



Dissection of Structural Reorganization of Wheat 5B Chromosome Associated With Interspecies Recombination Suppression

Elena Salina^{1,2*}, Alexander Muterko¹, Antonina Kiseleva^{1,2}, Zhiyong Liu³ and Abraham Korol⁴

¹Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia, ²Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia, ³Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China, ⁴Institute of Evolution, University of Haifa, Haifa, Israel

OPEN ACCESS

Edited by:

Ruslan Kalendar,
University of Helsinki, Finland

Reviewed by:

Ming Hao,
Sichuan Agricultural University, China
Khalil Kashkush,
Ben-Gurion
University of the Negev, Israel

*Correspondence:

Elena Salina
salina@bionet.nsc.ru

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 26 February 2022

Accepted: 08 April 2022

Published: 04 May 2022

Citation:

Salina E, Muterko A, Kiseleva A,
Liu Z and Korol A (2022) Dissection
of Structural Reorganization of Wheat
5B Chromosome Associated With
Interspecies Recombination
Suppression.
Front. Plant Sci. 13:884632.
doi: 10.3389/fpls.2022.884632

Chromosomal rearrangements that lead to recombination suppression can have a significant impact on speciation, and they are also important for breeding. The regions of recombination suppression in wheat chromosome 5B were identified based on comparisons of the 5B map of a cross between the Chinese Spring (CS) variety of hexaploid wheat and CS-5Bdic (genotype CS with 5B substituted with its homologue from tetraploid *Triticum dicoccoides*) with several 5B maps of tetraploid and hexaploid wheat. In total, two regions were selected in which recombination suppression occurred in cross CS × CS-5Bdic when compared with other maps: one on the short arm, 5BS_RS, limited by markers BS00009810/BS00022336, and the second on the long arm, 5BL_RS, between markers Ra_c10633_2155 and BS00087043. The regions marked as 5BS_RS and 5BL_RS, with lengths of 5 Mb and 3.6 Mb, respectively, were mined from the 5B pseudomolecule of CS and compared to the homoeologous regions (7.6 and 3.8 Mb, respectively) of the 5B pseudomolecule of Zavitan (*T. dicoccoides*). It was shown that, in the case of 5BS_RS, the local heterochromatin islands determined by the satellite DNA (119.2) and transposable element arrays, as well as the dissimilarity caused by large insertions/deletions (chromosome rearrangements) between 5BSs *aestivum/dicoccoides*, are likely the key determinants of recombination suppression in the region. Two major and two minor segments with significant loss of similarity were recognized within the 5BL_RS region. It was shown that the loss of similarity, which can lead to suppression of recombination in the 5BL_RS region, is caused by chromosomal rearrangements, driven by the activity of mobile genetic elements (both DNA transposons and long terminal repeat retrotransposons) and their divergence during evolution. It was noted that the regions marked as 5BS_RS and 5BL_RS are associated with chromosomal rearrangements identified earlier by C-banding analysis of intraspecific polymorphism of tetraploid emmer wheat. The revealed divergence in 5BS_RS and 5BL_RS may be a consequence of interspecific hybridization, plant genetic adaptation, or both.

Keywords: wheat, *dicoccoides*, recombination suppression, chromosome 5B, DNA transposons, LTR retrotransposons, tandem repeats, heterochromatin

INTRODUCTION

The evolutionary history of the wheat genome has been widely discussed over the last three decades (Goncharov, 2011; Gornicki et al., 2014; El Baidouri et al., 2017; Mirzaghaderi and Mason, 2017; Pont et al., 2019; Li et al., 2022), with genomic changes occurring during both the divergence of its diploid progenitors and the formation of tetraploid and hexaploid forms of wheat and their subsequent evolution. Overall, results suggest that although structural features of the genomes of the diploid species involved in amphiploidization were retained, certain changes also occurred in each round of polyploidization (Jiang and Gill, 1994; Devos et al., 1995; Salina et al., 2006; Marcussen et al., 2014; Jorgensen et al., 2017). The main genomic changes arising during the allopolyploids' formation were connected to the need for joint "accommodation" of the different genomes in one nucleus and cytogenetic diploidization. The wheat B genome and its progenitor in the diploid species *Aegilops speltoides* are rich in tandem repeats (families pSc119.2, pAs1, GAAGAG, etc.), which are involved in the formation of heterochromatin blocks; in contrast, the A genome is quite lacking in tandem repeat families. This imbalance in the content of tandem repeats in the genomes of diploid progenitors may significantly affect meiotic synchronization in the newly developed amphiploid nucleus. Variability in telomeric heterochromatin and centromeric regions has been seen in wheat-rye hybrids and in their progeny (Fu et al., 2013; Bashir et al., 2018). In the case of the bread wheat B genome, Spelt1 repeats specific to the subtelomeric regions of the diploid progenitor *Ae. speltoides* were lost in this process, which was also experimentally confirmed using newly synthesized amphiploids (Salina et al., 2004). Thus, the imbalance between homoeologous genomes A and B in the content of tandem repeats may decrease as a result of partial deletion of tandem repeats specific to one of the parents during the formation of amphiploids.

The situation with mobile elements is somewhat different. In particular, a study of long terminal repeat (LTR) retroelements in wheats and their progenitors has shown that most transposable elements (TEs) proliferated differentially in diploid progenitors and were then inherited by allopolyploids (Bariah et al., 2020a). Presumably, amphiploidization did not affect the proliferation of retrotransposons (Charles et al., 2008; Salina et al., 2011). However, certain changes in the short interspersed nuclear element (SINE) and non-LTR-retrotransposon families in the second generation of synthetic allohexaploid wheat, and their further active methylation in the third generation, have been reported (Ben-David et al., 2013). Thus, the reorganization of mobile elements at stages following the formation of the first amphiploids warrants further analysis.

One of the most important areas in allopolyploid research is determining the activity of the genes once they undergo duplication. Various interacting mechanisms can partially switch off genes *via* methylation or mutations of duplicated genes (Feldman and Levy, 2012). The consequences of such processes are of paramount importance in allopolyploid organisms. Thus, some 45S RNA genes are inactivated by methylation in chromosomes of one of the species constituting the allopolyploid

genome in order to maintain the necessary number of 45S RNA in the cells (Shcherban et al., 2008). The *Ph1* locus emerged in allopolyploid species to prohibit homoeologous pairing in allopolyploid meiosis. Candidates for the *Ph1* gene, as well as the molecular mechanisms underlying its effect, have been discussed (Griffiths et al., 2006; Bhullar et al., 2014; Fan et al., 2021). The formation of an active *Ph1* locus on chromosome 5B also reflects the consequences of interactions between homoeologous gene loci that include activation of genes on one chromosome accompanied by silencing of their orthologues on the homoeologous chromosomes. Chromosome 5B of common wheat is known for the fact that its reorganization throughout the course of evolution led to the emergence of the *Ph* locus, which then contributed to stabilization of the genome of the tetraploid wild emmer wheat *Triticum dicoccoides* and other widespread allopolyploid *Triticum* species. The emergence of *T. dicoccoides*, a wild wheat species of the emmer group, was accompanied by species-specific translocation involving the arms of chromosomes 4AL, 5AL, and 7BS, which was then transmitted to tetraploid and hexaploid species of the emmer group, including cultivated wheat (Naranjo et al., 1987; Devos et al., 1995; Nelson et al., 1995). Chromosome 5B was not involved in large translocation rearrangements during the allopolyploidization.

At the same time, the use of the LTR retrotransposon family in wheat as a genetic marker allowed to identify large-scale genomic rearrangements between the reference chromosome 5B sequences of wild emmer and bread wheat. Six cases of large-scale rearrangements were identified, including 4 cases of long deletions in bread wheat, the introduction of a new DNA fragment, and a single example of copy number variation of a long tandem repeat in chromosome 5B (Bariah et al., 2020b).

A high level of intraspecific polymorphism for chromosome 5B has been noted for the allopolyploid emmer wheat group (*T. dicoccoides*, *Triticum dicoccum*, *Triticum aestivum*). Translocations between the long arm of chromosome 7A and the short arm of chromosome 5B were detected in *T. dicoccoides* and *T. dicoccum* (Badaeva et al., 2015). There is a high level of population-specific and region-specific polymorphisms in the distribution of C-bands on the long arm of chromosome 5B in emmer wheat (Badaeva et al., 2015). A group of cultivars of the hexaploid wheat *T. aestivum* that carry 5BS:7BS and 5BL:7BL translocations resulting from intragenomic rearrangements have been bred and are still successfully grown in Western Europe (Badaeva et al., 2007). Thus, chromosome 5B has definite potential for reorganization, resulting in the emergence of intraspecific polymorphisms in common wheat that are utilized in breeding programs.

One significant consequence of intra- or interspecies chromosome rearrangements is the suppression of recombination in hybrids obtained from crossing forms that differ in such rearrangements. Chromosomal rearrangements that lead to recombination suppression can have a significant impact on genetic divergence and even speciation, but are also important for practical breeding.

In the present work, we evaluated the key points of chromosome 5B reorganization throughout evolution by identifying regions of recombination suppression in interspecific crosses. Sequence

analysis of 5B pseudomolecules of *T. aestivum* ($2n=6x$) and *T. dicoccoides* ($2n=4x$) in the studied regions on 5B allowed us to identify possible determinants of recombination suppression.

MATERIALS AND METHODS

Genetic Linkage Maps

In this work, we used the genetic linkage map of CS and a disomic substitution line in which the CS 5B chromosome is replaced with the *T. dicoccoides* 5B chromosome (CS-5Bdic; Salina et al., 2018). Tetraploid wheat maps developed with populations from a cross of *T. durum* and *T. dicoccoides* were also used. The first map, based on a Svevo \times Zv recombinant inbred line (RIL) population, was constructed and published by Avni et al. (2014), and the second map was based on a Langdon \times Hermon (TZ-2) RIL population, which was genotyping previously (Wang et al., 2018). Both maps were developed using MultiPoint-UltraDense software. The hexaploid wheat map was constructed using the same tools, with the genotyping data of the Chara \times Glenlea doubled haploid population kindly provided by E. Akhunov (Kansas State University) and previously published (Wang et al., 2014). The information about the mapping populations used in this study is summarized in **Supplementary Table S1**.

