SUPPORTING INFORMATION

FRIZZY PANICLE drives supernumerary spikelets in bread wheat (*T. aestivum* L.)

Oxana Dobrovolskaya, Caroline Pont, Richard Sibout, Petr Martinek, Ekaterina Badaeva, Florent Murat, Audrey Chosson, Nobuyoshi Watanabe, Elisa Prat, Nadine Gautier, Véronique Gautier, Charles Poncet, Yuriy Orlov, Alexander A. Krasnikov, Hélène Bergès, Elena Salina, Lyudmila Laikova, Jerome Salse.

Supplemental Information.

Plant materials - Seventeen bread wheat lines with the SS phenotype were used in this study. Of these lines, 13 are genetically related and present the MRS type of SS. All of these lines are fertile and produce seeds.

The MRS lines. The Rŭc163-1-02 (designated Ruc163 throughout the paper) and Rŭc167-1-02 (designated Ruc167 throughout the paper) lines were developed by Martinek and Bednar (2001) and were previously used for the genetic mapping of the *mrs1* gene, which determines the MRS trait (Dobrovolskaya *et al.,* 2009). Ten MRS lines, V3-79-08, V3-82-08, V3-83-08, V3-84-08, V3-85-08, V3-86-08, V3-87-08, V3-88-08, V3-89-08, and V3-90-08, carry the *mrs1* gene in different genetic backgrounds. These lines were characterized using SSR markers that are closely linked to the *mrs1* gene (Dobrovolskaya *et al.,* 2008). The near-isogenic line NIL-*mrs1* was recently developed by Prof. N. Watanabe (College of Agriculture, Ibaraki University, Ibaraki, Japan). The donor of the MRS trait, KM240, is related to the other MRS lines used in this work; the recurrent parent was the Russian spring bread wheat variety Novosibirskaya 67 (N67).

The other four lines are genetically unrelated. The lines are spontaneous or induced mutants.

The Skle128-09 line. The Skle128-09 (Skle128 throughout the paper) line is derived from the Tibetan triple-spikelet wheat (TTSW) (*Triticum aestivum* L. concv. *tripletum* nom. nud.), a landrace of common wheat collected from Tibet, China. The donor of this material was Dr. Bao-Rong Lu (Fudan University, Shanghai). The line is characterized by genetically stable supernumerary spikelet trait; most of the supernumerary spikelets are presented by triple spikelets. The trait is under the control of two recessive genes (Yang *et al.,* 2005); one of the genes, *qTS2A-1*, was genetically mapped to chromosome 2AS (Li *et al.,* 2011). The SS phenotype of this line is determined by a spontaneous mutation(s).

The MC1611 line. The MC1611 line is a mutant produced through chemical mutagenesis and provided by Dr. V.M. Melnick (Altai Research Institute of Agriculture, Siberian Branch of Russian Academy of Agricultural Sciences, Barnaul, Russia). Seeds of Russian spring variety Saratovskaya 29 (S29) were treated with nitrosomethylurea (NMU). The SS trait of this line is affected by a gene localized to chromosome 2D through monosomic analysis (Laikova *et al.,* 2005). The SS trait of MC1611 is weaker than other SS lines used in our study. Under greenhouse conditions, the additional spikelets developed only at several (1-3) nodes at the lower part of the spike (Figure S1).

The So164-02 line. The So164-02 (So164 throughout the paper) line is characterized by the SS trait of *horizontal spikelets* (HS) (syn. '*tetrastichon* sessile spikelets'), occurring when two or three spikelets are in a horizontal position at a spike rachis node. The initial SS form was obtained by Dr. Svetka Korić (Agriculture University of Zagreb, Zagreb, Croatia) from the cross of a branched tetraploid with a common bread wheat line. The donor of this line (accession N 7344-85) was Prof. Slavko Borojević (University of Novi Sad, Novi Sad, Serbia).

The Rŭc204-08 line. The Rŭc204-08 line (Ruc204 throughout the paper) is characterized by genuine branching (GB) (syn. '*turgidum* type of branching) under field conditions in Kroměříž, Czech Republic (2008) and Novosibirsk, Russia (2009, 2010). In addition to the development of supernumerary spikelets,

lateral branches bearing spikelets are formed in the basal part of the spike. Under greenhouse conditions, this line exhibits a weaker mutant phenotype; additional spikelets only develop at a rachis node (Figure S1). This line originated from China under the name '47hh-C', but the details of its origin are unknown.

Light microscopy and SEM analysis - The developing spikes from mutant lines, MC1611, NIL-*mrs1*, and normal spiked cultivars, Saratovskaya 29 (S29), Novosibirskaya 67 (N67), Chinese Spring were dissected with a scalpel under a binocular microscope (Altami PS0745, "Altami", St. Petersburg, Russia). The analysis of inflorescence structure of MC1611, NIL-*mrs1*, S29, and N67 was performed using the Carl Zeiss Stereo Discovery V12 (Carl Zeiss Microscopy GmbH, Germany) light microscope. The images were captured using the AxioCam MRc-5 (Carl Zeiss Microscopy GmbH, Germany). Scanning electron microscope, SEM (TM-1000, Hitachi Co. Hitachi,Ltd, Japan), was used to observe the morphological features of the inflorescences of NIL-*mrs1* and N67. Dissected young spikes were examined under low vacuum conditions (30-50Pa) and an accelerating voltage of 15kV.

C-banding procedure - C-banding analysis was carried out as described earlier (Badaeva *et al.,* 1994). The slides were analysed on the Leitz Wetzlar microscope, and the selected metaphase plates were captured using CCD Leica DFC 280. Chromosomes were classified according to standard genetic nomenclature (Gill *et al.,* 1991).

Primers - All primers developed and used in this study are listed in Table S3. Information on the SSR markers used here is available at http://www.graingenes.org.

Mapping procedure - The mapping populations were genotyped using SSR markers. PCR reactions were performed as described in Nicot et al., 2004. The amplification products were detected using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, Calif.). The ABI GeneScan, version 2.1, and Genotyper, version 2.0, software programs (Applied Biosystems, Foster City, CA, USA) were used to determine the size of the fragments based on an internal lane standard. Genetic maps were constructed using MAP-MAKER/EXP v3.0b software (Lander *et al.,* 1987) with the Kosambi (Kosambi et al., 1943) mapping function and a LOD threshold of 3.00. The QTL mapping of the trait was performed using MapQTL ® 5 software (Ooijen 2004).

COS-SSCP analysis - The COS-SSCP analysis was performed as described in Quraishi *et al.,* (2009) and Pont *et al.,* (2013).

BAC clone screening, sequencing and annotation - A BAC library from *T. aestivum* cv. Chinese Spring, "Tae-B-Chinese spring", developed at the INRA – CNRGV (http://cnrgv.toulouse.inra.fr/) was used in this study. The library has a 9.0x genome coverage and consists of 1147776 clones organized in 2988 384-well plates. The screening was performed as previously described (Dibari *et al.,* 2012). BAC sequencing was performed on the GS Junior (Roche) using the 454 sequencing technology according to the procedure described by the manufacturer at the genotyping platform GENTYANE (UMR-1095 GDEC, Clermont-Ferrand). Genes and repeated elements (TEs) were identified by computing and integrating the results based on BLAST algorithms (Altschul *et al.,* 1990), predictor programs and the software described below. The gene structure and putative functions were identified by combining the results of BLASTN and BLASTX alignments against dbEST (NCBI; http://www.ncbi.nlm.nih.gov/) and

SwissProt databases (Expasy; http://expasy.org/sprot/) with the results of the predictor program, FgeneSH (Salamov and Solovyev, 2000), with default parameters. Known genes were named based on the BLASTX results of proteins with known functions (SwissProt). The characterization of the protein functional domain was performed using InterPro (Hunter *et al.,* 2009) and SMART database (Schultz *et al.,* 1998; Letunic *et al.,* 2009), while protein motifs were analyzed using ScanProsite (De Castro *et al.,* 2006). ChloroP1.1 and ProtComp 9.0 were used to identify signal peptides and to predict protein subcellular localization, respectively. Transposable elements (TEs) were detected by a comparison with two databases of repetitive elements: TREP (Graingenes: A database for Triticeae and Avena; http://wheat.pw.usda.gov./ITMI/Repeats/) and Repbase (Jurka 2000). The boundaries of TEs were identified with the REPET package (Flutre *et al.,* 2011). The insertion profile of the TEs was identified using a modified version of svg_ltr.pl script (Kronmiller and Wise 2008).

Resequencing - Sequences of the three *WFZP* orthologous genes (~3000 bp each, including the coding part of the genes, promoter regions and 3'-regions) of seven SS mutant lines and four normal spiked bread wheat cultivars were obtained by direct Sanger sequencing of PCR products using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems), following the manufacturer's instructions. Fluorescently terminated extension products were separated using a capillary ABI-3730 Bioanalyzer (Applied Biosystems).

RT-PCR - Total RNA was extracted from the developing spikes, leaves and roots of cv. Chinese Spring (Cs). The developing spikes were dissected with a scalpel under a binocular microscope (Altami PS0745, "Altami", St. Petersburg, Russia). Five to eight isolated young spikes per sample were pooled. Total RNA was also isolated from the vegetative organs, young leaves and roots of Cs at the floret differentiation stage of spike development. RNA extractions were performed on three independent biological replicates according to the ZR Plant RNA MiniPrep^{TM} (Zymoresearch) protocol; DNase set (Qiagen) was used for RNA purification. If it was necessary, the residual DNA in the RNA samples was removed using DNase set (Qiagen) and RNeasy Minelute Cleanup (Qiagen) kits. cDNA synthesis was performed with a Transcriptor first strand cDNA synthesis kit (Roche) following the procedure described by the manufacturer with 0.5 µg of RNA in 20 µl of reaction. The thermal cycling conditions were 10′ at 25°C/ 30′ at 55°C/ 5′ at 85°C. The reverse transcription reaction was diluted to a final volume of 200 µl, and 4 µl of synthesized cDNA was used as a template for real-time PCR using the LightCycler® 480 (Roche Diagnostics). The reactions were performed in 10 µl containing 1 x LightCycler® 480 DNA SYBR Green I Master (Roche Diagnostics). Three homoeologous gene-specific primer pairs used in quantitative Real-Time PCR analysis are listed in Table S3. All samples were prepared in three technical replicates, and a negative control using water as a template was included. The quantification cycles (Cq) were analyzed using the LightCycler®480 software version 1.5.0 and normalized with a reference gene Ta.304.3 (primers CCTCTGAGAGCCTCCACGTC; TGGTAAGCAGGACACCCAAAC) using an 'Advanced Relative Quantification' module to obtain a normalized ratio E_t^{-CqR}/E_r^{-CqR} (with CqT/CqR: cycle number at target/reference detection threshold (crossing point) and E_t/E_r: efficiency of target/reference amplification (10^{-1}/slope)). The specificity of the amplification was confirmed via

melting curve analysis of the final PCR products by increasing the temperature from 65°C to 95°C. The PCR efficiency was calculated for each gene using a standard curve of serial dilutions and was used in the relative expression analysis. All observations were expressed as means $\pm S_D$.

