Science* 345:1249721. 10.1126/science.1249721

Christian ML, and Willis WG. 1993. Survival of wheat streak mosaic virus in grass hosts in Kansas from wheat harvest to fall wheat emergence. *Plant Dis* 77:239-242. 10.1094/PD-77-0239

Clague MJ, and Urbe S. 2017. Integration of cellular ubiquitin and membrane traffic systems: focus on deubiquitylases. *FEBS J* 284:1753-1766. 10.1111/febs.14007

Dubcovsky J, and Dvorak J. 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316:1862-1866. 10.1126/science.1143986

Duplan V, and Rivas S. 2014. E3 ubiquitin-ligases and their target proteins during the regulation of plant innate immunity. *Front Plant Sci* 5:42. 10.3389/fpls.2014.00042

Eddy SR. 2011. Accelerated Profile HMM Searches. *PLoS Comput Biol* 7:e1002195. 10.1371/journal.pcbi.1002195

Fahim M, Mechanicos A, Ayala-Navarrete L, Haber S, and Larkin PJ. 2011. Resistance to *Wheat streak mosaic virus* – a survey of resources and development of molecular markers. *Plant Pathol* 61:425-440. 10.1111/j.1365-3059.2011.02542.x

Furlan G, Klinkenberg J, and Trujillo M. 2012. Regulation of plant immune receptors by ubiquitination. *Front Plant Sci* 3:238. 10.3389/fpls.2012.00238

Furniss JJ, and Spoel SH. 2015. Cullin-RING ubiquitin ligases in salicylic acid-mediated plant immune signaling. *Front Plant Sci* 6:154. 10.3389/fpls.2015.00154

Gill BS, Friebe B, Qi LL, Wilson DL, Raupp WJ, Fritz AK, Seifers DL, Martin TJ, and Pumphrey MO. 2008. Notice of release of KS08WGGRC50 wheat streak mosaic virus- and Triticum mosaic virus-resistant hard red winter wheat germ plasm. *Annual Wheat Newsletter* 54:170.

Gill BS, Friebe B, Wilson DL, and Cox TS. 1995. Registration of KS93WGRC27 wheat streak mosaic virus resistant T4DL.4Ai 2S wheat germplasm. *Crop Sci* 35:1236.

Gordon SP, Contreras-Moreira B, Woods DP, Des Marais DL, Burgess D, Shu S, Stritt C, Roulin AC, Schackwitz W, Tyler L, Martin J, Lipzen A, Dochy N, Phillips J, Barry K, Geuten K, Budak H, Juenger TE, Amasino R, Caicedo AL, Goodstein D, Davidson P, Mur LAJ, Figueroa M, Freeling M, Catalan P, and Vogel JP. 2017. Extensive gene content variation in the Brachypodium distachyon pan-genome correlates with population structure. *Nat Commun* 8:2184. 10.1038/s41467-017-02292-8

Gregory TR. 2004. Insertion-deletion biases and the evolution of genome size. *Gene* 324:15-34. 10.1016/j.gene.2003.09.030

Haley SD, Johnson JJ, Peairs FB, Stromberger JA, Heaton EE, Seifert SA, A. KR, Rudolph JB, Bai G, Bowden RL, Chen M-S, Chen X, Jin Y, Kolmer JA, Chen R, and eabourn BW. 2011. Registration of 'Snowmass' wheat. *J Plant Reg* 5. 10.3198/jpr2010.03.0175crc

Haley SD, Martin TJ, Quick JS, Seifers DL, Stromberger JA, Clayshulte SR, Clifford BL, Peairs FB, Rudolph JB, Johnson JJ, Gill BS, and Friebe B. 2002. Registration of CO960293-2 wheat germplasm resistant to wheat streak mosaic virus and Russian wheat aphid. *Crop Sci* 42:1381–1382.