Comparative Analysis of Genetic Maps for Chromosome 5B

To compare marker order between genetic linkage maps, The Genetic Map Comparator (Holtz et al., 2017) and BioMercator v4 (Sosnowski et al., 2012) were used. To identify chromosomal regions showing considerable recombination suppression, we compared distances between corresponding markers in CS \times CS-5Bdic and the tetraploid and hexaploid wheat maps mentioned above. The regions in CS \times CS-5Bdic where the distances between the markers were reduced by more than 3–3.5 times compared to the tetraploid and hexaploid wheat were selected for the following analysis. Information about markers flanking these regions of suppression is presented in **Supplementary Table S2**.

The recombination regions mapped to deletion bins were performed based on CS \times CS-5Bdic genetic mapping data using SNP, SSR, ISBP (insertion site-based polymorphism) markers and localization SSR, ISBP, and SNP markers in bins (Gadaleta et al., 2014; Salina et al., 2018).

Comparative DNA Analysis for Target Regions of Chromosome 5B

Regions of chromosome 5B where recombination was suppressed according to the results of the genetic analysis were extracted from the reference pseudomolecules of *T. aestivum* cv. CS (IWGSC RefSeq v2.1)¹ and *T. dicoccoides* cv. Zv (WEWSeq v2.0)² based on localization of flanking DNA markers. Genome

assembly of CS line 42 (CS42, PRJNA392179) was retrieved from the NCBI server.³ A sequence similarity search against pseudomolecules of *T. aestivum* varieties Arina (Switzerland), Jagger (United States), Julius (Germany), and Landmark (Canada) was carried out on the IPK server.⁴ Multiple sequence alignment was carried out using MAFFT 7.311 software (Kato and Standley, 2013).

Tandem repeats were identified with Tandem Repeats Finder software (Benson, 1999) using the default settings. All nested repeats were grouped. Groups separated by less than 100 bp were merged into the same cluster. The obtained clusters containing tandemly repeated DNA sequences were filtered by length \geq 1 kb. For each cluster containing repeats of the pSc119.2 satDNA family, the sequences of the repeated motif were aligned, and a consensus sequence was obtained.

To calculate the map coverage, local DNA sequence alignments were generated using NUCmer, included in MUMmer (Kurtz et al., 2004) with the following parameters: “-mum -b 150 -c 600 -g 30 -l 24,” and local alignments with \geq 80% similarity to the reference sequence were studied.

DNA-sequence alignments for the collinearity analysis were obtained with BLASTn (Zhang et al., 2000; Camacho et al., 2009) and filtered by length (\geq 3,000 bp) and identity (\geq 75%). The assemblies containing non-overlapping fragments ordered on the same strand were calculated, and the most representative assembly was taken for downstream analysis.

TEs were recognized in RepeatModeler (v2.0.2, Smit et al., 2008) and RepeatMasker (v4.1.2-p1, Smit et al., 2013) using the RepBase database (RepBaseRepeatMaskerEdition- 20181026, Bao et al., 2015).

Dot plots of 5BS_RS and 5BL_RS regions were calculated in Gepard (v1.40, Krumsiek et al., 2007) using a word length of 30 bp. Circular data visualization was performed in Circos (Krzywinski et al., 2009).

DNA Shape Analysis

Modeling of predicted DNA helical paths and calculations of DNA shape features were performed according to previously published methods (Muterko, 2017). The curvature distribution was analyzed in a sliding window of 40 bp with a 1-bp step. DFT was performed with a sampling frequency of 1,024. Three-dimensional DNA structures were visualized by the PyMOL Molecular Graphics System, version 1.7.2 Schrödinger, LLC.

Evaluation of LTR-Retrotransposon Divergence

Putative LTR retrotransposons with LTR–LTR divergence \geq 60% were predicted with LTRharvest (Ellinghaus et al., 2008), and elements with a LTR length \geq 180 bp were selected. LTR of each element was aligned, and free-end gaps of 150-bp fragments were used in computing divergence. LTR TE divergence was calculated from LTR–LTR mismatches, assuming InDels as single evolutionary events. The mean LTR–LTR

¹<https://wheat-urgi.versailles.inra.fr/Seq-Repository>

²<http://weewseq.wixsite.com/consortium>

³<ftp://ftp.ncbi.nlm.nih.gov/genomes>

⁴http://webblast.ipk-gatersleben.de/wheat_ten_genomes

divergence (age) is the local maximum of the Gamma distribution, with the shape and scale parameters obtained from the negative binomial distribution fitted to the LTR–LTR mismatch distribution. Furthermore, to evaluate the fractions of LTR–LTR divergences (ages) and improve the fit to the LTR–LTR mismatch distribution, the mixture of Poisson distributions was resolved with the use of the expectation–maximization algorithm. Goodness of fit was evaluated using the coefficient of determination (r^2). LTR TE carrying a number of mismatches equal to or less than the mean LTR divergence was considered young, while all others were considered old. The over- and under-representation of young LTR elements was evaluated in a sliding window of 10 elements with a 1 element step, using hypergeometric test.

RESULTS

Comparative Analysis of Chromosome 5B Genetic Maps

To identify chromosome 5B rearrangements that have occurred during evolution, we compared chromosome 5B recombination frequency in hybridizations of tetraploid (*T. durum* × *T. dicoccoides*) and hexaploid (*T. aestivum* × *T. aestivum*) wheat and in hybrids of hexaploid wheat cv. Chinese Spring (CS) and a disomic substitution line in which CS chromosome 5B has been replaced with *T. dicoccoides* chromosome 5B (CS-5Bdic). Only populations genotyped with Illumina Infinium arrays were used in this work (Supplementary Table S1). All of the maps were constructed using MultiPoint-UltraDense software.

A comparison of the maps demonstrated that the used markers were arranged in the same order (Figure 1). Although the total number of markers in the Chara × Glenlea and Svevo × Zavitan (Zv) maps compared to that of CS × CS-5Bdic was larger, the number of skeletal markers (markers representing groups of cosegregating markers occupying separate positions on the constructed linkage map) varied to a lesser degree. A comparative analysis of the length of the 5B maps and the distances between corresponding markers in the studied tetraploid and hexaploid populations relative to CS × CS-5Bdic showed that the latter was 1.5- to 1.7-fold shorter.

The regions with suppressed recombination are evident on the 5B map of CS × CS-5Bdic (Figure 1). In these regions, the distances between the markers are reduced by more than 3–3.5 times, as compared to the maps constructed for tetraploid (Svevo × Zavitan and Langdon × Hermon) and hexaploid (Chara × Glenlea) wheats and exceed the whole map reduction coefficient (1.5–1.7) by more than twofold. Two regions of recombination suppression on the short arms (between markers BS00009810 and BS00022336) and long arms (between markers Ra_c10633_2155 and BS00087043) are observed in all comparisons of intra- versus interspecies chromosome maps (all markers cosegregating with these basic markers are summarized in Supplementary Table S2).

Recombination suppression is presumably associated with the fact that hybridization in the case of CS × CS-5Bdic involved chromosomes 5B of common wheat and the wild species

Triticum dicoccoides, which may have discrepancies along the 5B sequence. In addition, heterochromatin and possibly associated epigenetic regulation may be involved in recombination suppression (Termolino et al., 2016) and divergence of sequences of related wheat species. For a better understanding of the potential causes of recombination suppression in specific regions of chromosome 5B, the 5B pseudomolecules of *Triticum* species need to be compared.

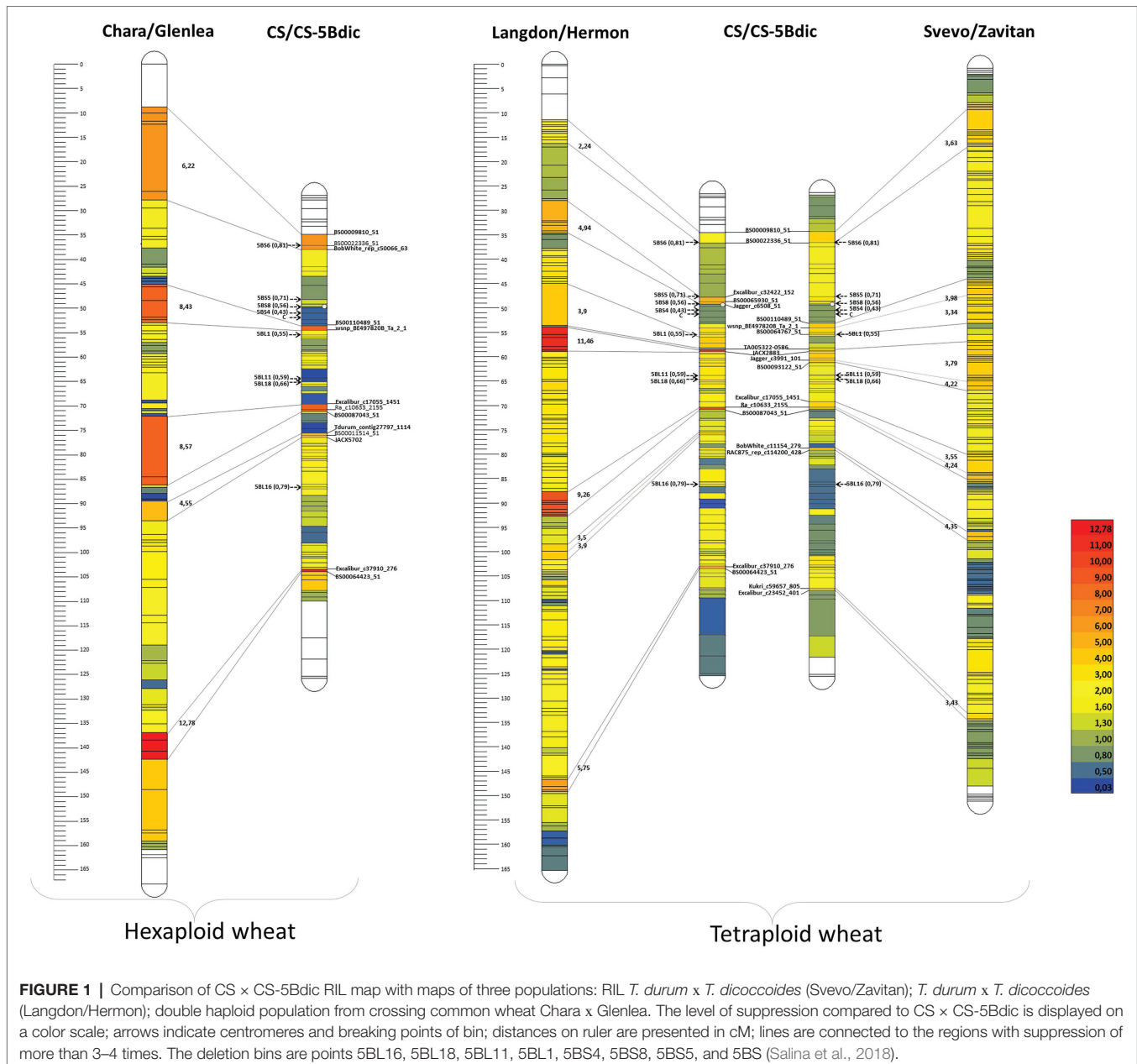
Comparative Analysis of *Triticum aestivum* and *Triticum dicoccoides* Chromosome 5B Regions Associated With Recombination Suppression in Interspecific Cross