***Brachypodium* mutant identification and sequencing.** The *Brachypodium* mutant library available at http://www-ijpb.versailles.inra.fr/en/plateformes/crb/index.html was PCR screened with *FZP* primers. PCR products were sequenced and mutants phenotyped for spike architecture. 500 families (3000 plants with 6 plants per family) were grown in greenhouse and mature plants were screened for phenotypes in spikes development. Two independant families (Bd8202 and Bd8972) were identified with a clear supernumerary spikelet phenotype. In order to verify the co-segregation of the mutated alleles and phenotypes, 12 M2 seeds from each family were grown and genotyped. All plants with the homozygous mutation allele showed the supernumerary spikelet phenotype as well as their own progenies (2 generation were tested).

***qSS* mapping -** To evaluate the genetic contributions of the 2DS deletions in Ruc204 and Skle128 to their SS traits, interval mapping was performed on F_2 mapping populations from the Ruc204xS29 (149 F_2s) and Skle128 x S29 (134 F_2s) crosses. The F_1 progenies of both crosses had spikes of the conventional type, and the F_2s segregated into several phenotypic classes.

The Skle128xS29 cross. In addition to the Skle128 and S29 parental phenotypic classes, two intermediate SS classes were identified (Figure S6a). The ratio of SS segregants of the Skle128 type (10 plants) to the other F_2 plants (124 plants) conformed to the 1:15 segregation ratio (χ^2=0.34, 0.75<P<0.5), suggesting that the SS trait of Skle128 is under the control of two recessive genes. One of the genes was recently mapped as a QTL on 2AS (Li et al., 2011), and here, we demonstrated that the region of 2DS carrying a terminal deletion (Figure S4) explained 49.6% of the phenotypic variance (LOD = 12.4) (Figure S7).

The Ruc204 x S29 cross. F_2s from the Ruc204xS29 cross showed several phenotypic classes (Figure S6b.). In this cross, the segregation ratio of the parental (Ruc204) type segregants (21) to the other F_2s (128) deviated from both the 1:15 (χ^2=15.6) and 1:3 (χ^2=9.5) ratios. The interval mapping for the SS trait detected two QTLs on chromosomes 2DS and 2AS, which explained 47.4% (LOD = 13.25) and 16% (LOD = 3.8) of the phenotypic variance, respectively, suggesting that the main contribution was from the 2D QTL (Figure S7). The 2D QTL was localized to the region of the interstitial deletion (Figures S4 and S7). Thus, both of the 2DS deletions contribute to the SS phenotype of the Ruc204 and Skle128 lines.

The MC1611 x Skala cross. The F_1 plants had spikes of the conventional type. The F_2 progeny showed either SS or the conventional spike type. The segregation ratio was close to 3 normal spiked to 1 SS (χ^2=8.6; 0.01<P<0.001) (Table S1), suggesting that, in addition to the major gene, a minor gene(s) affects the mutant phenotype. Using monosomic analysis, a gene responsible for the SS phenotype of MC1611 was previously assigned to 2D (Laikova et al., 2005), and here we defined the position of the QTL for SS on the 2DS map (Supplemental Figure 5) using interval mapping.

COS mapping - COS-SSCP analysis. A set of 62 COS -SSCP markers developed for the short arm of the 2-group chromosomes (Quraishi et al., 2009) were tested for polymorphisms between the parents of the Ruc163 x So149 and Ruc167 x So149 mapping populations. Polymorphic markers were detected and

applied for genotyping of the F_2s of both populations; seven markers for the Ruc167 x So149 population and four markers for the Ruc163 x So149 population were chosen. One marker, BG263521-5p, was mapped on the 2D maps of the two mapping populations (Figure 2A). This marker was located proximally to *mrs1,* and to obtain a COS marker distally flanking this gene, we performed *in silico* mapping of the sequenced RFLPs from the 2DS consensus map (Figure 2A). The map was constructed based on the 2DS consensus map from (Dobrovolskaya et al., 2009) and includes the *mrs1* gene and the BG263521-5p (COS1) marker (Figure 2A).

The *WFZP* expression pattern - Total RNA was extracted from young spikes of Chinese Spring plants grown under greenhouse conditions. The spikes were at the following developmental stages:

1) *Spikelet differentiation stage*. At this stage, the spikelets are differentiated, and glume primordia are initiated at the base of each spikelet;
2) *Early floret differentiation stage 1.* Floret primordia are initiated in each spikelet; lemma primordia start to be visible;
3) *Early floret differentiation stage 2.* Florets are forming on opposite sides of the spikelet axis, rachilla; floret meristem starts to produce floral organs;
4) *Late floret differentiation stage.* Floral organs are differentiated in each floret; inflorescence meristem initiates the terminal spikelet.

Quantitative RT-PCR. To understand the expression pattern of *WFZP*, we performed quantitative RT–PCR analysis of *WFZP* transcription in different organs, young spikes at various stages of development and young leaves and roots of bread wheat cv. Chinese Spring (Cs), using A, B, and D homoeoallele-specific primers (sequences of the primers are given in Table S3). The *WFZP* gene showed organ-specific and stage-specific expression patterns. *WFZP* transcripts were detected in young spikes with the highest level of expression at the early floret developmental stage (Figures 4D). Moreover, the results of the quantitative RT-PCR showed a consistently higher relative expression of *WFZP-D* compared to that of *WFZP-A* and *WFZP-B*: up to a twenty-fold increase was detected for *WFZP-D/WFZP-B* and three to four-fold increases were detected for *WFZP-D/WFZP-A*. The *WFZP-B* relative expression was very low at any stage of development; this low expression might be a consequence of an ME insertions into the proximal promoter of the *WFZP-B* gene.

REFERENCES

Altschul, S.F., Gish, W., Miller W., Myers, E.W., Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol*., **215,** 403-410.

Badaeva, E.D., Badaev, N.S., Gill, B.S., Filatenko, A.A. (1994). Intraspecific karyotype divergence in *Triticum araraticum* (Poaceae). *Plant Syst. Evol*., **192,** 117-145.

De Castro, E., Sigrist, C..JA., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch A., Hulo N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional residues in proteins. *Nucleic Acids Res.*, **34,** 362-365.

Dibari, B., Murat, F., Chosson, A., Gautier, V., Poncet, C., Lecomte, P., Mercier, I., Bergès, H., Pont, C., Blanco, A., Salse, J. (2012). Deciphering the genomic structure, function and evolution of carotenogenesis related phytoene synthases in grasses. *BMC Genomics,* **13**, 221.

Dobrovolskaya, O., Martinek, P., Röder, M.S., Börner, A. (2008). Microsatellite mapping of a mutant gene (*mrs*) for multirow spike in wheat (*T. aestivum*). In Proceeding of International conference

"Conventional and molecular breeding of field and vegetable crops" 22-27 November 2008, Novi Sad, Serbia, pp 133-136.
Dobrovolskaya, O., Martinek, P., Voylokov, A.V., Korzun, V., Röder, M.S., Börner, A. (2009). Microsatellite mapping of genes that determine supernumerary spikelets in wheat (*T. aestivum*) and rye (*S. cereale*). *Theor. Appl. Genet.,* **119,** 867-874.
Flutre, T., Duprat, E., Feuillet, C., Quesneville, H. (2011). Considering transposable element diversification in *de novo* annotation approaches. *PLoS One*, **6,** e16526.
Gill, B.S., Friebe, B., Endo, T.R. (1991). Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome,* **34,** 830-839**.**
Hunter, S., Apweiler, R., Attwood, T.K., *et al.,* (2009). InterPro: the integrative protein signature database. *Nucleic Acid Res*., **37,** 224-228.
Jurka, J. (2000). Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet*., **16,** 418-420.
Kosambi, D.D. (1943). The estimation of map distances from recombination values. *Ann. Eugen*., **12,** 172-175.
Kronmiller, B.A., Wise, R.P. (2008). TEnest: Automated Chronological annotation and Visualization of Nested Plant Transposable Elements. *Plant Physiol*., **146,** 45-59.
Laikova, L.I., Arbuzova, V.S., Popova, O.M., Efremova, T.T., Melnick, V.M. (2005). Study on spike branching in the *T. aestivum* mutant lines (cv. Saratovskaya29). In: Goncharov, P.L., Zilke, R.A., Gordeeva, T.N. (eds) Proceeding of the IX workshop on genetics and breeding, Novosibirsk, Russia, pp 388-393.
Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daley, M., Lincoln, S., Newburg L. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics,* **1,** 174-181.
Letunic, I., Doerks, T., Bork, P. (2009). SMART 6: recent uploads and developments. *Nucleic Acid Res.,* **37,** 229-232.
Li, J., Wang, Q., Wei, H., Hu, X., Yang, W. (2011) SSR Mapping for locus conferring on the triple-spikelet trait of the Tibetan triple-spikelet wheat (*Triticum aestivum* L. concv. tripletum). *Triticeae Genomics. Genet.*, 2(1), 1-6.
Nicot, N., Chiquet, V., Gandon, B., Amilhat, L., Legeai, F., Leroy, F., Bernard, M., Sourdille, P. (2004). Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theor. Appl. Genet.,* **109,** 800-805.
Ooijen, J.W. (2004) MapQTL® 5 Software for the mapping of quantitative trait loci in experimental populations. Wageningen: Kyazma BV.
Pont, C., Murat, F., Guizard, S., *et al.,* (2013). Wheat Syntenome Unveils New Evidences of Contrasted Evolutionary Plasticity Between paleo- and Neoduplicated Subgenomes. *Plant Journal,* **76**(6), 1030-44.
Quraishi, U.M., Abrouk, M., Bolot, S., *et al.,* (2009). Genomics in cereals: from genome wide conserved orthologous set (COS) sequences to candidate genes for trait dissection. *Funct Integr Genomics*, **9,** 473-484.
Salamov, A., Solovyev, V. (2000). Ab initio gene finding in *Drosophila* genomic DNA. *Genome Res.,* **10,** 516-522.
Schultz, J., Milpets, F., Bork, P., Ponting, C.P. (1998). SMART, a simple modular architecture research tool: identification of signaling domain. *Proc Natl Acad Sci U S A*., **95,** 5857-5864.
Yang, W.-Y., Lu, B.-R., Hu, X.-R., Yu, Y., Zhang, Y. (2005). Inheritance of the triple-spikelet character in a Tibetan landrace of common wheat. *Genet. Resour. Crop. Ev.,* **52,** 847-851.