Hua Z, Doroodian P, and Vu W. 2018. Multiple gene duplications contributed unevenly to the expansion of ubiquitin and ubiquitin-like protein modifiers in plants. *Plant J* (resubmission).

530 Hua Z, Zou C, Shiu SH, and Vierstra RD. 2011. Phylogenetic comparison of F-Box (FBX) gene
531 superfamily within the plant kingdom reveals divergent evolutionary histories indicative of
532 genomic drift. *PLoS One* 6:e16219. 10.1371/journal.pone.0016219
533 International Brachypodium I. 2010. Genome sequencing and analysis of the model grass
534 Brachypodium distachyon. *Nature* 463:763-768. 10.1038/nature08747
535 Ito D, Miller Z, Menalled F, Moffet M, and Burrows M. 2012. Relative susceptibility among
536 alternative host species prevalent in the great plains to wheat streak mosaic virus. *Plant*
537 *Dis* 96:1185-1192. 10.1094/PDIS-09-11-0746-RE
538 Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz
539 DC, Tanaka T, Wu J, Zhou S, Childs KL, Davidson RM, Lin H, Quesada-Ocampo L,
540 Vaillancourt B, Sakai H, Lee SS, Kim J, Numa H, Itoh T, Buell CR, and Matsumoto T.
541 2013. Improvement of the Oryza sativa Nipponbare reference genome using next
542 generation sequence and optical map data. *Rice* 6:4. 10.1186/1939-8433-6-4
543 Kirik A, Salomon S, and Puchta H. 2000. Species-specific double-strand break repair and
544 genome evolution in plants. *EMBO J* 19:5562-5566. 10.1093/emboj/19.20.5562
545 Li B, Lu D, and Shan L. 2014. Ubiquitination of pattern recognition receptors in plant innate
546 immunity. *Mol Plant Pathol* 15:737-746.
547 Li L, Stoeckert CJ, Jr., and Roos DS. 2003. OrthoMCL: identification of ortholog groups for
548 eukaryotic genomes. *Genome Research* 13:2178-2189. 10.1101/gr.1224503
549 Liu YG, Mitsukawa N, Oosumi T, and Whittier RF. 1995. Efficient isolation and mapping of
550 *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR.
551 *Plant J* 8:457-463. 10.1046/j.1365-313X.1995.08030457.x
552 Lu H, Price J, Devkota R, Rush C, and Rudd J. 2011. A dominant gene for resistance to wheat
553 streak mosaic virus in winter wheat line CO960293-2. *Crop Sci* 51:5-12.
554 10.2135/cropsci2010.01.0038
555 Marino D, Peeters N, and Rivas S. 2012. Ubiquitination during plant immune signaling. *Plant*
556 *Physiol* 160:15-27. 10.1104/pp.112.199281
557 Martin TJ, Fritz AK, Seifers D, and Shroyer JP. 2007. 'RonL' hard white wheat. *Agricultural*
558 *Experiment Station Publications, Kansas State University Manhattan, KS*.
559 Martin TJ, Zhang G, Fritz AK, Miller R, and Chen M. 2014. Registration of ‘Clara CL’ Wheat. *J*
560 *Plant Reg* 81:38-42. 10.3198/jpr2013.07.0040crc
561 McNeil JE, French R, Hein GL, Baenziger PS, and Eskridge KM. 1996. Characterization of
562 genetic variability among natural populations of wheat streak mosaic virus.
563 *Phytopathology* 86:1222-1227.