To define the possible determinants of recombination suppression, we analyzed the primary structure of the two chromosome 5B regions of *T. aestivum* and *T. dicoccoides*: (a) 5BS_RS, between markers BS00009810 (same position as BS00083715: 10489937–10,490,037 bp (CS chr5B pseudomolecule), 9,941,766–9,941,866 bp (Zv chr5B pseudomolecule)) and BS00022336 (15585936–15,586,036 bp (CS), 17,514,968–17,515,068 (Zv)); and (b) 5BL_RS, between markers Ra_c10633_2155 [same position as Tdurum_contig 25513_123: 549917016–549,917,116 bp (CS), 561,877,394–561,877,494 bp (Zv)] and BS00087043 [553533144–553,533,244 bp (CS), 565,677,935–565,678,035 bp (Zv)]. The regions of recombination suppression on the short and long arms, with lengths of 5 Mb and 3.6 Mb, respectively, were mined from the 5B pseudomolecule of CS and compared to the homoeologous regions (7.6 and 3.8 Mb, respectively) of the 5B pseudomolecule of *T. dicoccoides* Zv.

5BS_RS Regions

Comparison of the 5BS_RS regions of CS and Zv revealed two areas of tandem repeat sequences and several areas that had either very low similarity or did not map to CS chromosomes at all (Figure 2A). The latter included different chromosomal rearrangements, with a predominance of insertions/deletions (InDels; Figure 2B). This also followed from the difference in length of the 5BS_RS region on pseudomolecules of CS and Zv. In particular, the 5BS_RS of Zv includes around 2 Mb of the DNA sequences, which did not map to the collinear segment of the CS chromosome, suggesting that they are the major contributors to recombination suppression in the 5BS_RS region through loss of similarity.

Extended arrays of tandemly repeated DNA sequences (satellite DNA [satDNA]) were localized to the beginning of the 5BS_RS region. A comparison of fractions containing clusters of tandem repeats with repetitive elements of different lengths, as well as an evaluation of the total length of repetitive sequences in each fraction, showed that repeated sequences within 5BS_RS are mainly represented by clusters of a 118-bp tandem unit or its dimer (235–236 bp) or trimer (353–354 bp). Primary structure analysis has shown that this highly repetitive sequence belongs to the pSc119.2 satDNA family, which is widely distributed within the *Triticeae*, as well as in some *Aveneae* species, and forms a large and evolutionarily old component of the genome (Contento et al., 2005). For a more detailed investigation of

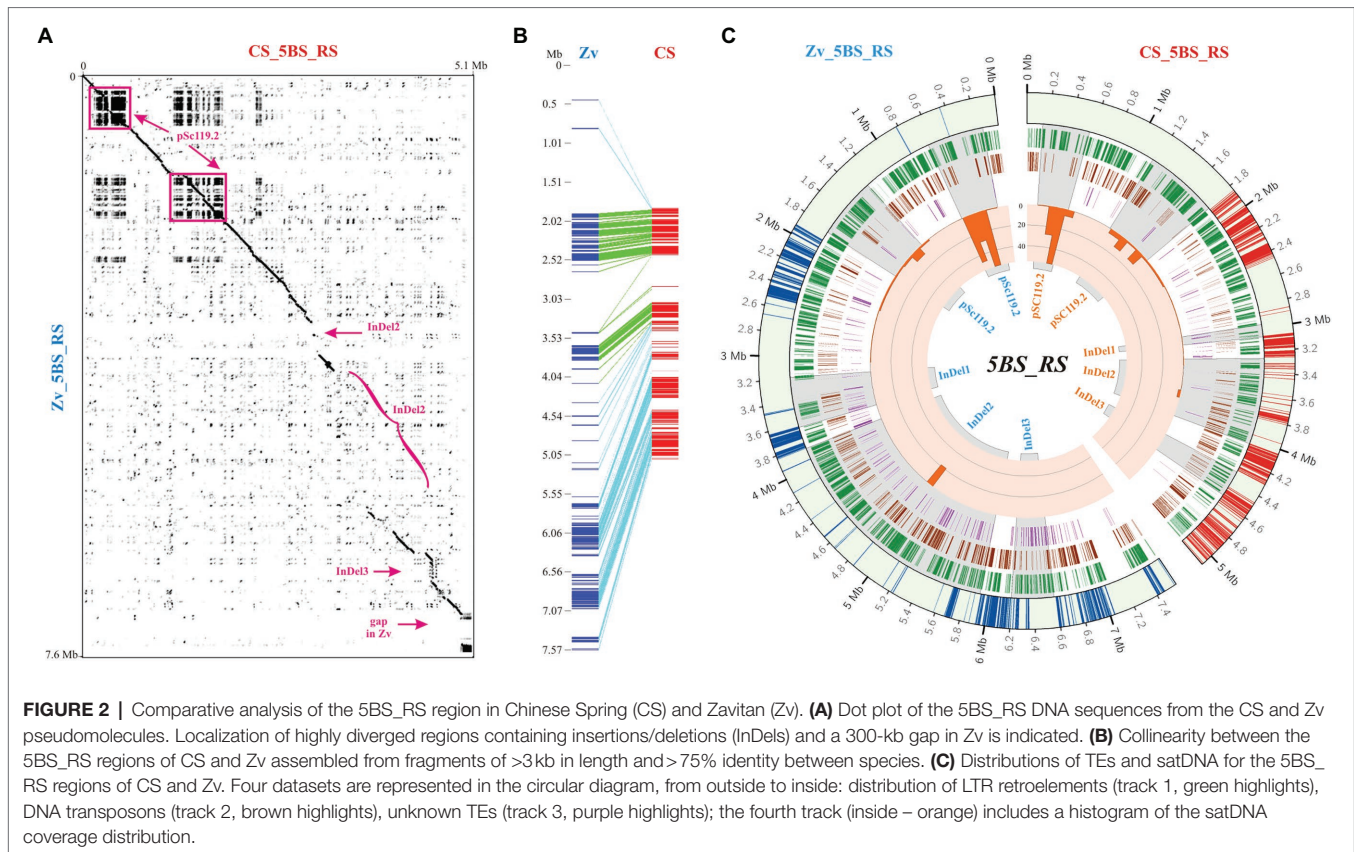


distribution and genetic diversity, the arrays of 118-bp tandem repeats of the pSc119.2 family were clustered and filtered, as described in Experimental Procedures, and a consensus sequence was deduced for each cluster. The analysis of pSc119.2 distribution identified three large pools (sized ~500, 600, and 90 kb) located on half of the 5BS_RS region (Figures 2A,C; Supplementary Figure S1). A total of 113 clusters of pSc119.2 were identified for CS and 82 for *Zv* (Supplementary Figure S1), even though both provided similar coverage (Figure 2C).

With the exclusion of single clusters, the 118-bp motif was highly conserved in each pool (identity of 86–93%) and between pools (94%) for both CS and *Zv*. Similarity between consensus DNA sequences of pSc119.2 pools of CS and *Zv* was 95–97%. Multiple sequence alignment found that pool 2 in CS contains

an array of inverted clusters (Supplementary Figure S1). This is an important feature distinguishing pool 2 of the pSc119.2 clusters in CS and *Zv*. The total length of this inverted region was more than 100 kb. Furthermore, the chromosome region within pool 3 in both CS and *Zv* was inverted compared to pool 1 (as well as pool 2 of *Zv* and a part of pool 2 from CS). Although the presence of inverted repeats can lead to their instability in the genome, high-level identity in both the position and primary structure was revealed for the studied pSc119.2 repeats between wild tetraploid and cultivar hexaploid wheat (Supplementary Figures S1, S2).