Table S1: Segregation data for the numbers of normal-spiked (NS) and supernumerary spikelet (SS) plants in the F_2 and BC_1 progenies.

	Parents of crosses (allelic composition)	Type of cross	F1	F2/BC1					
				Total	Observed segregation		Segregation ratio	χ2	P
					NS	SS	NS to SS		
1	**Ruc167* (*wfzp* -A.1 *wfzp* -D.1) x So149 (*wfzp* -A.1 *WFZP* -D) [1]**	**F2****	**NS**	106	80	26	3 to 1	0,013	0.90-0.95
2	**Ruc163* (*wfzp -A.1 wfzp -D.1*) x So149* (*wfzp-A.1 WFZP-D*) [1]**	**F2****	**NS**	100	78	22	3 to 1	0,48	0.25-0.50
3	**NIL-mrs1* (*wfzp -A.1 wfzp -D.1*) x Skle128 *(wfzp -A.1 wfzp -D.1*)**	**F2*****	**SS**	140	-	140	-	-	-
4	**NIL-mrs1* (*wfzp -A.1 wfzp -D.1*) x Ruc204 *(wfzp -A.1 wfzp -D.1*)**	**F2*****	**SS**	140	-	140	-	-	-
5	**NIL-mrs1* (*wfzp -A.1 wfzp -D.1*) x MC1611 (*WFZP -A wfzp -D.2*)**	**F2*****	**SS**	142	26	116	1 to 3	4,2	0.05-0.025
6	**MC1611 (*WFZP -A wfzp -D.2*) x S29 (*WFZP -A WFZP -D*)**	**F2**	**NS**	158	121	37	3 to 1	0,21	0.50-0.75
7	**MC1611 (*WFZP -A wfzp -D.2*) x S29 (*WFZP -A WFZP -D*)**	**BC1 (MC1611/S29// MC1611)**	**NS**	155	77	78	1 to 1	0,013	0.90-0.95
8	**MC1611 (*WFZP -A wfzp -D.2*) x S29 (*WFZP -A WFZP -D*)**	**BC1 (MC1611/S29// S29)**	**NS**	140	140	-	-	-	-
9	**MC1611 (*WFZP -A wfzp -D.2*) x Skala (*WFZP -A WFZP -D*)**	**F2****	**NS**	135	116	19	3 to 1 15 to 1	8,6 14,2	0.001-0.01
10	**Skle128 *(wfzp -A.1 wfzp -D.1*) x S29 (*WFZP-A WFZP-D*)**	**F2****	**NS**	134	124'	10	15 to 1	0,34	0.50-0.75
11	**Ruc204*(wfzp -A.1 wfzp -D.1*) x S29 (*WFZP-A WFZP-D*)**	**F2****	**NS**	149	128'	21	3 to 1 15 to 1	9,5 15,6	

** used as a mapping population
*** used for complementation test
' this class of segregants consist of all the F2s that do not have the SS parental phenotype
in red are mutant alleles
in green are wild alleles

[1] Dobrovolskaya, O. et al. Microsatellite mapping of genes that determine supernumerary spikelets in wheat *(T. aestivum)* and rye (*S. cereale*). Theor. Appl. Genet. 119, 867-874 (2009).

Table S2: List of orthologous genes (rice, sorghum, maize, *Brachypodium*) at the *Mrs1* locus.

Markers name	COS	Rice7	Brachy1	Sorghum2	Maize2	Maize7	Rice function
CDO405	BE499478	LOC_Os07g45064	Bradi1g19650.1chr01	Sb02g041160			ATM-like protein, putative, expressed
		LOC_Os07g45080		Sb02g041170	ZmEvi110931-2		expressed protein
		LOC_Os07g45090		Sb02g041180	ZmEvi110933-2	ZmEvi068263-7	NADH-ubiquinone oxidoreductase 51 kDa subunit, mitochondrial precurs.
		LOC_Os07g45260		Sb02g041310			transferase, transferring glycosyl groups, putative, expressed
		LOC_Os07g45280	Bradi1g19570.1chr01	Sb02g041320			protein Mo25, putative, expressed
		LOC_Os07g45290		Sb02g041330			cytochrome P450 72A1, putative, expressed
		LOC_Os07g45320	Bradi1g19530.1chr01	Sb02g041360		ZmEvi039280-7	COX VIIa-like protein, putative, expressed
		LOC_Os07g45350	Bradi1g19510.1chr01	Sb02g041380		ZmEvi039308-7	ZCF61, putative, expressed
		LOC_Os07g45360	Bradi1g19500.1chr01	Sb02g041390		ZmEvi093709-7	ATP-dependent RNA helicase DDX52, putative, expressed
		LOC_Os07g46160	Bradi1g19390.1chr01	Sb02g041470		ZmEvi093725-7	speckle-type POZ protein, putative, expressed
		LOC_Os07g46170	Bradi1g19380.1chr01	Sb02g041460	ZmEvi120564-2		transmembrane protein 115, putative, expressed
		LOC_Os07g46180	Bradi1g19370.1chr01_pseudomolecule brac version0 15495039-15483590 BestGuessCds				PWWP domain containing protein, expressed
		LOC_Os07g46220	Bradi1g19340.1chr01	Sb02g041420	ZmEvi024687-2	ZmEvi093711-7	expressed protein
		LOC_Os07g46240	Bradi1g19300.1chr01	Sb02g041530	ZmEvi120572-2		expressed protein
		LOC_Os07g46280	Bradi1g19270.1chr01	Sb02g041550		ZmEvi093724-7	non-cyanogenic beta-glucosidase precursor, putative, expressed
		LOC_Os07g46300	Bradi1g19260.1chr01	Sb02g041570			expressed protein
		LOC_Os07g46310	Bradi1g19220.1chr01_pseudomolecule brac version0 15396067-15401935 BestGuessCds				magnesium-chelatase subunit H, putative
		LOC_Os07g46340		Sb02g041600			methylase, putative, expressed
		LOC_Os07g46350	Bradi1g19230.1chr01_pseudomolecule brac version0 15409996-15413928 BestGuessCds				serine carboxypeptidase 1 precursor, putative, expressed
		LOC_Os07g46360		Sb02g041630			lactoylglutathione lyase, putative, expressed
		LOC_Os07g46370	Bradi1g19210.1chr01	Sb02g041640			EMB2221, putative, expressed
		LOC_Os07g46380		Sb02g041650			glycosyltransferase, putative, expressed
		LOC_Os07g46410	Bradi1g19140.1chr01_pseudomolecule brac version0 15337267-15332131			ZmEvi120118-7	bifunctional thioredoxin reductase/thioredoxin, putative, expressed
		LOC_Os07g46450	Bradi1g19100.1chr01	Sb02g041730			rho GTPase activator, putative, expressed
		LOC_Os07g46460	Bradi1g19080.1chr01_pseudomolecule brac version0 15284037-15301108 BestGuessCds				ferredoxin-dependent glutamate synthase, chloroplast precursor, putat. expr.
		LOC_Os07g46480		Sb02g041760			aspartic proteinase nepenthesin-1 precursor, putative, expressed
		LOC_Os07g46490	Bradi1g19050.1chr01	Sb02g041770			phosphatidylinositol-4-phosphate 5-kinase 4, putative, expressed
		LOC_Os07g46500	Bradi1g19030.1chr01	Sb02g041780		ZmEvi128244-7	tankyrase 2, putative, expressed
		LOC_Os07g46520	Bradi1g19010.1chr01	Sb02g041820			rhythmically expressed gene 2 protein, putative, expressed
		LOC_Os07g46540	Bradi1g19000.1chr01_pseudomolecule brac version0 15203222-15211438 BestGuessCds				condensin complex subunit 1, putative, expressed
		LOC_Os07g46560	Bradi1g18970.1chr01	Sb02g041870		ZmEvi131694-7	ubiquitin ligase SINAT5, putative, expressed
		LOC_Os07g46570		Sb02g041880			electron transporter, putative, expressed
		LOC_Os07g46590	Bradi1g18910.1chr01	Sb02g041910			chromodomain-helicase-DNA-binding protein, putative, expressed
		LOC_Os07g46600	Bradi1g18900.1chr01	Sb02g041930			RNA pseudouridylate synthase family protein, expressed
		LOC_Os07g46610		Sb02g041935			cis-zeatin O-glucosyltransferase, putative, expressed
		LOC_Os07g46630	Bradi1g18810.1chr01	Sb02g041950			deaminase, putative, expressed
		LOC_Os07g46640	ecule brac version0 15	Sb02g041960			UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminepho-sphotransferase, putative, expressed
		LOC_Os07g46700	Bradi1g18760.1chr01	Sb02g042010			RNA-binding protein, putative, expressed
		LOC_Os07g46710	Bradi1g18740.1chr01_pseudomolecule brac version0 15009973-15011923 BestGuessCds				expressed protein
		LOC_Os07g46720	Bradi1g18730.1chr01	Sb02g042030			nucleolar protein Nop56, putative, expressed
		LOC_Os07g46750	Bradi1g18690.1chr01	Sb02g042050			elongation factor 1-beta, putative, expressed
		LOC_Os07g46760		Sb02g042060			activator of basal transcription 1, putative, expressed
		LOC_Os07g46770	Bradi1g18670.1chr01_pseudomolecule brac version0 14968812-14967178 BestGuessCds				pentatricopeptide repeat protein PPR868-14, putative, expressed
		LOC_Os07g46780	Bradi1g18660.1chr01	Sb02g042070			tyrosine-specific transport protein, putative, expressed
		LOC_Os07g46790	Bradi1g18650.1chr01	Sb02g042100			4-alpha-glucanotransferase, putative, expressed
		LOC_Os07g46980		Sb02g042150			sex determination protein tasselseed-2, putative, expressed
		LOC_Os07g46990	Bradi1g18340.1chr01	Sb02g042170			superoxide dismutase 2, putative, expressed
		LOC_Os07g47100	Bradi1g18360.1chr01	Sb02g042190			sodium/hydrogen exchanger 2, putative, expressed
		LOC_Os07g47110	Bradi1g18380.1chr01	Sb02g042210			phosphoric diester hydrolase/transcription factor, putative, expressed
		LOC_Os07g47120	Bradi1g18390.1chr01	Sb02g042220			beta-amylase, putative, expressed
		LOC_Os07g47150	Bradi1g18420.1chr01	Sb02g042240	ZmEvi106880-2		expressed protein
		LOC_Os07g47201	Bradi1g18460.1chr01_pseudomolecule brac version0 14822704-14816849 BestGuessCds				adenyl cyclase, putative, expressed
		LOC_Os07g47210	Bradi1g18470.1chr01	Sb02g042280		ZmEvi127679-7	anther-specific proline-rich protein APG, putative, expressed
		LOC_Os07g47250	Bradi1g18490.1chr01	Sb02g042310		ZmEvi131619-7	lipase precursor, putative, expressed
		LOC_Os07g47270	Bradi1g18500.1chr01	Sb02g042320		ZmEvi131624-7	protein kinase APK1B, chloroplast precursor, putative, expressed
		LOC_Os07g47280	Bradi1g18510.1chr01	Sb02g042330			DNA polymerase zeta catalytic subunit, putative, expressed
		LOC_Os07g47284		Sb02g042330			DNA polymerase zeta catalytic subunit, putative
		LOC_Os07g47290	Bradi1g18550.1chr01_pseudomolecule brac version0 14881092-14876604			ZmEvi131630-7	xylose isomerase, putative, expressed
	fzpF4R2	**LOC_Os07g47330**	**Bradi1g18580.1**	**Sb02g042400**	**ZmEvi005249-2**	**ZmEvi123374-7**	**branched silkless1, putative, expressed**
		LOC_Os07g47350	Bradi1g18600.1chr01	Sb02g042430		ZmEvi123380-7	potassium transporter 7, putative, expressed
		LOC_Os07g47360		Sb02g042440			CW-type Zinc Finger family protein, expressed
		LOC_Os07g47370	Bradi1g18310.1chr01	Sb02g042450			pentatricopeptide repeat protein PPR1106-17, putative, expressed
		LOC_Os07g47420	Bradi1g18250.1chr01_pseudomolecule brac version0 14684818-14682607 BestGuessCds				60S ribosome subunit biogenesis protein NIP7, putative, expressed
		LOC_Os07g47450	Bradi1g18240.1chr01	Sb02g042480	ZmEvi041656-2	ZmEvi131818-7	flowering promoting factor-like 1, putative, expressed
		LOC_Os07g47470	Bradi1g18220.1chr01	Sb02g042500			fertility restorer homologue A, putative
		LOC_Os07g47490	Bradi1g18200.1chr01_pseudomolecule brac version0 14651322-14655522			ZmEvi131828-7	indole-3-acetic acid-amido synthetase GH3.11, putative, expressed
		LOC_Os07g47500	Bradi1g18180.1chr01	Sb02g042540			histone-arginine methyltransferase CARM1, putative, expressed
		LOC_Os07g47510	Bradi1g18170.1chr01	Sb02g042550			stress-related protein, putative, expressed
		LOC_Os07g47530	Bradi1g18160.1chr01	Sb02g042560			ATP binding protein, putative, expressed
		LOC_Os07g47570	Bradi1g18130.1chr01	Sb02g042590			expressed protein
		LOC_Os07g47580	Bradi1g18120.1chr01	Sb02g042600	ZmEvi044907-2		TGF-beta-inducible nuclear protein 1, putative, expressed
		LOC_Os07g47590	Bradi1g18090.1chr01	Sb02g042630			expressed protein
		LOC_Os07g47620					universal stress protein, putative, expressed
		LOC_Os07g47630	Bradi1g18070.1chr01_pseudomolecule brac version0 14579020-14583908 BestGuessCds				pre-mRNA-splicing factor SF2, putative, expressed
		LOC_Os07g47670	Bradi1g18030.1chr01	Sb02g042700			hypoxia induced protein conserved region containing protein, expressed
		LOC_Os07g47710	Bradi1g18010.1chr01	Sb02g042710		ZmEvi118863-7	60S ribosomal protein L22-2, putative, expressed
		LOC_Os07g47750	Bradi1g17980.1chr01	Sb02g042730	ZmEvi022861-2		expressed protein
		LOC_Os07g47760		Sb02g042740			phylloplanin precursor, putative, expressed