564 Meyer RS, DuVal AE, and Jensen HR. 2012. Patterns and processes in crop domestication: an
565 historical review and quantitative analysis of 203 global food crops. *New Phytol* 196:29-
566 48. 10.1111/j.1469-8137.2012.04253.x
567 Murray GM, and Brennan JP. 2009. Estimating disease losses to the Australian wheat industry.
568 *Australasian Plant Pathol* 38:558-570. 10.1071/AP09053
569 Navia D, Mendonça RSd, Skoracka A, Szydło W, Knihinicki D, Hein GL, Pereira PRVdS, Truol
570 G, and Lau D. 2013. Wheat curl mite, Aceria tosichella, and transmitted viruses: an
571 expanding pest complex affecting cereal crops. *Experimental and Applied Acarology*
572 59:95-143.
573 Price JA, Workneh F, Evett SR, Jones DC, Arthur J, and Rush CM. 2010. Effects of wheat
574 streak mosaic virus on root development and water-use efficiency of hard red winter
575 wheat. *Plant Dis* 94:766-770. 10.1094/PDIS-94-6-0766
576 Seifers DL, Haber S, Martin TJ, and Zhang G. 2013a. New sources of temperature-sensitive
577 resistance to Wheat streak mosaic virus in wheat. *Plant Dis* 97:1051-1056.
578 10.1094/PDIS-11-12-1029-RE
579 Seifers DL, Haber S, Martin TJ, and Zhang G. 2013b. Temperature-sensitive resistance to
580 wheat streak mosaic virus in CO960333 and KS06HW79 wheat. *Plant Dis* 97:983-987.
581 10.1094/PDIS-10-12-0971-RE
582 Seifers DL, Martin TJ, Haber S, and Haley SD. 2006. Temperature sensitivity and efficacy of
583 wheat streak mosaic virus resistance derived from CO960293 wheat. *Plant Dis* 90:623-
584 628. 10.1094/PD-90-0623
585 Seifers DL, Martin TJ, Harvey TL, and Gill BS. 1995. Temperature sensitivity and efficacy of
586 wheat streak mosaic virus resistance derived from *Agropyron intermedium*. *Plant Dis*
587 79:1104-1106. 10.1094/PD-79-1104
588 Sharp GL, Martin JM, Lanning SP, Blake NK, Brey CW, Sivamani E, Qu R, and Talbert LE.
589 2002. Field evaluation of transgenic and classical sources of wheat streak mosaic virus
590 resistance. *Crop Sci* 43:105-110. 10.2135/cropsci2002.1050
591 Sim T, Willis WG, and Eversmeyer MG. 1988. Kansas plant disease survey. *Plant Dis* 72:832-
592 836. 10.1094/PD-72-0832
593 Tan CT, Assanga S, Zhang G, Rudd JC, Haley SD, Xue Q, Ibrahim A, Bai G, Zhang X, Byrne P,
594 Fuentealba MP, and Liu S. 2017. Development and validation of KASP markers for
595 wheat streak mosaic virus resistance gene *Wsm2*. *Crop Sci* 57:340-348.
596 10.2135/cropsci2016.04.0234

Triebe B, Mukai Y, Dhaliwal HS, Martin TJ, and Gill BS. 1991. Identification of alien chromatin specifying resistance to wheat streak mosaic and greenbug in wheat germ plasm by C-banding and in situ hybridization. *Theor Appl Genet* 81:381-389. 10.1007/BF00228680

Wosula EN, Tatineni S, Wegulo SN, and Hein GL. 2017. Effect of temperature on wheat streak mosaic disease development in winter wheat. *Plant Dis* 101:324-330. 10.1094/PDIS-07-16-1053-RE

Yue JX, Li J, Aigrain L, Hallin J, Persson K, Oliver K, Bergstrom A, Coupland P, Warringer J, Lagomarsino MC, Fischer G, Durbin R, and Liti G. 2017. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat Genet* 49:913-924. 10.1038/ng.3847

Zhang G, Martin TJ, Fritz AK, Miller R, Chen M, Bowden RL, and Johnson J. 2015a. Registration of 'Oakley CL' wheat. *J Plant Reg* 9:190-195. 10.3198/jpr2014.04.0023crc

Zhang G, Martin TJ, Fritz AK, Miller R, Chen MS, Bowden RL, and Bai G. 2016a. Registration of 'Joe' hard white winter wheat. *J Plant Reg* 10:283-286. 10.3198/jpr2016.02.0007crc

Zhang G, Seifers DL, and Martin TJ. 2015b. Inheritance of wheat streak mosaic virus resistance in KS03HW12. *Austin J Plant Biol* 1:1001.