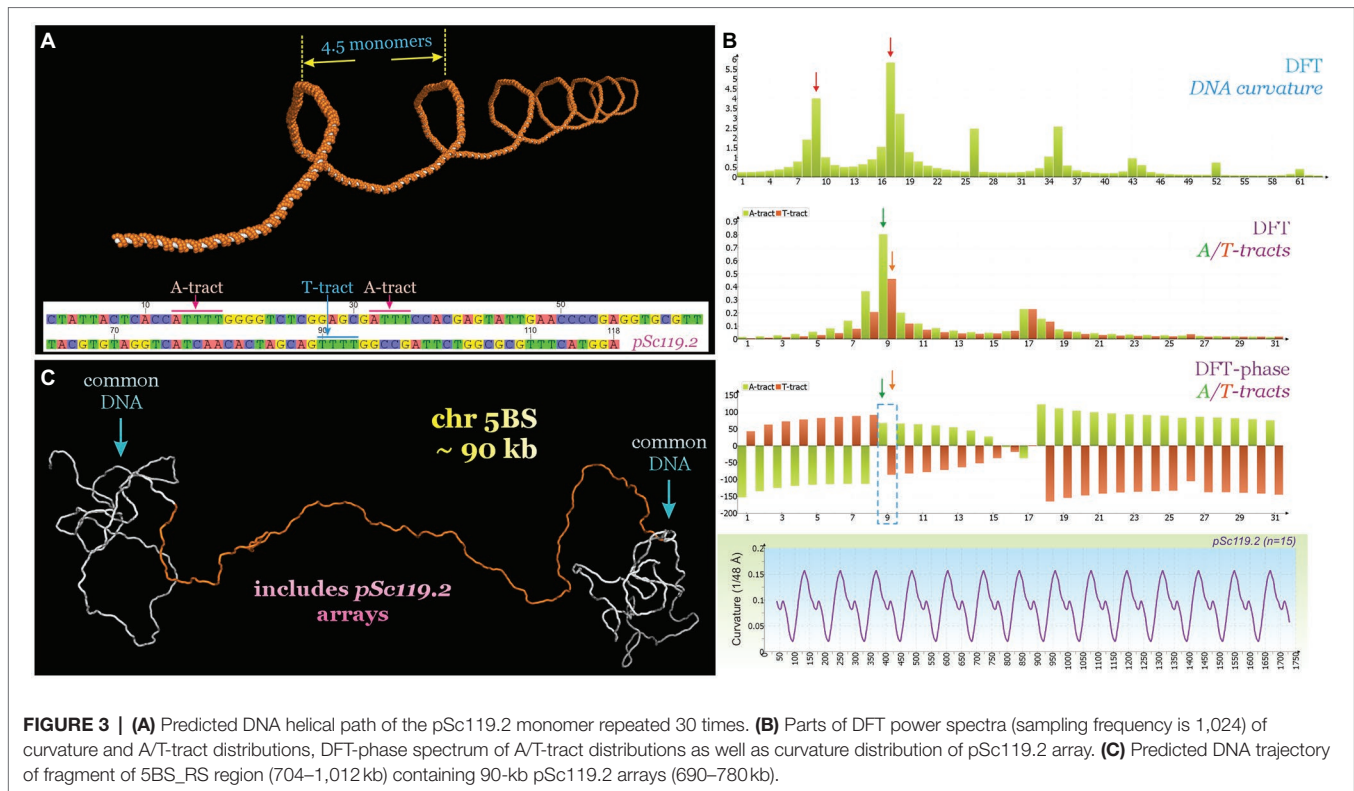
Such stability of the tandem repeat areas is presumably determined by the tertiary structure of the DNA. The DNA sequence of pSc119.2 includes three A-tracts ($AnTn$, $n > 3$).



A-tracts are known as key determinants of the local bends and global curvature of the DNA molecule (Marini et al., 1982; Koo et al., 1986). This global DNA bending can increase significantly when A-tracts are repeated in phase with the helical screw (Marini et al., 1982; Koo et al., 1986; Koo and Crothers, 1988), i.e., are localized in a molecule such that the intrinsic curvature progressively increases in a certain direction. The tandemly repeated sequence containing A-tracts can easily satisfy such conditions. To evaluate the contribution of this satDNA to molecular bending, the predicted DNA helical path for the 118-bp motif of pSc119.2 was calculated and the distribution of the DNA curvature was analyzed (**Figure 3A**). We found that assuming a DNA helical period of 10 bp, the 118–120-bp periodicity of the pSc119.2 motif is in phase with the DNA helical screw. Two consecutive A-tracts within the analyzed pSc119.2 motif also alternated in phase with the DNA helix (20-bp distance), leading to an increase in the macroscopic DNA curvature. Since A-tracts and T-tracts produce bends in opposite directions, we suggest that those alternating at the same frequency and in phase with the helical screw should mutually compensate for each other. However, analysis of the discrete Fourier transform (DFT)-phase spectra of A- and T-tract alternations showed a phase shift of up to 150° (**Figure 3B**). Thus, T-tracts alternating out of phase with A-tracts, but producing a bend in the opposite direction, will increase the macroscopic curvature of the DNA molecule. This prediction also follows

from the DFT power spectrum of the curvature distribution, where the two detected major frequency bins were 9 and 17, corresponding to 118- and 60-bp periodicity, respectively. One of the important features of regularity is the formation of extended homogeneous rigid structures. In the considered case, this structure is almost a straight helix with a 4.5-monomer period. To demonstrate this, the three-dimensional DNA helical path was inferred for a 90-kb fragment of pSc119.2 arrays (pool 1) flanking the common DNA sequences (**Figure 3C**). The assumption of involving of satDNA in heterochromatin formation was supported with the use of publicly available for CS the distribution of the chromatin states, determined with a multivariate hidden Markov model for different epigenetic marks (Li et al., 2019). It was shown that the pSc119.2 repeat-containing regions within 5BS_RS are associated with the E13 chromatin state, whose signature is the modified histone H3K9me2 (**Supplementary Figure S3**), found in constitutive heterochromatin of many eukaryotes. Furthermore, the DNase I profiling (Li et al., 2019) also indicates the low chromatin accessibility within these regions of 5BS_RS.

Although the presence of pSc119.2 tandem repeats providing denser DNA packaging in heterochromatin areas does not explain the observed suppression of recombination, the presence of constitutive heterochromatin can enhance the influence of adjacent regions that are associated with the recombination suppression.



As noted above, a region with low similarity between the homologues can significantly contribute to recombination suppression in heterozygotes. Comparison of the 5BS_RS regions in CS and Zv revealed three large, highly diverged consecutive segments containing InDel-1: ~0.11 and 0.46 Mb, InDel-2: 0.77 and 2.1 Mb, and InDel-3: 0.25 and 0.38 Mb on CS and Zv pseudomolecules, respectively (Figure 2A). Detailed analysis of these regions characterized predominant insertions of TEs (Figure 2C). InDel-1 contained a 0.39-Mb insertion in CS and 0.28-Mb insertion in Zv. Within InDel-2, a total insertion length of 1.6 Mb was detected for Zv and 0.4 Mb for CS. InDel-3 included insertion of repeated DNA sequences in CS, most of which were nested. Zv also contained repeats in this region; however, they were not similar to those in CS and differed in lengths.

The mobile genetic elements identified within the first and second InDel regions were predominantly the TEs of the LTR order of the class retrotransposons, superfamily Gypsy, and Copia (Supplementary Figure S2). However, it is not clear when during the evolution of polyploid wheat these chromosomal rearrangements occurred. To shed light on this problem, the LTR-retrotransposon age distribution was analyzed along the 5BS_RS chromosome region, and the over- and under-representation of “young” LTR elements was evaluated within the segments where rearrangements in CS and Zv are localized (Figure 4). Analysis of LTR–LTR mismatch distribution in CS found the prevalent fraction of LTR divergence to be five mismatches (Figure 4A). Thus, the elements carrying five or less LTR–LTR mismatches were considered “young,” whereas

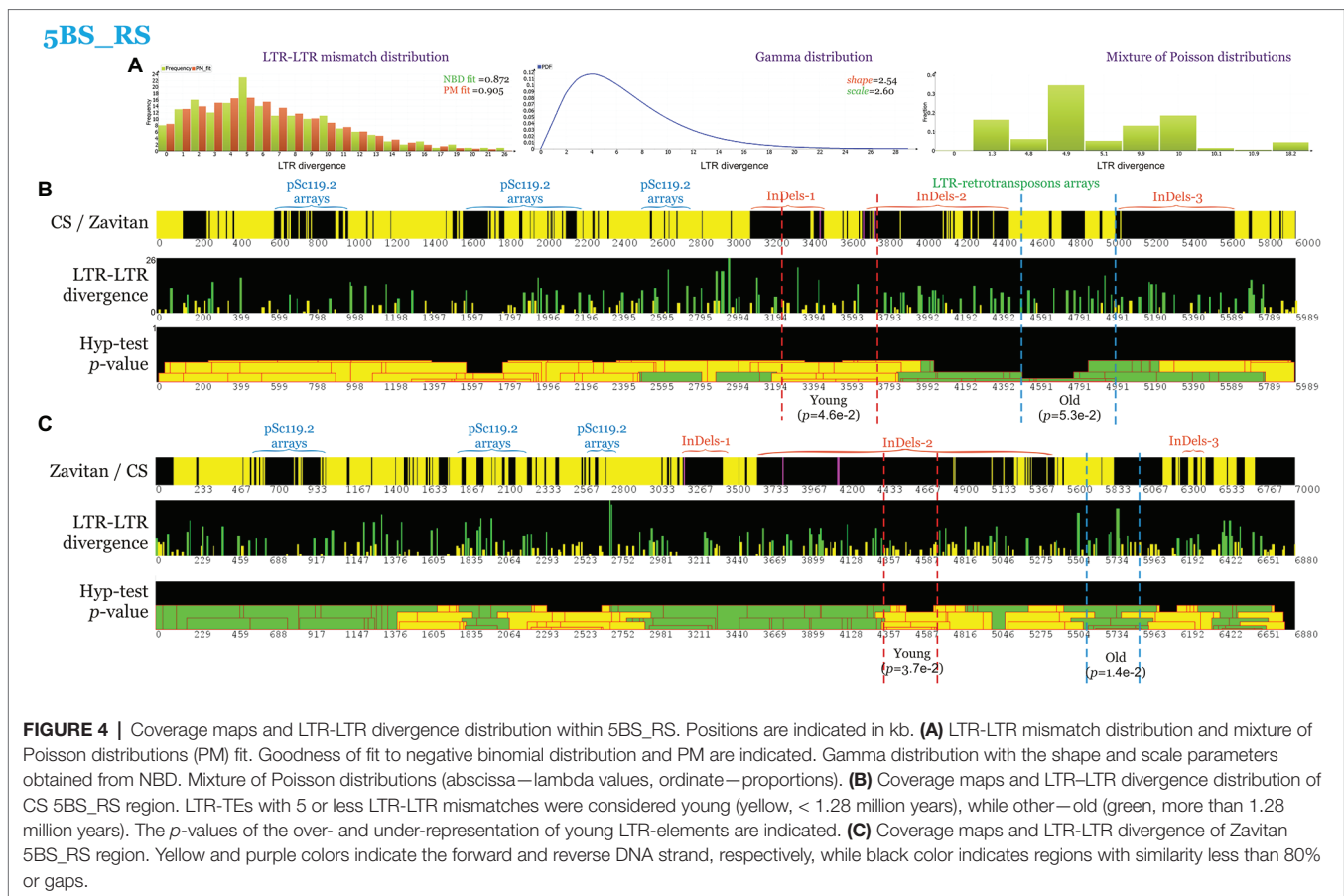
longer ones were regarded as “old.” Assuming a substitution rate of 1.3×10^{-8} mutations per site per year for the plant LTR retrotransposons (Ma and Bennetzen, 2004), this age threshold corresponds to insertion of the LTR retrotransposons before and after 1.3 million years ago for “young” and “old” retrotransposons, respectively.