Table S3: Primer sequences derived from *FZP-A-B-D* gene copies. Primer name (first column), primer sequence (second column) and location (third column).

Primer Name	Sequence 5' – 3'	location
BAC library screening		
FZP_F1	CACTGGCTCGGCACCTTC	CDS
FZP_R1	CTCAGGTACCCGGAGTTGTCG	
FZP_F1	CACTGGCTCGGCACCTTC	CDS
FZP_R2	AGGTCGTCCACGTCCTCCC	
Amplicon sequensing		
WFZP-A		
WFZP_2A_F1	CATGGGCAAATCGGTTAATG	5' region
WFZP_2A_R1	TGGATGAGATGGCGAGGTAG	
WFZP_F2	TCTTGTCAGTGGCAGGCATC	5' region
WFZP_2A_R2	TGGCAGAAGTGAAGTGAGGT	
WFZP_F3	GCTCACAGTCTCAGCAACCA	5'-UTR -CDS
WFZP_2A_R3	CACTGGGCACCGGCATGGAA	
WFZP_2AD_F4	CAGCCAACCTCACTTCACT	CDS
WFZP_2A_R4	GCTAGGGCACCGAAACAAC	
WFZP_F5	ACGACATGGTCGCCTCGT	CDS-3' region
WFZP_2A_R5	GGATCGGGGTGGATAGATTG	
WFZP_2AD_F5	CTCAGAGCCTCAGACCCATT	3' region
WFZP_2A_R5	GGATCGGGGTGGATAGATTG	
WFZP-B		
WFZP_2B_F1	TGGGCTCTGCCTTCACAATCAG	5' region
WFZP_2B_R1	GATGAGATGGCGACCTTGG	
WFZP_2B_F2	GGCACACAAATCCAAACACA	5 'region -CDS
WFZP_2B_R2	GCCGTGATCCGCGGCATTGA	
WFZP_2B_F3	ACAGTGCTCTCAGCCTCTCA	5'-UTR -CDS
WFZP_2B_R3	CGGTGCATTTGCTTCAGTGT	
WFZP_F3	GCTCACAGTCTCAGCAACCA	5'-UTR -CDS
WFZP_2B_R2	GCCGTGATCCGCGGCATTGA	
WFZP_2B_F4	GAGTGCTCCATGCCGGTGCT	CDS-3' region
WFZP_2B_R4	GCCACTTTAATTCCGCAGGACT	
WFZP_F5	ACGACATGGTCGCCTCGT	3' region
WFZP_2B_R5	GCATTGCTTAATTGGGTGAT	
WFZP_2B_F5	GAAGATCGTGAAGAAGAGTGG	3' region
WFZP_2B_R5	GCATTGCTTAATTGGGTGAT	
WFZP-D		
WFZP_2D_F1	TGTGCGCGCCGCAAAATCTT	5' region
WFZP_2D_R1	GGATTTGTGTGCCCTAACCTA	
WFZP_F6	TCTTGGCAGGGGCTTCAAT	5' region
WFZP_2D_R2	GCCCTGCTGGTGGTAGCCGA	
WFZP_2D_F2	CTCCTCATCTCTTTGGTCCT	5'-CDS
WFZP_2D_R2	GCCCTGCTGGTGGTAGCCGA	
WFZP_2D_F3	TCAGTTCTGCCATGAGCATC	5'UTR-CDS
WFZP_2D_R2	GCCCTGCTGGTGGTAGCCGA	

WFZP_2D_F2	CTCCTCATCTCTTTGGTCCT	CDS-3' UTR
WFZP_2D_R3	GCCACTCTTCTTCTTCGTCG	
WFZP_F5	ACGACATGGTCGCCTCGT	3' region
WFZP_2D_R4	ACGAAGTCACTACAAGCACA	
Quantitative RT- PCR ***WFZP-A***		
WFZP_F3	GCTCACAGTCTCAGCAACCA	5'-UTR-CDS
WFZP_2A_R2	TGGCAGAAGTGAAGTGAGGT	
WFZP-B		
WFZP_F5	ACGACATGGTCGCCTCGT	CDS-3'UTR
WFZP_2B_R3	CGGTGCATTTGCTTCAGTGT	
WFZP-D		
WFZP_F5	ACGACATGGTCGCCTCGT	CDS-3'UTR
WFZP_2D_R5	CTGGCTGGTGCATTTGTTG	
Detection of the *wfzp-D.1* allele		
WFZP_2D_F3	TCAGTTCTGCCATGAGCATC	CDS
wfzp-D.1_R*	TGGTCGGGTCGCGTATCTC	
Detection of the *wfzp-A.1* allele**		
WFZP-F7	ATGGCCTTCTCGGAGCATTC	CDS
WFZP-R3	TGGTCGGGTCGCGTATCTT	

* - *wfzp-D.1* allele specific primer;

** - *wfzp-A.1* allele is discriminated using 2% high-resolution MetaPhor agarose.

Table S4: *WFZP-A* alleles characterized in bread wheat cultivars. Cultivar name (first column), origin (second column) and allele (third column).

Cultivar name	Country of origin	*WFZP-A* allele
AURORE	AUS	w
BALKAN	YUG	w
SHORTANDINKA	KAZ	w
CHYAMTANG	NPL	w
COPPADRA	TUR	w
COTIPORA	**BRA**	***WFZP-A.1***
CTIPORA	BRA	w
FRUMENTO CUPO	ITA	w
GLENLEA-CAN	CAN	w
GODOLLOI 15	HUN	w
MISKAAGANI	LIB	w
MOCHO DE ESPIGA BIANCA	PRT	w
NP120	IND	w
NYU BAY	**JPN**	***WFZP-A.1***
OPAL	DEU	w
PITIC 62	MEX	w
SEU SEUN 27	KOR	w
HOPEA	FIN	w
XERES	ESP	w
ZANDA	BEL	w
ORNICAR	FRA	w
APACHE	FRA	w
CHINESE SPRING	CHN	w
RENAN	FRA	w
SO149	**CZE**	***WFZP-A.1***
SARATOVSKAYA 29	RUS	w
NOVOSIBIRSKAYA 67	RUS	w
SKALA	RUS	w
BIYSKAYA OZIMAYA	RUS	w
OMSKAYA23	RUS	w
CHELABA75	RUS	w
VASSA	RUS	w
RASSVET	BLR	w
FESTIVAL	BLR	w
PAMIATI VAVENKOVA	RUS	w
TULAIKOV 100	RUS	w
PHILATOVKA	RUS	w
NOVOSIBIRSKAYA 40	RUS	w
BEZENCHUKSKAYA 98	RUS	w
NOVOSIBIRSKAYA 32	RUS	w
FISHT	RUS	w
TANYA	RUS	w
KULUNDINKA	RUS	w
BAGRATION	RUS	w
NOVOSIBIRSKAYA 15	RUS	w
NOVOSIBIRSKAYA 31	RUS	w
NOVOSIBIRSKAYA 44	RUS	w
NOVOSIBIRSKAYA 9	RUS	w

'w' (third column) designated the *WFZP-A* wild allele.