Zhang X, Bai G, Xu R, and Zhang G. 2016b. Wheat streak mosaic virus resistance in eight wheat germplasm lines. *Plant Breeding* 135:26-30. 10.1111/pbr.12334

Table 1(on next page)

Resistant percentages (%) of rice cultivar Nipponbare, *Brachypodium* accession Bd21-3, wheat cultivars RonL and Tomahawk after inoculation with WSMV isolate Sidney 81.

Table 1. Resistant percentages (%) of rice cultivar Nipponbare, *Brachypodium* accession Bd21-3, wheat cultivars RonL and Tomahawk after inoculation with WSMV isolate Sidney 81.

Temperature	Genotype	Rep I[a]	Rep II	Mean[b]
18°C	Nipponbare (rice)	83.3 (12)	100.0 (8)	91.7A
	Bd21-3 (*Brachypodium*)	20.0 (10)	14.3 (7)	17.2B
	RonL (Wheat)	90.0 (10)	90.9 (11)	90.5A
	Tomahawk (wheat)	0.0 (11)	0.0 (11)	0.0B
	LSD (0.01)			39.7
22°C	Nipponbare (rice)	100.0 (2)	100.0 (1)	100.0A
	Bd21-3 (*Brachypodium*)	12.5 (8)	30.0 (10)	21.3B
	RonL (wheat)	0.0 (12)	0.0 (11)	0.0B
	Tomahawk (wheat)	0.0 (10)	0.0 (10)	0.0B
	LSD (0.01)			36.1

[a] Number in the parenthesis indicates the size of sample in each replication

[b]Genotypes not having the same letter in common are significantly different at $p < 0.01$.

Table 2(on next page)

Infectivity variation of WSMV on 45 *Brachypodium* natural variants

Table 2. Infectivity variation of WSMV on 45 *Brachypodium* natural variants

Accessions	Total plants	Infected plants	Infection rate (%)	ELISA[a]	GHV[b]
ABRS	7	7	100	0.54±0.09	78.9±12.6
ABR4	7	7	100	0.55±0.05	80.4±7.5
Adi-2	4	4	100	0.60±0.10	88.6±14.0
Adi-10	5	5	100	0.49±0.07	71.6±10.6
ARN1	9	9	100	0.53±0.08	78.0±11.5
Bd1-1	5	5	100	0.57±0.05	83.6±6.9
Bd2-3	5	5	100	0.50±0.06	73.1±9.2
Bd3-1	5	5	100	0.54±0.05	79.1±7.4
Bd21-1	7	7	100	0.52±0.23	76.8±33.9
Bd21-3	7	7	100	0.59±0.08	87.2±11.5
Bd29-1	9	9	100	0.40±0.17	58.9±24.3
Bd30-1	7	7	100	0.50±0.03	73.2±4.7
Bis1	5	5	100	0.58±0.04	85.5±6.4
Foz1	3	3	100	0.50±0.03	74.1±3.7
Gaz8	5	5	100	0.51±0.04	75.6±6.3
Kah-1	6	6	100	0.48±0.03	71.0±4.2
Kah-S	6	6	100	0.47±0.08	69.2±11.6
Koz1	8	8	100	0.48±0.06	70.8±8.5
Koz3	7	7	100	0.50±0.11	73.4±15.6
Luc1	7	7	100	0.48±0.10	71.2±14.7
Mig3	9	9	100	0.49±0.04	71.3±6.5
Mon3	5	5	100	0.76±0.12	112.1±17.8
Mur1	7	7	100	0.43±0.21	63.5±30.2
Per1	7	7	100	0.53±0.05	77.8±7.0
RON2	8	8	100	0.61±0.06	89.6±8.5
Sig2	7	7	100	0.60±0.08	87.8±11.7
TEK-2	7	7	100	0.45±0.05	66.4±6.6
TEK-4	8	8	100	0.64±0.07	94.3±9.8
TEK11	7	7	100	0.59±0.07	86.2±10.7
TR2B	7	7	100	0.55±0.03	81.2±4.9
TR3C	8	8	100	0.61±0.04	89.9±6.5
TR7a	5	5	100	0.42±0.24	62.4±34.6
TR8i	9	9	100	0.46±0.17	66.9±24.5
TR9K	7	7	100	0.44±0.28	65.2±41.3
TR10C	6	6	100	0.60±0.08	87.6±11.4
TR11A	8	8	100	0.55±0.07	80.4±10.5
TR11G	8	8	100	0.51±0.05	74.8±7.7
TR12C	5	5	100	0.38±0.20	56.4±29.6
TRBa	6	6	100	0.58±0.12	85.5±17.4
TR13C	8	8	100	0.63±0.10	91.2±14.9
TR26	6	6	100	0.53±0.08	77.9±12.3
TRIi	9	9	100	0.58±0.05	85.8±7.5
TRSi	7	7	100	0.61±0.05	90.1±8.0
Uni2	8	8	100	0.68±0.12	100.5±17.7
18-1	6	6	100	0.57±0.04	83.1±6.3
Healthy control	5	0	0	0.0068±0.008	