Statistically significant over-representation of “young” LTR elements ($p=0.046$ and 0.037 , respectively) was shown for two regions (300 and 500 kb), including InDel-1 in CS and the proximal part of InDel-2 in Zv (Figure 4B). The under-representation of young LTR elements was shown for a 1-Mb region between InDel-2 and InDel-3 in CS and Zv (Figures 4B,C).

Thus, given a wild emmer emergence time of about 0.5 MYA (Feldman and Levy, 2012) and according to the estimated time of movement of mobile elements, the rearrangements observed in the studied regions of chromosome 5BS took place in diploid progenitors of allopolyploid wheat and accompanied the processes of evolution and the formation of allopolyploids.

5BL_RS Regions

Despite the apparent similarity of the 5BL_RS segment between the CS and Zv genome assemblies, our analysis revealed several highly diverged regions (HDRs, Figure 5A). This loss of similarity was mainly caused by insertions and deletions (InDels) as it followed from the different lengths of 5BL_RS in CS and Zv pseudomolecules (3,617 and 3,802 Mb, respectively) and displacements in the distribution of collinear segments (Figure 4B). In particular, from 0.3 to 2 Mb (Figure 5B), the

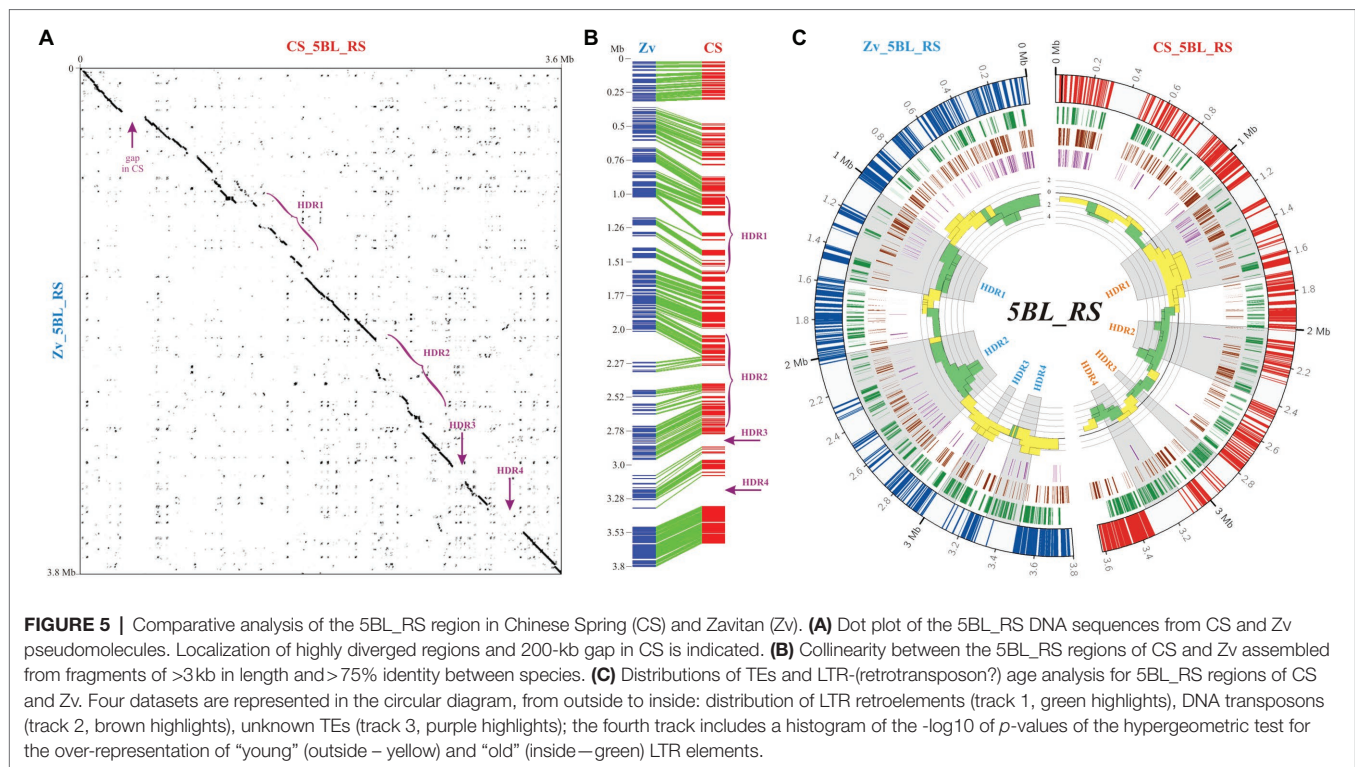


displacement of collinear segments implied insertions in CS (or deletions in *Zv*), and from 2 Mb on, insertions in *Zv* (or deletions in CS). DNA sequence analysis of these regions revealed that in the first case, the initial offset of collinear segments is caused by the prolonged gap in CS (0.3–0.48 Mb), while all following displacements in this area, beginning from 1.16 Mb, are caused by insertions of additional sequences in the CS genome (Figure 5B). We conclude that in the second case, insertions of additional DNA sequences in the *Zv* genome beginning from a position of 2 Mb occurred at least twice, leading to the displacement of collinear segments within the alignment in the opposite direction.

Overall, two major and two minor segments with significant loss of similarity could be recognized within the 5BL_RS region (HDR1–4, Figure 5A). The first (HDR1) is localized proximal to the Tdurum_contig25513_123 marker (1–1.57 Mb) and includes around 550 kb on the 5B pseudomolecules of CS and *Zv*. The second (HDR2) begins near the center of 5BL_RS and expands to the BS00087043_51 marker (2–2.7 Mb). It encompasses almost 350 kb in CS and is twice as long in *Zv* (close to 700 kb), assuming insertions with a total length of 350 kb within the *Zv* genome. Although both of these segments are enriched with TEs, different types of TEs are not equally represented. In particular, in HDR1, the DNA transposons are prevalent over LTR retrotransposons (Figure 5C), and among them, elements of the DNA/CMC-EnSpm family are represented

almost exclusively. Two short (around 0.1 and 0.2 Mb) distal segments (HDR3 and HDR4, respectively) with a complete absence of similarity were enriched in LTR retrotransposons (Figure 5C).

To investigate the possible reasons for divergence of HDRs within 5BL_RS, analysis of the LTR age distribution was carried out exactly as described above for 5BS_RS. The results obtained for *Zv* showed that in the HDRs of 5BL_RS, where LTR elements are few and DNA transposons are prevalent (HDR1 and HDR2), there is a statistically significant over-representation of the “old” LTR retrotransposons, whereas in the HDRs where “young” LTR elements are over-represented (HDR3 and HDR4), the DNA transposons are less frequent (Figure 5C, left side). Except for HDR2, opposite results were obtained for CS (Figure 5C, right side). In fact, this opposition is in good accordance with the occurrence of HDRs, which allows us to explain the suppression of recombination on 5BL_RS for homologous chromosomes of the considered wheat species. In particular, the over-representation of the “young” retroelements in HDR1 of CS agrees with the insertion of additional DNA sequences in this region (rather than their selective deletion in *Zv*) and hence contributes to the lowering of similarity with a collinear chromosomal segment in *Zv* for which the “old” LTR retrotransposons are over-represented on this region. Similarly, over-representation of the “young” LTR elements in HDR3 and HDR4 of *Zv* corresponds to over-representation



of the “old” LTR elements in the collinear regions of CS (Figure 5C).

Based on these data, we suggest that the loss of similarity, which can lead to recombination suppression in the 5BL_RS region, is caused by chromosomal rearrangements, driven by the activity of mobile genetic elements (both DNA transposons and LTR retrotransposons) and their divergence during evolution.

It should be noted that in CS, HDR2 includes a scaffold of around 400 kb (scaffold2564–2), which is flanked by a long span of simple tandem repeats (STRs) of the same type: (TCTTCT) $_n$, $n > 500$. Since flanked STRs are known to cause problems for correct scaffold orientation during *de novo* genome assembly, the orientation of this scaffold on pseudomolecule 5B of CS (IWGSC assembly) until the appearance of version 1.2 was questionable. The direction of this genomic region in common wheat cultivars, such as *T. aestivum* varieties Arina, Jagger, Julius, and Landmark, is similar to that in Zv (Walkowiak et al., 2020). However, in all of these cases, the scaffolds break at the flanked satDNA. On the other hand, this inversion is absent in the 5B pseudomolecule of the CS42 genome assembly (Zimin et al., 2017), which, together with optical maps, was used to improve the IWGSC CS assembly during the release of RefSeq v2.1 (Zhu et al., 2021). Although these data most likely indicate the absence of an inversion within HDR2, we cannot rule out its presence on chromosomes 5B of CS and CS-5Bdic used in the cross in the present study. Inversion of a scaffold would be a good explanation for the recombination suppression on the 5BL_RS region, although we do not yet have enough evidence to confirm this assumption.

DISCUSSION

Recombination suppression is an important phenomenon involved in processes such as the formation of sex chromosomes, evolution of hybrid zones, and speciation (Mank, 2012; Charlesworth, 2016). In plant breeding, the transfer of agronomically important genes in non-recombinant linkage groups occurs when they are localized on alien translocations (Wingen et al., 2017). There are diverse causes of recombination suppression; the most commonly discussed are InDels, inversions, heterochromatin regions, and epigenetic mechanisms (Hammarlund et al., 2005; Ziolkowski et al., 2015; Guo et al., 2016).