Table S5: SS phenotypes and associated mutations in *WFZP-D*.

Line	Phenotype (type of the SS trait)	Origin	Mutation in *WFZP-D*	*WFZP-D* allele
NIL-*mrs1* [1] Ruc163 (*mrs1*) Ruc167 (*mrs1*) ten V3 lines (*mrs1*)[2]	MRS MRS MRS and GB MRS	chemical mutagenesis	non-synonimous in AP2/ERF	*wfzp-D.1*
Ruc204	HS and GB[3]	unknown	complete deletion	*null*-allele
Skle128	HS (triple-spikelet)	spontaneous	complete deletion	*null*-allele
So164	HS (tetrastichon)	cross of a branched tetraploid with a common bread wheat line	complete deletion	*null*-allele
MC1611[4]	HS (weak)	chemical mutagenesis	premature stop codon in AP2/ERF	*wfzp-D.2*

[1] - near isogenic line (recurrent parent is cv. Novosibirskaya 67). The MRS trait was transferred from the hexaploid wheat gene resource 'Ra1', a mutant that was produced through chemical mutagenesis (Dobrovolskaya et al., 2009), and all the MRS lines bear the *mrs1* gene in different genetic backgrounds.

[2] - V3-79-08, V3-82-08, V3-83-08, V3-84-08, V3-85-08, V3-86-08, V3-87-08, V3-88-08, V3-89-08, and V3-90-08 lines.

[3] – GB under field conditions

[4] - MC1611 derived from cv. Saratovskaya 29 via chemical mutagenesis.

MRS – multirow spike; GB – genuine branching; HS – horizontal spikelets

Figure S1: Spike morphology of bread wheat lines with different SS phenotyes. The middle part of spikes and schematic illustration of spike structure of (A) cv. Novosibirskaya 67 (N67), (B) NIL-*mrs1*, (C) Ruc204 grown under greenhouse (left) and field (right) conditions, (D) Skle128, (E) So164, (F) MC1611. Abbreviation: RS – ramified spike, HS – horizontal spikelet, MRS – multirow spike; SS - supernumerary spikelets, WT – wild type.

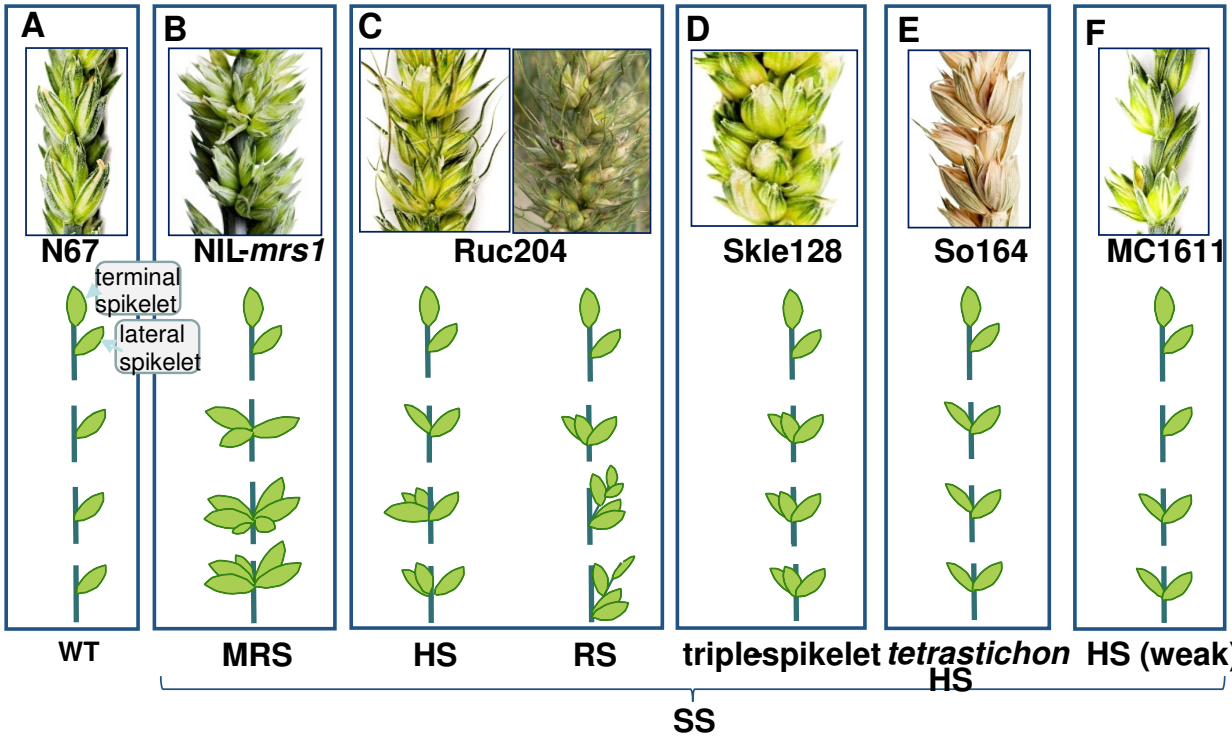


Figure S2: Genetic mapping of the NIL-*mrs1* line on the chromosome 2D. (**A**) Chromosome 2DS consensus map (**B**) from Dobrovolskaya et al. (2009). Green and red bars indicate chromosomal regions of the N67 recurrent parent, and KM240 (donor of MRS), respectively.

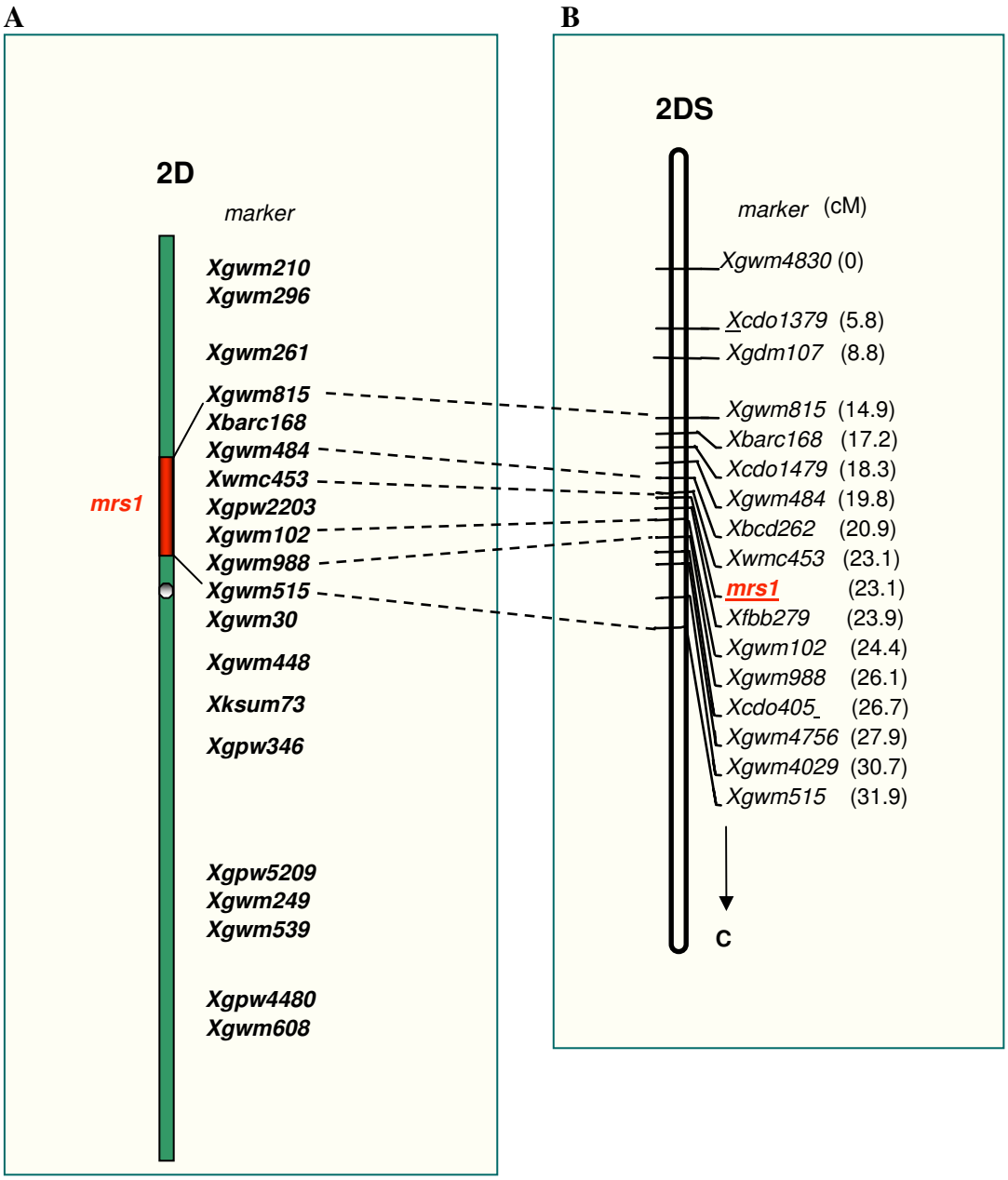


Figure S3: Light microscopy image of the MC1611 inflorescence showing developing ectopic spikelets.. Abbreviations: es – ectopic spikelet; sm, fm, gl – spikelet meristem, floret meristem and glume of a primary spikelet; sm*, fm*, gl* – spikelet meristem, floret meristem and glume of an ectopic spikelet