[a]ELISA, enzyme linked immunosorbent assay, mean absorbance ± SD

[b]GHV = Sample ELISA value / Healthy ELISA value, mean ± SD

Table 3(on next page)

Back assay of WSMV infection on Bd21-3 with a series of inoculum dilutions

Table 3. Back assay of WSMV infection on Bd21-3 with a series of inoculum dilutions

Inoculum dilution rate (wt/vol)	Total plants	Infected plants	Infection rate (%)	ELISA[a]	GHV[b]
1:5	9	8	88.9	0.29±0.11	26.1±10.4
1:10	8	8	100.0	0.33±0.06	29.7±5.2
1:20	8	8	100.0	0.29±0.04	26.4±3.6
1:40	6	5	83.3	0.24±0.13	22.0±12.0
1:80	7	7	100.0	0.27±0.05	24.4±5.0
1:160	8	4	50.0	0.13+0.13	11.6+11.9
Healthy control	5			0.01±0.01	1.0±0.53

[a]ELISA, enzyme linked immunosorbent assay, mean absorbance ± SD

[b]GHV = Sample ELISA value / Healthy ELISA value, mean ± SD

Figure 1

A diagram showing the analysis procedures and summary of results in identifying *Wsm2* ancestor candidates.

Figure 1; Zhang and Hua, 2018

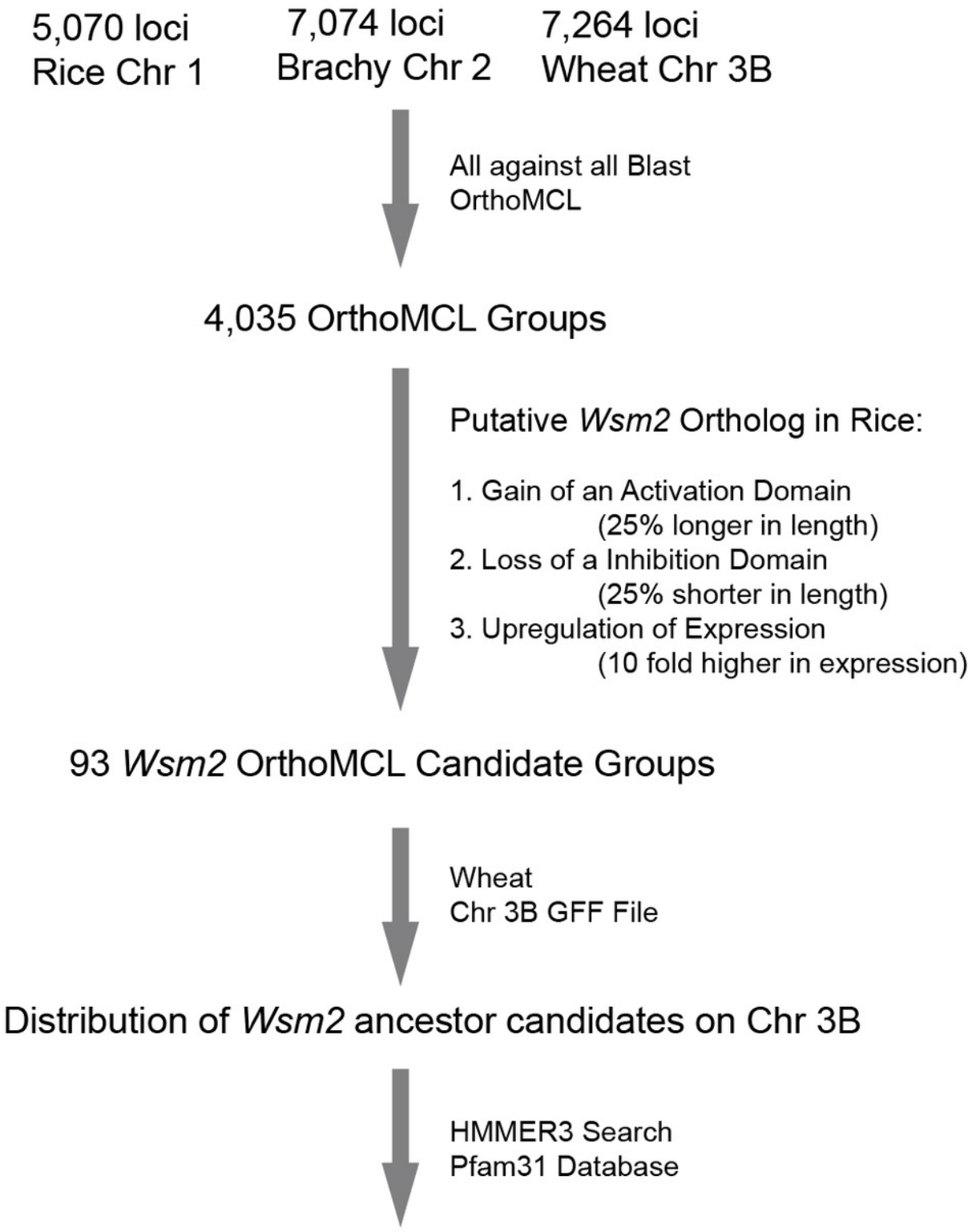

Figure 2

Physical relationship of 93 candidates of *Wsm2* ancestor loci with *Wsm2* locus on wheat chromosome 3B.

A) Positions of 16 candidates that may acquire an activation domain to become *Wsm2*. B) Locations of 67 candidates that may form *Wsm2* by deletion of an inhibition domain. C) Distribution of 10 candidates that may be changed as *Wsm2* through fusion with a strong promoter in the *Wsm2* locus. D) The place of the Wsm2 locus on chromosome 3B that is represented by its tightly linked 8 SNP markers. E) Schematic representation of the structure of chromosome 3B adopted from Choulet et al. (Choulet et al. 2014) . Circle dots: 93 candidates of *Wsm2* ancestor loci. Black diamonds: 8 SNP markers that are tightly linked with *Wsm2*. Red diamonds: beginning and end of chromosome 3B. R1, R2, and R3: three regions of chromosome 3B with different recombination rates. C: centromeric/ pericentromeric region. Red/yellow shaded regions: two distal regions (R1 and R3) with high recombination rates. +/-: Watson/Crick DNA strands of chromosome 3B.

Figure 2; Zhang and Hua, 2018