In our case, recombination-suppression regions on wheat chromosome 5B were detected in a cross between CS and CS-5Bdic (genotype CS with 5B substitution by its homologue from tetraploid *T. dicoccoides*). It is currently not possible to use information on the 5Bdic chromosome donor to establish the possible causes of recombination suppression, although this substitution line has been actively used in various genetic studies (Faris et al., 2000; Lu et al., 2006; Salina et al., 2018). At the same time, much experimental material has been accumulated on intraspecific polymorphism of tetraploid species of the emmer group, which may be useful for identifying the causes of recombination suppression observed in CS \times CS-5Bdic crosses. Thus, a comparative chromosome analysis by C-banding in 446 accessions of *T. dicoccum* and 105 accessions of *T. dicoccoides* from seven countries, covering its whole present-day natural distribution range, revealed population-specific and region-specific

polymorphisms (Badaeva et al., 2015). The following intraspecific polymorphisms were observed for chromosome 5B in *T. dicoccum* and *T. dicoccoides*: (1) a 7A:5B translocation involving the short arm of chromosome 5B (T5BL.5BS-7AL); (2) differences in the distribution of diagnostic C-bands on the long chromosome arm. It should be noted that the 5BS_RS and 5BL_RS regions identified in our study also fall into the region of high intraspecific polymorphism. Thus, a more detailed analysis of the T5BL.5BS-7AL breakpoint on chromosome 5BS showed that it is located distal to the minor pSc119.2 site and the 5S rRNA gene locus, whereas the major pSc119.2 site was relocated to chromosome 7A (Dedkova et al., 2009; Badaeva et al., 2015). A comparison of these findings with our data on the location of the recombination-suppression region relative to the distribution of repeats on chromosome 5BS (**Supplementary Figure S2**) suggests that the region of the chromosome that we labeled 5BS_RS is involved in intraspecific rearrangement in *T. dicoccum* and *T. dicoccoides*. We cannot rule out the possibility that the breakpoint runs in close proximity to the tandem pSc119.2 repeats identified in this region (**Figure 2**).

Intraspecific polymorphism on the long arm of chromosome 5B was revealed by variation in the location of C bands, the total number of which on 5BL is eight (Badaeva et al., 2015). In *T. dicoccum* accessions, the region attributable to intraspecific polymorphism covers the interstitial and distal region of 5BL, whereas for *T. dicoccoides* accessions, this region is significantly wider and covers more than 70% of 5BL. The 5BL_RS region that we studied, located in the 5BL18 bin (**Figure 1**), also falls within the region of intraspecific polymorphism of *T. dicoccum* and *T. dicoccoides* species. In addition, as previously shown, alien translocations often occur in the 5BL18 bin of chromosome 5B during interspecific hybridization (Timonova et al., 2013). Taken together, this indicates the relative instability of the regions marked as 5BS_RS and 5BL_RS during both plant adaptation to different agroclimatic zones and interspecific hybridization.

The reasons for such instability, leading to intraspecific differences and subsequently, to the suppression of recombination during interspecific hybridization, can be clarified by a detailed analysis of the structure of the chromosome regions involved in these processes. As possible factors of suppressed recombination in 5BS_RS, 5BL_RS or both, one can consider heterochromatinization associated with the presence of extended tracts of tandem repeats (119.2), large insertions, and highly divergent regions enriched with TEs.

Role of Tandem Repeats and TEs in Recombination Suppression

Tandem repeats and TEs are quite often associated with the heterochromatin regions of chromosomes. The possible participation of heterochromatin regions in recombination suppression has been discussed previously (Charlesworth et al., 1986; Kagawa et al., 2002). There is evidence that heterochromatin suppression of crossing-over is controlled in the chromatin structure (Westphal and Reuter, 2002).

The region of recombination suppression on the short arm of chromosome 5B is colocalized with a heterochromatin region at the boundary of two distal bins (5BS6–5BS5) according to physical and genetic mapping (**Figure 1**; Salina et al., 2018). Analysis of 5BS_RS revealed that nearly 2.8 Mb of this region is covered by pSc119.2 tandem repeats. The other part of the region is covered by TEs and displays very low between-species similarity due to extended InDels.

One of the most common features of satDNA, which is shared among many eukaryotic organisms, is its intrinsic curvature. In previous studies, it has been repeatedly noted that clusters of tandemly repeated nucleotide sequences probably play a key role in DNA stabilization, as well as in protein packaging and higher-order chromatin condensation (Fitzgerald et al., 1994; Tolstorukov et al., 2005; Mrazek, 2010). It is assumed that sequence-directed bends produced by the repeated A-tracts in the satDNA in phase with the helix facilitate the assembly of DNA into nucleosomes and therefore, represent an essential structural basis for heterochromatin condensation (Fitzgerald et al., 1994; Rohs et al., 2009; Segal and Widom, 2009; Struhl and Segal, 2013). The results obtained in the present study agree with the results of Vershinin and Heslop-Harrison (1998) indicating the phasing of nucleosomes along arrays of the pSc119.2 satDNA. Furthermore, the pSc119.2 repeat is known to form a major part of heterochromatin regions in the B genome of wheat (Mukai et al., 1993; Contento et al., 2005). In the present study, we demonstrated that the localization of A-tracts within the pSc119.2 motif leads to sequence-directed bending of the DNA molecule. We suggest that the observed periodicity of the DNA bend can facilitate molecular packing and may be a structural basis for local heterochromatin condensation in this region. The high degree of similarity between different clusters of pSc119.2 found herein supports the existing opinion that there is selective pressure acting on the maintenance of curved DNA (Charlesworth et al., 1994). Bearing in mind the overlapping location of the studied region with the heterochromatin stretch, we argue that large arrays of pSc119.2 satDNA within the 5BS_RS locus can lead to the formation of local islands of interstitial heterochromatin within the euchromatic chromosomal arm, contributing to recombination suppression in this and adjacent regions.

There are also other mechanisms supporting heterochromatin structure. Thus, direct or inverted repeats can lead to the creation of a heterochromatic state (Bennetzen et al., 1994; Dorer and Henikoff, 1994; Tikhonov et al., 1999). The methylation of histone H3mK9 (Jenuwein and Allis, 2001; Martienssen and Colot, 2001; Vermaak et al., 2003), DNA methylation (Martienssen and Colot, 2001; Law and Jacobsen, 2010) and histone variant H2A.W (Yelagandula et al., 2014) play an important function in regulating heterochromatin and act in synergy to maintain transposon silencing. In particular, in the present study it was shown that chromatin of the pSc119.2 repeat-containing regions within 5BS_RS is marked with histone H3K9me2. This last is the signature of repeat-rich constitutive heterochromatin for many eukaryotes (Naumann et al., 2005; Nicetto and Zaret, 2019) and is an evolutionarily conserved, specific mark of nuclear peripheral heterochromatin (Polshko et al., 2019).

Furthermore, it has been shown that heterochromatin in *Arabidopsis* is determined by TEs and related tandem repeats under the control of the chromatin-remodeling ATPase DDM1 (Lippman et al., 2004). The enrichment of TE sequences in heterochromatin is widespread among higher eukaryotes. In plant genomes, even relatively small retrotransposon blocks are methylated (Bennetzen et al., 1994; SanMiguel et al., 1996; Tikhonov et al., 1999) and can be considered heterochromatic regions. Moreover, small tandem expansions, such as adjacent duplication of a TE from a monomer to a dimer, or an increase in the copy number within a simple sequence repeat, can lead to altered epigenetic status within and around these sequences, and these may also be considered heterochromatin (Dorer and Henikoff, 1994).

In the present study, during the analysis of the LTR–LTR divergence, we found indirect evidence for heterochromatin formation in the 5BS_RS region carrying extended TE arrays. It was shown that local heterochromatin islands, determined by the satDNA and TE arrays, as well as dissimilarity caused by the large InDels (chromosome rearrangements), are likely key determinants of recombination suppression in the 5BS_RS region.

Our results allow suggesting that mobile genetic elements (both DNA transposons and LTR retrotransposons) played a key role in the loss of similarity in the region labeled 5BL_RS. Overall, two major and two minor segments with significant loss of similarity can be recognized within the 5BL_RS region (HDR1–4, **Figure 3A**). The HDRs are enriched with TEs, but each region is characterized by a different history of integration and movement of LTR retroelements, and a different distribution ratio of DNA transposons and LTR retrotransposons. The results are consistent with recent studies indicating a negative correlation between regional genomic variation in TE density and frequency of recombination. The strength of this correlation varies depending on the TE type; TE insertion polymorphism may be an important factor determining intraspecific variations in recombination rate in the surrounding genomic regions, with all of the consequences of this event (Kent et al., 2017).

CONCLUSION

Chromosomal rearrangements that lead to recombination suppression can have a significant impact on speciation, and they are also important for breeding. In our case, recombination-suppression regions on wheat chromosome 5B (5BS_RS and 5BL_RS) were detected in a cross between CS and CS-5Bdic (genotype CS with 5B substitution by its homologue from tetraploid *T. dicoccoides*). It was shown that the loss of similarity, which can lead to suppression of recombination in the studied regions, is caused by chromosomal rearrangements, driven by the activity of mobile genetic elements (both DNA transposons and long terminal repeat retrotransposons) and their divergence during evolution. Also, the local heterochromatin islands determined by the satellite DNA (pSc119.2), as well as the dissimilarity caused by large

insertions/deletions (chromosome rearrangements) between *aestivum/dicoccoides*, are likely the key determinants of recombination suppression in the region.

It was noted that the regions marked as 5BS_RS and 5BL_RS are associated with chromosomal rearrangements identified earlier by C-banding analysis of intraspecific polymorphism of tetraploid emmer wheat. Taken together, this indicates the relative instability of the studied regions during both plant adaptation to different agroclimatic zones and interspecific hybridization.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

ES, AM, and AKo: conceptualization. AKi: formal analysis and investigation (comparative mapping data). ZL: formal analysis and investigation (Langdon × Hermon RIL mapping data). AM: formal analysis and investigation (RS regions of pseudomolecules, satDNA, TEs, large InDels, and DNA shape analyses). ES: project administration and supervision. ES, AM, and AKi: writing-original draft preparation. ES, AM, ZL, and AKo: writing, review, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

Genetic linkage maps and comparative chromosome analysis were supported by RSF (Russian Science Foundation) project no. 21-76-30003. Analysis of the 5BS_RS and 5BL_RS regions was funded by the Kurchatov Genomics Center of IC&G (075-15-2019-1662). DNA shape analysis was performed within the budgetary project FWNR-2022-0017.

ACKNOWLEDGMENTS

We thank Camille Vainstein for English editing of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.884632/full#supplementary-material>

Supplementary Figure S1 | (A) Scatter graph of distribution of number of pSc119.2 clusters with length not less than 1 kb, within the 5BS_RS region extracted from Cs and Zv pseudomolecules and divided on 100-kb segment. **(B)** Scatter graph of coverage by the sequences of the pSc119.2 clusters evaluated for each 100-kb segment. Arrows indicate direction (forward/inverse) of the pSc119.2 satDNA arrays.