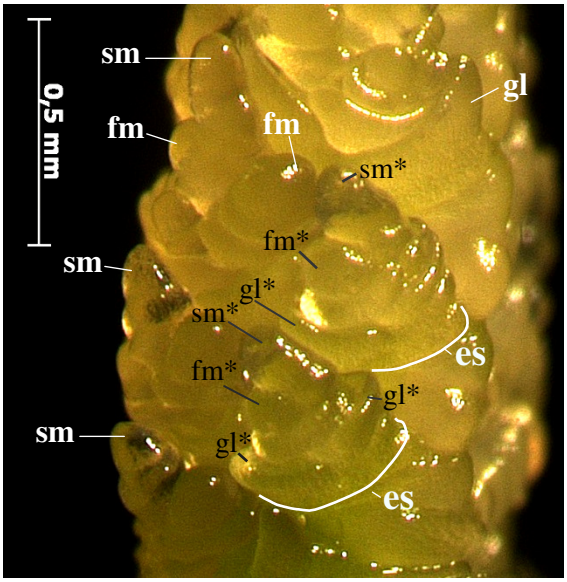
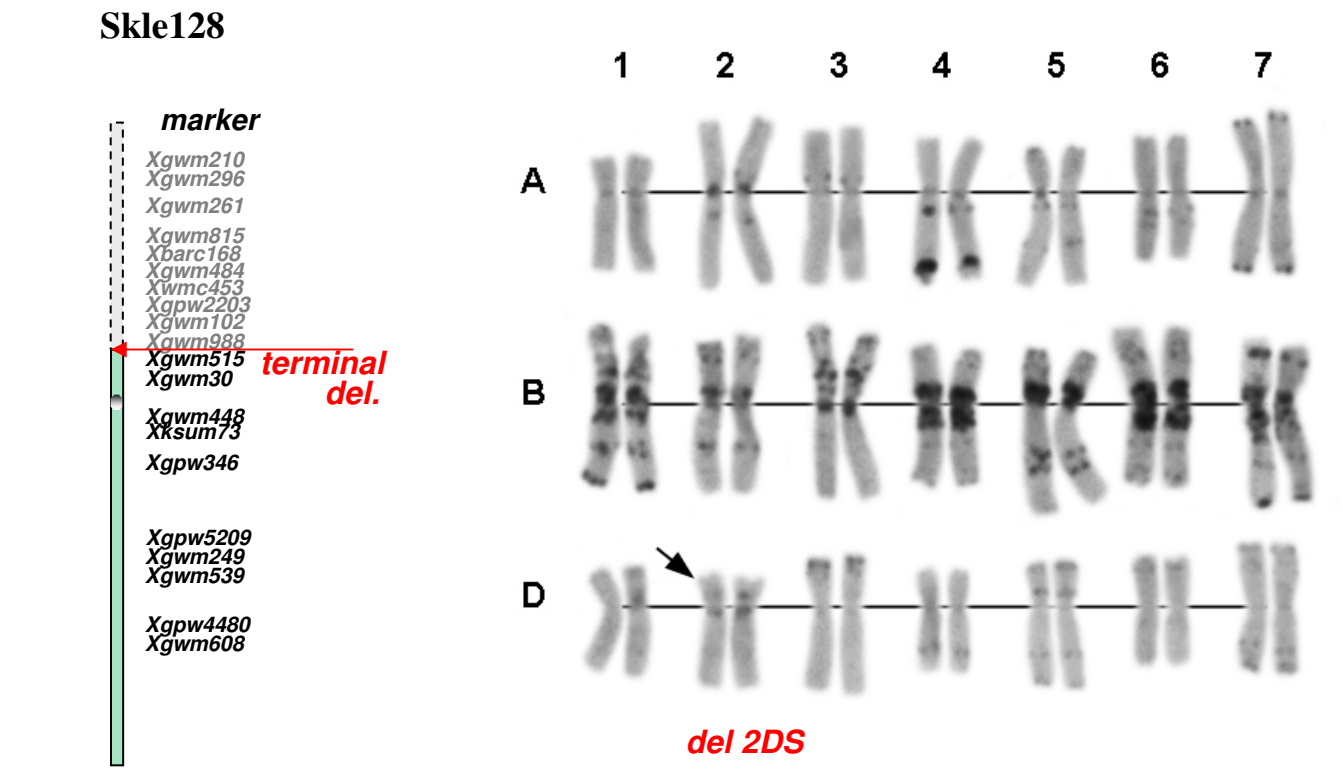
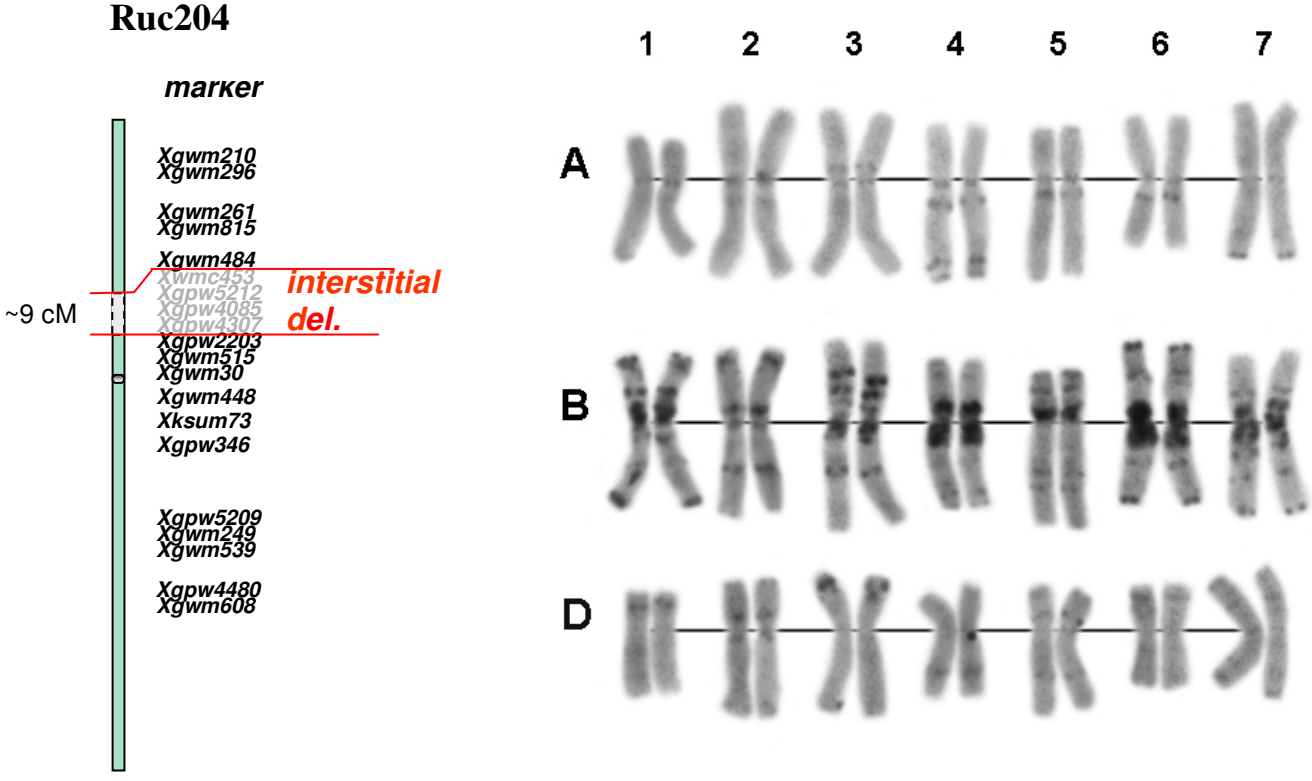


Figure S4: Deletion breakpoint mapping and ideograms for the Skle128 (a) and Ruc204 (b) lines on chromosome 2D. Deleted SSR loci are in grey. (c) Comparison of the lengths of chromosomes 2D of Skle128 and Ruc204 with that of cv. Pitic (control) showing ~10% and ~50% reductions, respectively.

A



B



C

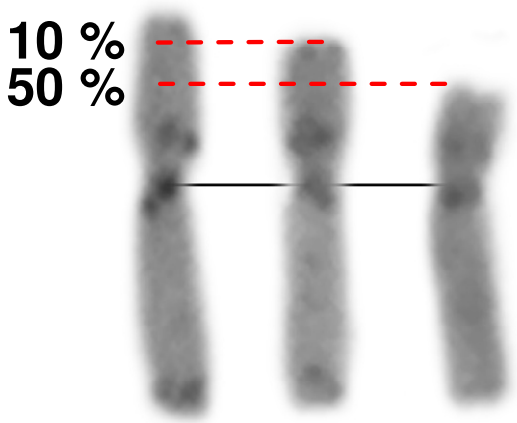


Figure S5: Ideogram of the So164 line (see legend Figure S4).

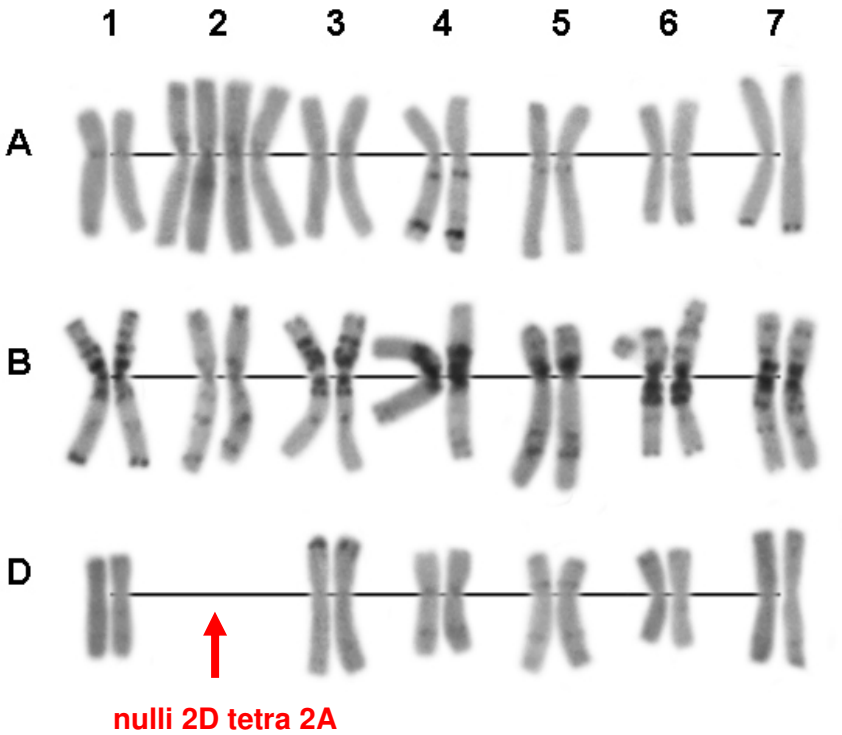
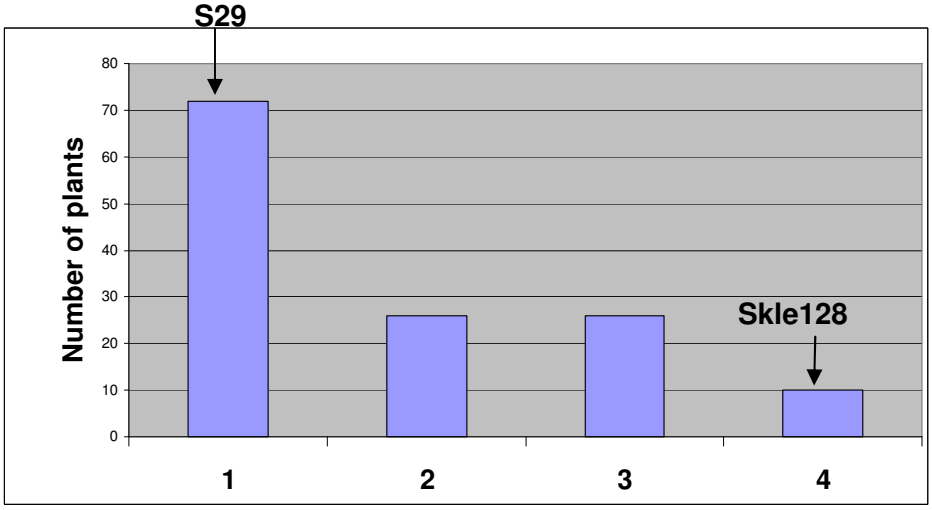


Figure S6: Frequency distribution of supernumerary spikelets in Skle128 x S29 and Ruc204 x S29 F_2 mapping populations. Plants from the classes #2 were similar to the normal spiked parent, S29, and had one or two supernumerary spikelets per a spike; plants from class #3 had numerous SS spikelets; class #5 includes plans with more severe SS phenotypes than SS parent.