Supplementary Figure S2 | Distribution of the pSc119.2 arrays on pseudomolecule of 5BS. Localization of telomere, centromere, and the recombination suppression region are indicated. Forward and reverse orientation of pSc119.2 is indicated by the blue and red colors, respectively. The pSc119.2 was mapped and plotted in MUMmer 3.0 (Kurtz et al., 2004).

Supplementary Figure S3 | Distribution of the E12 and E13 chromatin states within 5BS_RS and their signatures (Li et al., 2019).

Supplementary Table S2 | Markers flanking regions of suppression. The markers flanking 5BS_RS and 5BL_RS are in red.

REFERENCES

- Avni, R., Nave, M., Eilam, T., Sela, H., Alekperov, C., Peleg, Z., et al. (2014). Ultra-dense genetic map of durum wheat × wild emmer wheat developed using the 90K iSelect SNP genotyping assay. *Mol. Breed.* 34, 1549–1562. doi: 10.1007/s11032-014-0176-2
- Badaeva, E. D., Dedkova, O. S., Gay, G., Pukhalskiy, V. A., Zelenin, A. V., Bernard, S., et al. (2007). Chromosomal rearrangements in wheat: their types and distribution. *Genome* 50, 907–926. doi: 10.1139/g07-072
- Badaeva, E. D., Keilwagen, J., Knüpfner, H., Waßermann, L., Dedkova, O. S., Mitrofanova, O. P., et al. (2015). Chromosomal passports provide new insights into diffusion of emmer wheat. *PLoS One* 10:e0128556. doi: 10.1371/journal.pone.0128556
- Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* 6:11. doi: 10.1186/s13100-015-0041-9
- Bariah, I., Keidar-Friedman, D., and Kashkush, K. (2020a). Where the wild things are: transposable elements as drivers of structural and functional variations in the wheat genome. *Front. Plant Sci.* 11:585515. doi: 10.3389/fpls.2020.585515
- Bariah, I., Keidar-Friedman, D., and Kashkush, K. (2020b). Identification and characterization of largescale genomic rearrangements during wheat evolution. *PLoS One* 15:e0231323. doi: 10.1371/journal.pone.0231323
- Bashir, T., Chandra Mishra, R., Hasan, M. M., Mohanta, T. K., and Bae, H. (2018). Effect of hybridization on somatic mutations and genomic rearrangements in plants. *Int. J. Mol. Sci.* 19:3758. doi: 10.3390/ijms19123758
- Ben-David, S., Yaakov, B., and Kashkush, K. (2013). Genome-wide analysis of short interspersed nuclear elements SINES revealed high sequence conservation, gene association and retrotranspositional activity in wheat. *Plant J.* 76, 201–210. doi: 10.1111/tj.12285
- Bennetzen, J. L., Schrick, K., Springer, P. S., Brown, W. E., and SanMiguel, P. (1994). Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. *Genome* 37, 565–576. doi: 10.1139/g94-081
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573
- Bhullar, R., Nagarajan, R., Bennypaul, H., Sidhu, G. K., Sidhu, G., Rustgi, S., et al. (2014). Silencing of a metaphase I-specific gene results in a phenotype similar to that of the pairing homoeologous 1 (*Ph1*) gene mutations. *Proc. Natl Acad. Sci.* 111, 14187–14192. doi: 10.1073/pnas.1416241111
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC* 10:421. doi: 10.1186/1471-2105-10-421
- Charles, M., Belcram, H., Just, J., Huneau, C., Viollet, A., Couloux, A., et al. (2008). Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. *Genetics* 180, 1071–1086. doi: 10.1534/genetics.108.092304
- Charlesworth, D. (2016). The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evol. Appl.* 9, 74–90. doi: 10.1111/eva.12291
- Charlesworth, B., Langley, C. H., and Stephan, W. (1986). The evolution of restricted recombination and the accumulation of repeated DNA sequences. *Genetics* 112, 947–962. doi: 10.1093/genetics/112.4.947
- Charlesworth, B., Snegowski, P., and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371, 215–220. doi: 10.1038/371215a0
- Contento, A., Heslop-Harrison, J. S., and Schwarzacher, T. (2005). Diversity of a major repetitive DNA sequence in diploid and polyploid Triticeae. *Cytogenet. Genome Res.* 109, 34–42. doi: 10.1159/000082379
- Dedkova, O. S., Badaeva, E. D., Amosova, A. V., Ruanet, V. V., Mitrofanova, O. P., and Pukhal'skii, V. A. (2009). Diversity and the origin of the European population of *Triticum dicoccum* (Schrank) Schuebl. as revealed by chromosome analysis. *Russ. J. Genet.* 45, 1082–1091. doi: 10.1134/S1022795409090099
- Devos, K. M., Dubcovsky, J., Dvořák, J., Chinoy, C. N., and Gale, M. D. (1995). Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor. Appl. Genet.* 91, 282–288. doi: 10.1007/BF00220890
- Dorer, D. R., and Henikoff, S. (1994). Expansions of transgene repeats cause heterochromatin formation and gene silencing in *drosophila*. *Cell* 77, 993–1002. doi: 10.1016/0092-8674(94)90439-1
- El Baidouri, M., Murat, F., Veyssiere, M., Molinier, M., Flores, R., Burlot, L., et al. (2017). Reconciling the evolutionary origin of bread wheat (*Triticum aestivum*). *New Phytol.* 213, 1477–1486. doi: 10.1111/nph.14113
- Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC* 9:18. doi: 10.1186/1471-2105-9-18
- Fan, C., Hao, M., Jia, Z., Neri, C., Chen, X., Chen, W., et al. (2021). Some characteristics of crossing over in induced recombination between chromosomes of wheat and rye. *Plant J.* 105, 1665–1676. doi: 10.1111/tj.15140
- Faris, J. D., Haen, K. M., and Gill, B. S. (2000). Saturation mapping of a gene-rich recombination hot spot region in wheat. *Genetics* 154, 823–835. doi: 10.1093/genetics/154.2.823
- Feldman, M., and Levy, A. A. (2012). Genome evolution due to allopolyploidization in wheat. *Genetics* 192, 763–774. doi: 10.1534/genetics.112.146316
- Fitzgerald, D. J., Dryden, G. L., Bronson, E. C., Williams, J. S., and Anderson, J. N. (1994). Conserved patterns of bending in satellite and nucleosome positioning DNA. *J. Biol. Chem.* 269, 21303–21314. doi: 10.1016/S0021-9258(17)31963-4
- Fu, S., Yang, M., Fei, Y., Tan, F., Ren, Z., Yan, B., et al. (2013). Alterations and abnormal mitosis of wheat chromosomes induced by wheat-rye monosomic addition lines. *PLoS One* 8:e70483. doi: 10.1371/journal.pone.0070483
- Gadaleta, A., Giancaspro, A., Nigro, D., Giove, S. L., Incerti, O., Simeone, R., et al. (2014). A new genetic and deletion map of wheat chromosome 5A to detect candidate genes for quantitative traits. *Mol. Breed.* 34, 1599–1611. doi: 10.1007/s11032-014-0185-1
- Goncharov, N. P. (2011). Genus *Triticum* L. taxonomy: the present and the future. *Plant Syst. Evol.* 295, 1–11. doi: 10.1007/s00606-011-0480-9
- Gornicki, P., Zhu, H., Wang, J., Challa, G. S., Zhang, Z., Gill, B. S., et al. (2014). The chloroplast view of the evolution of polyploid wheat. *New Phytol.* 204, 704–714. doi: 10.1111/nph.12931
- Griffiths, S., Sharp, R., Foote, T. N., Bertin, I., Wanous, M., Reader, S., et al. (2006). Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752. doi: 10.1038/nature04434
- Guo, C., du, J., Wang, L., Yang, S., Mauricio, R., Tian, D., et al. (2016). Insertions/deletions-associated nucleotide polymorphism in *Arabidopsis thaliana*. *Front. Plant Sci.* 7:1792. doi: 10.3389/fpls.2016.01792
- Hammarlund, M., Davis, M. W., Nguyen, H., Dayton, D., and Jorgensen, E. M. (2005). Heterozygous insertions alter crossover distribution but allow crossover interference in *Caenorhabditis elegans*. *Genetics* 171, 1047–1056. doi: 10.1534/genetics.105.044834
- Holtz, Y., David, J., and Ranwez, V. (2017). The genetic map comparator: a user-friendly application to display and compare genetic maps. *Bioinformatics* 33, btw816–btw1388. doi: 10.1093/bioinformatics/btw816
- Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074–1080. doi: 10.1126/science.1063127
- Jiang, J., and Gill, B. S. (1994). Different species-specific chromosome translocation in *Triticum timopheevii* and *T. turgidum* support diphyletic origin of polyploid wheats. *Chromosom. Res.* 2, 59–64. doi: 10.1007/BF01539455
- Jorgensen, C., Luo, M. C., Ramasamy, R., Dawson, M., Gill, B. S., Korol, A. B., et al. (2017). A high-density genetic map of wild emmer wheat from the Karaca dağ region provides new evidence on the structure and evolution of wheat chromosomes. *Front. Plant Sci.* 8:1798. doi: 10.3389/fpls.2017.01798