A



B

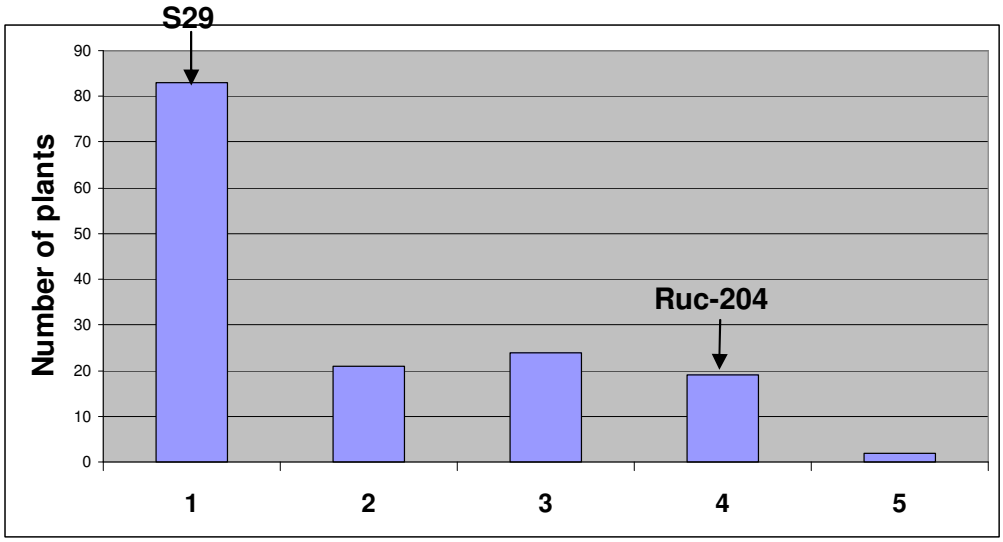


Figure S7: *mrs1* gene and QTL mapping in five F_2 mapping populations showing collocation of *mrs1*, SS-QTLs and *WFZP-D* on the chromosome 2D. (**A**, **B**) - 2D genetic maps including *mrs1* (the Ruc163 x So149 and Ruc167 x So149 crosses, respectively); (**C**) graphical representation of 2D in NIL-*mrs1* (green bars), showing a region of introgression (blue bar); (**D**, **E**, **F**) 2D genetic map including the ***qSS-D*** QTLs (the MC1611 x Skala, Skle128 x S29, Ruc204 x S29 crosses, respectively; Ruc204 and Skle128 carry deletions in 2DS). Red boxes delimit regions with calculated LOD score values indicated at the bottom of maps.

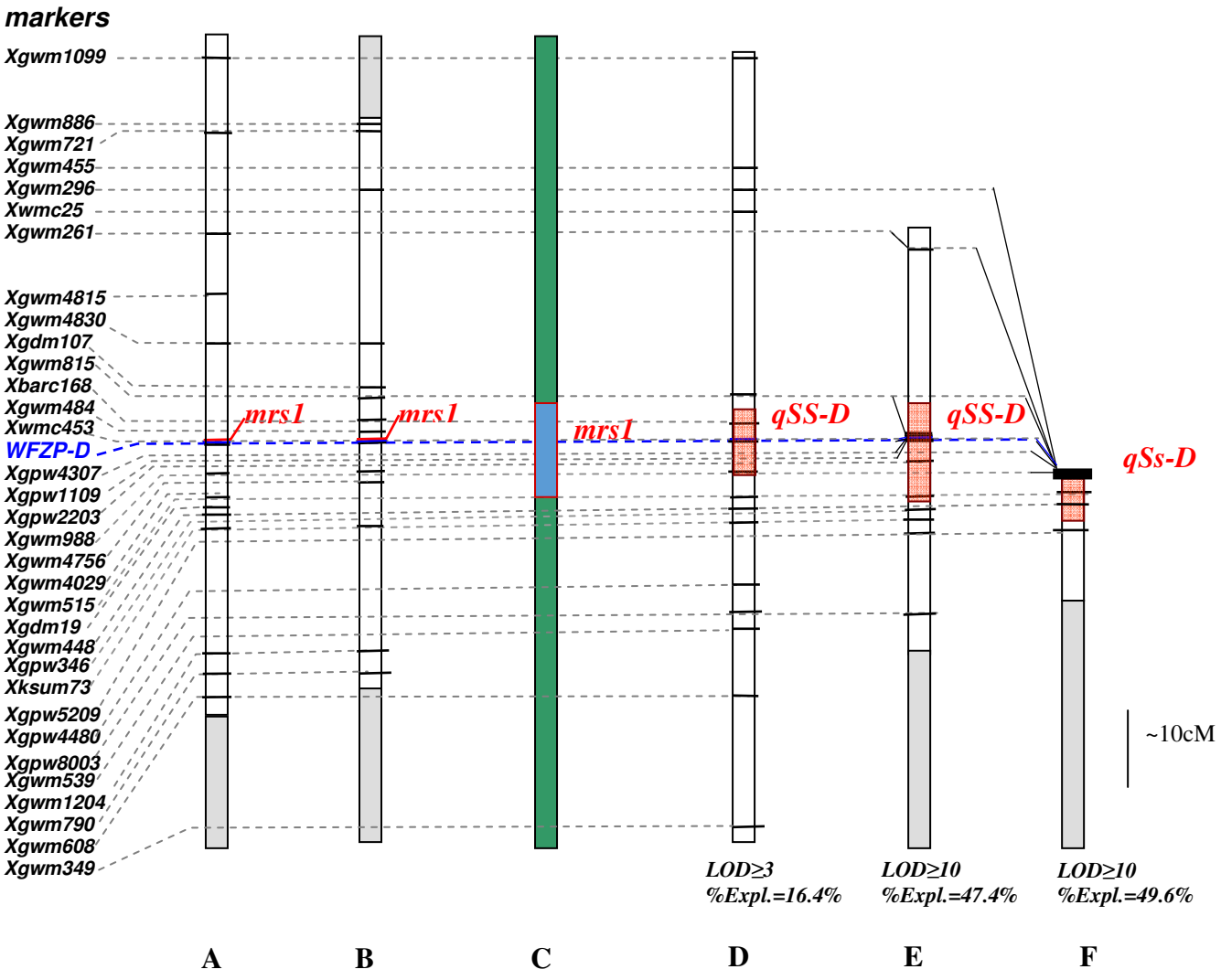
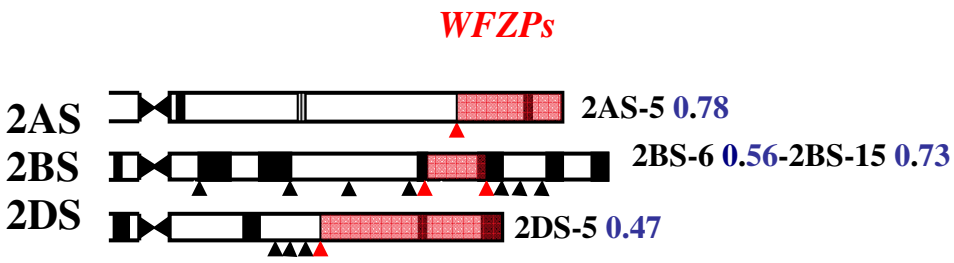


Figure S8: Alignment of wheat *WFZP* gene copies with rice and maize orthologs. The AP2/ERF domain is underlined in red.

.

Figure S9: Deletion bins and genetic positions of *WFZP* homoeologous genes on chromosomes 2AS, 2BS, 2DS. (**A**) Arrows indicate deletion breakpoints; dark bands on the chromosome indicate the location of C-bands (Gill et al., 1991); red arrows delimit the deletion bins that contain the *WFZP* homoeo-allelic genes. (**B**) Partial 2AS, 2BS and 2D genetic maps including the *WFZP* genes. Genetic mapping was performed on a subset of 96 F_2 individuals derived from a cross between cv. Chinese Spring and cv. Renan.

A



B

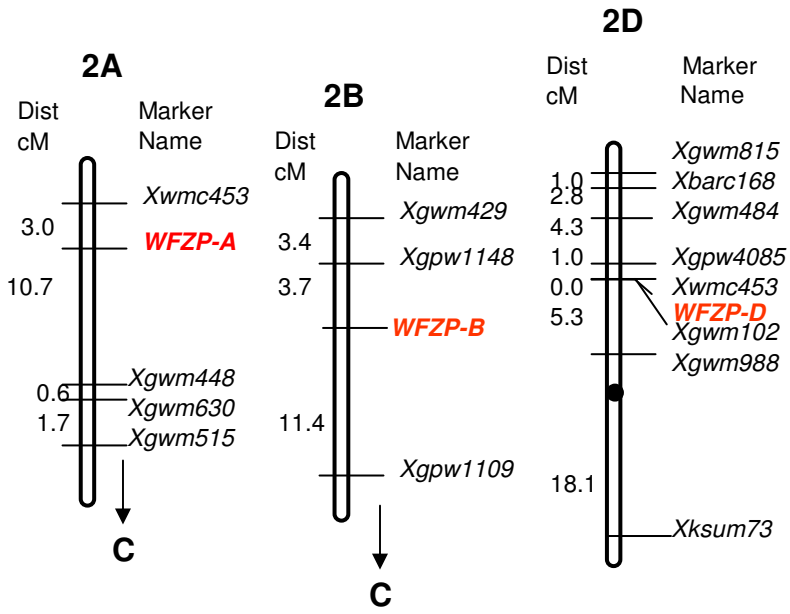
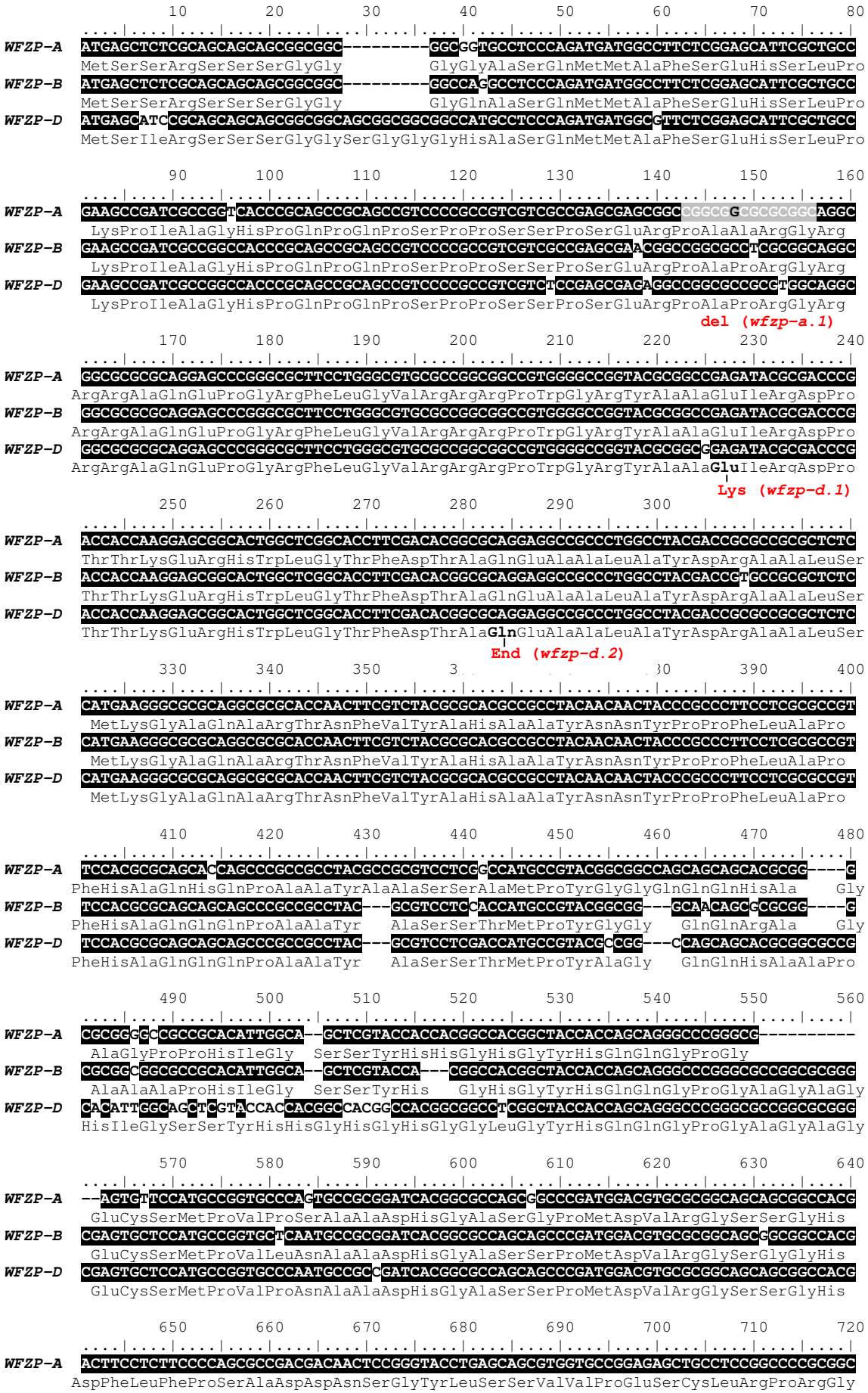
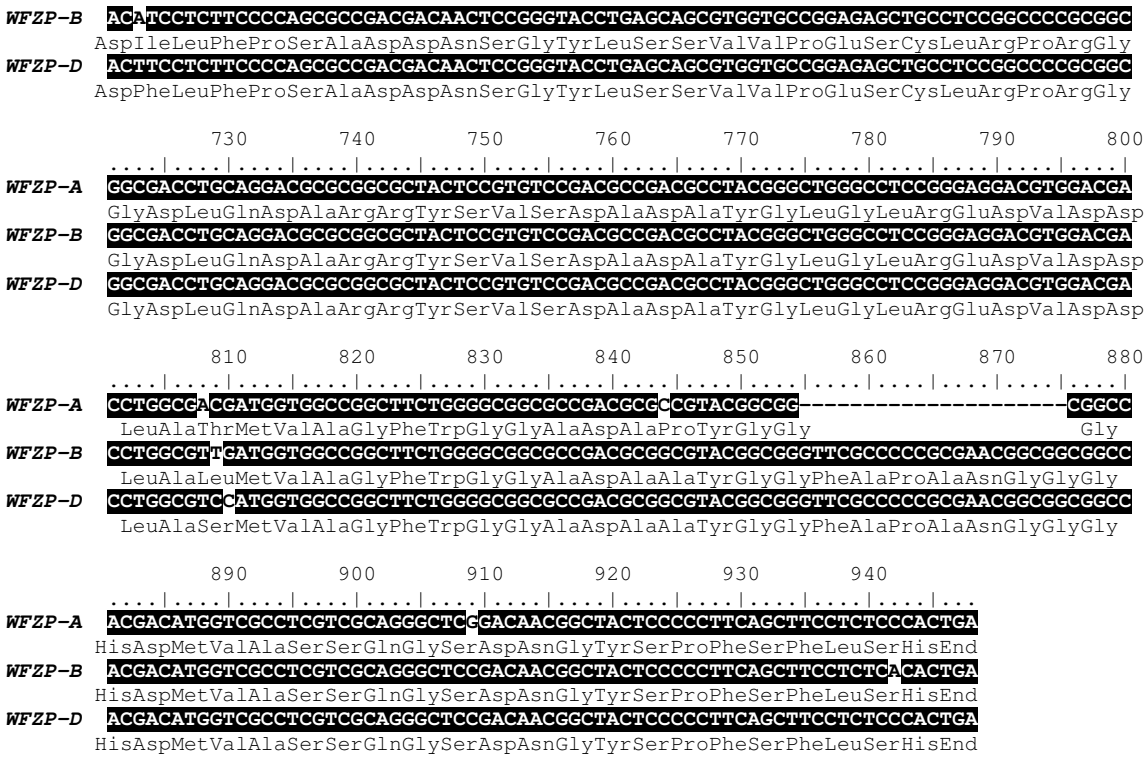


Figure S10: Alignment of deduced *WFZP-A, WFZP-B* and *WFZP-D* amino-acid sequences. Mutation are indicated in red.




WFZP-B ACATCCTCTTCCCCAGCGCCGACGACAACTCCGGGTACCTGAGCAGCGTGGTGCCGGAGAGCTGCCTCCGGCCCCGCGGC
AspIleLeuPheProSerAlaAspAspAsnSerGlyTyrLeuSerSerValValProGluSerCysLeuArgProArgGly
WFZP-D ACTTCCTCTTCCCCAGCGCCGACGACAACTCCGGGTACCTGAGCAGCGTGGTGCCGGAGAGCTGCCTCCGGCCCCGCGGC
AspPheLeuPheProSerAlaAspAspAsnSerGlyTyrLeuSerSerValValProGluSerCysLeuArgProArgGly
730 740 750 760 770 780 790 800
WFZP-A GGCGACCTGCAGGACGCGCGGCGCTACTCCGTGTCCGACGCCGACGCCTACGGGCTGGGCCTCCGGGAGGACGTGGACGA
GlyAspLeuGlnAspAlaArgArgTyrSerValSerAspAlaAspAlaTyrGlyLeuGlyLeuArgGluAspValAspAsp
WFZP-B GGCGACCTGCAGGACGCGCGGCGCTACTCCGTGTCCGACGCCGACGCCTACGGGCTGGGCCTCCGGGAGGACGTGGACGA
GlyAspLeuGlnAspAlaArgArgTyrSerValSerAspAlaAspAlaTyrGlyLeuGlyLeuArgGluAspValAspAsp
WFZP-D GGCGACCTGCAGGACGCGCGGCGCTACTCCGTGTCCGACGCCGACGCCTACGGGCTGGGCCTCCGGGAGGACGTGGACGA
GlyAspLeuGlnAspAlaArgArgTyrSerValSerAspAlaAspAlaTyrGlyLeuGlyLeuArgGluAspValAspAsp
810 820 830 840 850 860 870 880
WFZP-A CCTGGCGACGATGGTGGCCGGCTTCTGGGGCGGCGCCGACGCGCCGTACGGCGG----------CGGCC
LeuAlaThrMetValAlaGlyPheTrpGlyGlyAlaAspAlaProTyrGlyGly Gly
WFZP-B CCTGGCGTTGATGGTGGCCGGCTTCTGGGGCGGCGCCGACGCGGCGTACGGCGGGTTCGCCCCCGCGAACGGCGGCGGCC
LeuAlaLeuMetValAlaGlyPheTrpGlyGlyAlaAspAlaAlaTyrGlyGlyPheAlaProAlaAsnGlyGlyGly
WFZP-D CCTGGCGTCCATGGTGGCCGGCTTCTGGGGCGGCGCCGACGCGGCGTACGGCGGGTTCGCCCCCGCGAACGGCGGCGGCC
LeuAlaSerMetValAlaGlyPheTrpGlyGlyAlaAspAlaAlaTyrGlyGlyPheAlaProAlaAsnGlyGlyGly
890 900 910 920 930 940
WFZP-A ACGACATGGTCGCCTCGTCGCAGGGCTCGGACAACGGCTACTCCCCCTTCAGCTTCCTCTCCCACTGA
HisAspMetValAlaSerSerGlnGlySerAspAsnGlyTyrSerProPheSerPheLeuSerHisEnd
WFZP-B ACGACATGGTCGCCTCGTCGCAGGGCTCCGACAACGGCTACTCCCCCTTCAGCTTCCTCTCACACTGA
HisAspMetValAlaSerSerGlnGlySerAspAsnGlyTyrSerProPheSerPheLeuSerHisEnd
WFZP-D ACGACATGGTCGCCTCGTCGCAGGGCTCCGACAACGGCTACTCCCCCTTCAGCTTCCTCTCCCACTGA
HisAspMetValAlaSerSerGlnGlySerAspAsnGlyTyrSerProPheSerPheLeuSerHisEnd