- Kagawa, N., Nagaki, K., and Tsujimoto, H. (2002). Tetrad-fish analysis reveals recombination suppression by interstitial heterochromatin sequences in rye (*secale cereale*). *Mol. Gen.* 267, 10–15. doi: 10.1007/s00438-001-0634-5
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kent, T. V., Uzunović, J., and Wright, S. I. (2017). Coevolution between transposable elements and recombination. *Phil. Trans. Roy. Soc. Lond B Biol. Sci.* 372:20160458. doi: 10.1098/rstb.2016.0458
- Koo, H. S., and Crothers, D. M. (1988). Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc. Natl. Acad. Sci. U. S. A.* 85, 1763–1767. doi: 10.1073/pnas.85.6.1763
- Koo, H. S., Wu, H. M., and Crothers, D. M. (1986). DNA bending at adenine-thymine tracts. *Nature* 320, 501–506. doi: 10.1038/320501a0
- Krumsiek, J., Arnold, R., and Rattei, T. (2007). Gepard: A rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 23, 1026–1028. doi: 10.1093/bioinformatics/btm039
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi: 10.1038/nrg2719
- Li, Z., Wang, M., Lin, K., Xie, Y., Guo, J., Ye, L., et al. (2019). The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements. *Genome Biol.* 20:139. doi: 10.1186/s13059-019-1746-8
- Li, L. F., Zhang, Z. B., Wang, Z. H., Li, N., Sha, Y., Wang, X. F., et al. (2022). Genome sequences of five Sitopsis species of Aegilops and the origin of polyploid wheat B subgenome. *Mol. Plant* 15, 488–503. doi: 10.1016/j.molp.2021.12.019
- Lippman, Z., Gendrel, A. V., Black, M., Vaughn, M. W., Dedhia, N., Richard McCombie, W., et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature* 430, 471–476. doi: 10.1038/nature02651
- Lu, H.-J., Fellers, J. P., Friesen, T. L., Meinhardt, S. W., and Faris, J. D. (2006). Genomic analysis and marker development for the Tsn1 locus in wheat using bin-mapped ESTs and flanking BAC contigs. *Theor. Appl. Genet.* 112, 1132–1142. doi: 10.1007/s00122-006-0215-4
- Ma, J., and Bennetzen, J. L. (2004). Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci.* 101, 12404–12410. doi: 10.1073/pnas.0403715101
- Mank, J. E. (2012). Small but mighty: the evolutionary dynamics of W and Y sex chromosomes. *Chromosome Res.* 20, 21–33. doi: 10.1007/s10577-011-9251-2
- Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., The International Wheat Genome Sequencing Consortium et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345:1250092. doi: 10.1126/science.1250092
- Marini, J. C., Levene, S. D., Crothers, D. M., and Englund, P. T. (1982). Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. U. S. A.* 79, 7664–7668. doi: 10.1073/pnas.79.24.7664
- Martienssen, R. A., and Colot, V. (2001). DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science* 293, 1070–1074. doi: 10.1126/science.293.5532.1070
- Mirzaghaderi, G., and Mason, A. S. (2017). Revisiting pivotal-differential genome evolution in wheat. *Trends Plant Sci.* 22, 674–684. doi: 10.1016/j.TPLANTS.2017.06.003
- Mrazek, J. (2010). Comparative analysis of sequence periodicity among prokaryotic genomes points to differences in nucleoid structure and a relationship to gene expression. *J. Bacteriol.* 192, 3763–3772. doi: 10.1128/JB.00149-10
- Mukai, Y., Nakahara, Y., and Yamamoto, M. (1993). Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence in situ hybridization using total genomic and highly repeated DNA probes. *Genome* 36, 489–494. doi: 10.1139/g93-067
- Muterko, A. F. (2017). Quaternion modeling of the helical path for analysis of the shape of the DNA molecule. *Vavilovskii Zhurnal Genetiki i Selektii = Vavilov. J. Genet. Breed.* 21, 878–886. doi: 10.18699/VJ17.308
- Naranjo, T., Roca, A., Goicoechea, P. G., and Giraldez, R. (1987). Arm homoeology of wheat and rye chromosomes. *Genome* 29, 873–882. doi: 10.1139/g87-149
- Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmeler, K., et al. (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. *EMBO J.* 24, 1418–1429. doi: 10.1038/sj.emboj.7600604
- Nelson, J. C., Sorrells, M. E., Van Deynze, A. E., Lu, Y. H., Atkinson, M., Bernard, M., et al. (1995). Molecular mapping of wheat: major genes and rearrangements in homoeologous groups 4, 5, and 7. *Genetics* 141, 721–731. doi: 10.1093/genetics/141.2.721
- Nicetto, D., and Zaret, K. S. (2019). Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Curr. Opin. Genet. Dev.* 55, 1–10. doi: 10.1016/j.gde.2019.04.013
- Poleshko, A., Smith, C. L., Nguyen, S. C., Sivaramakrishnan, P., Wong, K. G., Murray, J. I., et al. (2019). H3K9me2 orchestrates inheritance of spatial positioning of peripheral heterochromatin through mitosis. *elife* 8:e49278. doi: 10.7554/eLife.49278
- Pont, C., Leroy, T., Seidel, M., Tondelli, A., Duchemin, W., Armisen, D., et al. (2019). Tracing the ancestry of modern bread wheats. *Nat. Genet.* 51, 905–911. doi: 10.1038/s41588-019-0393-z
- Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S., and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* 461, 1248–1253. doi: 10.1038/nature08473
- Salina, E. A., Lim, K. Y., Badaeva, E. D., Shcherban, A. B., Adonina, I. G., Amosova, A. V., et al. (2006). Phylogenetic reconstruction of Aegilops section Sitopsis and the evolution of tandem repeats in the diploids and derived wheat polyploids. *Genome* 49, 1023–1035. doi: 10.1139/g06-050
- Salina, E. A., Nesterov, M. A., Frenkel, Z., Kiseleva, A. A., Timonova, E. M., Magni, F., et al. (2018). Features of the organization of bread wheat chromosome 5BS based on physical mapping. *BMC Genomics* 19:80. doi: 10.1186/s12864-018-4470-y
- Salina, E. A., Numerova, O. M., Ozkan, H., and Feldman, M. (2004). Alterations in subtelomeric tandem repeats during early stages of allopolyploidy in wheat. *Genome* 47, 860–867. doi: 10.1139/g04-044
- Salina, E. A., Sergeeva, E. M., Adonina, I. G., Shcherban, A. B., Belcram, H., Huneau, C., et al. (2011). The impact of Ty3-gypsy group LTR retrotransposons Fatima on B-genome specificity of polyploid wheats. *BMC Plant Biol.* 11:99. doi: 10.1186/1471-2229-11-99
- SanMiguel, P., Tikhonov, A., Jin, Y. K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., et al. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274, 765–768. doi: 10.1126/science.274.5288.765
- Segal, E., and Widom, J. (2009). Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* 19, 65–71. doi: 10.1016/j.sbi.2009.01.00
- Shcherban, A. B., Sergeeva, E. M., Badaeva, E. D., and Salina, E. A. (2008). Analysis of 5S rDNA changes in synthetic allopolyploids Triticum×Aegilops. *Mol. Biol.* 42, 536–542. doi: 10.1134/S0026893308040080
- Smit, A. F. A., Hubley, R., and Green, P. (2008). RepeatModeler Open-1.0. Seattle, USA: Institute for Systems Biology. Available at: <http://www.repeatmasker.org>
- Smit, A. F. A., Hubley, R., and Green, P. (2013). RepeatMasker Open-4.0. Available at: <http://www.repeatmasker.org>
- Sosnowski, O., Charcosset, A., and Joets, J. (2012). Biomecator V3: An upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics* 28, 2082–2083. doi: 10.1093/bioinformatics/bts313
- Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* 20, 267–273. doi: 10.1038/nsmb.2506
- Termolino, P., Cremona, G., Consiglio, M. F., and Conicella, C. (2016). Insights into epigenetic landscape of recombination-free regions. *Chromosoma* 125, 301–308. doi: 10.1007/s00412-016-0574-9
- Tikhonov, A. P., SanMiguel, P. J., Nakajima, Y., Gorenstein, N. M., Bennetzen, J. L., and Avramova, Z. (1999). Colinearity and its exceptions in orthologous adh regions of maize and sorghum. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7409–7414. doi: 10.1073/pnas.96.13.7409
- Timonova, E. M., Dobrovol'skaya, O. B., Sergeeva, E. M., Bildanova, L. L., Sourdille, P., Feuillet, C., et al. (2013). A comparative genetic and cytogenetic mapping of wheat chromosome 5B using introgression lines. *Russ. J. Genet.* 49, 1200–1206. doi: 10.1134/S1022795413120132

- Tolstorukov, M. Y., Virnik, K. M., Adhya, S., and Zhurkin, V. B. (2005). A-tract clusters may facilitate DNA packaging in bacterial nucleoid. *Nucleic Acids Res.* 33, 3907–3918. doi: 10.1093/nar/gki699
- Vermaak, D., Ahmad, K., and Henikoff, S. (2003). Maintenance of chromatin states: an open-and- shut case. *Curr. Opin. Cell Biol.* 15, 266–274. doi: 10.1016/s0955-0674(03)00043-7
- Vershinin, A. V., and Heslop-Harrison, J. S. (1998). Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Mol. Biol.* 36, 149–161. doi: 10.1023/a:1005912822671
- Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588, 277–283. doi: 10.1038/s41586-020-2961-x
- Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183
- Wang, Z. Z., Xie, J. Z., Guo, L., Zhang, D. Y., Li, G. Q., Fang, T. L., et al. (2018). Molecular mapping of YrTZ2, a stripe rust resistance gene in wild emmer accession TZ-2 and its comparative analyses with *Aegilops tauschii*. *J. Integr. Agric.* 17, 1267–1275. doi: 10.1016/S2095-3119(17)61846-X
- Westphal, T., and Reuter, G. (2002). Recombinogenic effects of suppressors of position-effect variegation in drosophila. *Genetics* 160, 609–621. doi: 10.1093/genetics/160.2.609
- Wingen, L. U., West, C., Leverington-Waite, M., Collier, S., Orford, S., Goram, R., et al. (2017). Wheat landrace genome diversity. *Genetics* 205, 1657–1676. doi: 10.1534/genetics
- Yelagandula, R., Stroud, H., Holec, S., Zhou, K., Feng, S., Zhong, X., et al. (2014). The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell* 158, 98–109. doi: 10.1016/j.cell.2014.06.006
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214. doi: 10.1089/10665270050081478
- Zhu, T., Wang, L., Rimbert, H., Rodriguez, J. C., Deal, K. R., de Oliveira, R., et al. (2021). Optical maps refine the bread wheat *Triticum aestivum* cv. *Chinese Spr. Gen. Assembly. Plant J* 107, 303–314. doi: 10.1111/tpj.15289
- Zimin, A. V., Puiu, D., Hall, R., Kingan, S., Clavijo, B. J., and Salzberg, S. L. (2017). The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *Gigascience* 6, 1–7. doi: 10.1093/gigascience/gix097
- Ziolkowski, P. A., Berchowitz, L. E., Lambing, C., Yelina, N. E., Zhao, X., Kelly, K. A., et al. (2015). Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during Arabidopsis meiosis. *Elife* 4:e03708. doi: 10.7554/eLife.03708

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Salina, Muterko, Kiseleva, Liu and Korol. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.