BMC Genomics

RESEARCH

Open Access

Genome-wide simple sequence repeat analysis and specific molecular marker development of rye

Zhi Li[1,2,3†], Liqi Zhao[2,3†], Tao Yang[2,3], Jingsha Tang[2,3], Yu Miao[2,3] and Tianheng Ren[1,2,3*]

Abstract

Background Rye (*Secale cereale* L.) is the most widely used related species in wheat genetic breeding, and the introduction of its chromosome fragments into the wheat genome through distant hybridization is essential for enriching the genetic diversity of wheat. Rapid and accurate detection of rye chromatin in the wheat genome is important for distant hybridization. Simple sequence repeats (SSRs) are widely distributed in the genome, and SSRs of different species often exhibit species-specific characteristics.

Results In this study, genome-wide SSRs in rye were identified, and their characteristics were outlined. A total of 997,027 SSRs were selected, with a density of 115.97 SSRs/Mb on average. There was no significant difference in the number of SSRs on each chromosome. The number of SSRs on 2R was the highest (15.29%), and the number of SSRs on 1R was the lowest (13.02%). The number of SSRs on each chromosome is significantly correlated with chromosome length. The types of SSR motifs were abundant, and each type of SSR was distributed on 7 chromosomes of rye. The numbers of mononucleotide simple sequence repeats (MNRs), dinucleotide simple sequence repeats (DNRs), and trinucleotide simple sequence repeats (TNRs) were the greatest, accounting for 46.90%, 18.37%, and 22.64% of the total number, respectively. Among the MNRs, the number of G/C repeats and the number of 10 bp motifs were the greatest, accounting for 26.24% and 31.32% of the MNRs, respectively. Based on the SSR sequences, a total of 657 pairs of primers were designed. The PCR results showed that 119 pairs of these primers were rye-specific and could effectively detect rye chromatin in the wheat genome. Moreover, 86 pairs of the primers could also detect one or more specific rye chromosomes.

Conclusion These results lay a foundation for both genomic evolution studies of rye and molecular breeding in wheat.

Keywords Rye, SSR, Molecular Markers, Chromosome-specfic markers, Universal primers of rye

†Zhi Li and Liqi Zhao contributed equally to this work.

*Correspondence:
Tianheng Ren
renth@sicau.edu.cn
[1]State key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, China
[2]College of Agronomy, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, China
[3]Key Laboratory of Plant Genetics and Breeding, Sichuan Agricultural University of Sichuan Province, Wenjiang, Chengdu 611130, Sichuan, China

Background

Wheat is one of the most important crops in the world. However, during the process of targeted selection, many useful genes in wheat are lost, resulting in a decline in genetic diversity [1]. The introduction of DNA from related species of wheat into the wheat genome is an important method for improving wheat genetic diversity [2, 3]. Rye is one of the most important related species of wheat and plays a key role in wheat genetic breeding worldwide because it has many favorable genes that give rise to increased yield, disease resistance, and other desirable agronomic traits [2, 4–6]. In recent years, the 1R to 7R chromosomes of different rye sources were successfully introduced into the wheat genome and have shown great value for wheat breeding [4–12]. Therefore, accurate identification of different types of rye DNA in the wheat genome has become the focus of research on wheat-rye distant hybridization. At present, the most commonly used methods to detect alien DNA in the wheat genome include cytological methods and molecular markers [3, 5, 11, 12]. Molecular cytology methods such as C-band, fluorescence in situ hybridization (FISH), and genome in situ hybridization (GISH) can also detect alien chromatin in the wheat genome [13]. Many different molecular markers, such as random amplification polymorphic DNA (RAPD) [14], amplified fragment length polymorphism marker (AFLP) [15], sequence-tagged site (STS) [16], sequence characterized amplified regions (SCAR) [16], diversity arrays technology (DArT) [17], and single nucleotide polymorphisms (SNPs) [3], have been used to detect alien chromatin in the wheat genome in recent years. However, the amplification conditions of some molecular markers are complex, difficult to operate, or can only be verified based on methods such as sequencing. SSR (simple sequence repeat) molecular markers are widely used because of their simple operation, high stability, and low cost, and they were designed based on repeat sequences in the genome [18].

An SSR, also known as a microsatellite, is a DNA sequence formed by a repeated series of 1 to 6 nucleotides in the genome [19]. It exists widely in the whole genome of eukaryotes and is randomly distributed, with abundant polymorphisms [20, 21]. Because SSRs are highly variable in the genome, they are broadly used in genetic molecular markers and population genetics studies [18, 21]. Although SSRs are commonly distributed in the genome; their composition, length, and distribution of different motifs have unique characteristics and maintain a certain degree of stability [20–22].

While there are many repeats in the rye genome [23], systematic analysis of SSRs in rye has not been reported. In this study, we identified genome-wide SSRs in 'Weining' rye and outlined their characteristics, and 119 pairs of rye-specific SSR primers were developed. All these 119 pairs of primers for SSRs can not only detect cultivated rye (*Secale cereale* L.) chromatin in the wheat genome but also detect weedy rye (*S. cereale* subsp. *Segetale*, etc) and wild rye (*S. cereale* subsp. *vavilovii*) chromosomes in the genetic background of wheat.

Materials and methods

Genomic data and plant materials

The reference genome of rye 'Weining' [23], rye 'Lo7' and the wheat Chinese Spring (CS) were used in this study [24, 25]. All genomic data were downloaded from the National Center for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov) and the Wheatomics database (http://202.194.139.32). Seven rye varieties: 'Weining', 'Jingzhou', 'Qinling', 'Aigan', 'Baili', 'Kustro', and accession PI436168, and two wheat cultivars: CS and 'Mianyang11' (MY11) were used in this study. 'Weining', 'Jingzhou', 'Qinling', 'Aigan', and 'Baili' rye were collected by our laboratory in China in the 1990s [26]. Rye accession PI436168 was kindly provided by the Germplasm Resources Information Network (GRIN) of the United States Department of Agriculture (USDA) and was collected from Chile [26]. The rye variety 'Kustro' was collected from the U.S.A. by our laboratory in the 1990s [26]. The 1R-7R wheat-rye monosomic addition lines or disomic addition lines, wheat-weedy rye and wheat-wild rye addition lines, and translocation lines were all provided by the Key Laboratory of Plant Genetics and Breeding at Sichuan Agricultural University of Sichuan Province, and their chromosome structures were identified by nondenaturing fluorescence in situ hybridization (ND-FISH) before use.

Identification of genome-wide SSRs

The genome-wide SSRs of 'Weining' rye were identified by using MicroSatellite (MISA) software under default parameters [22], and repetitive mononucleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide motifs were identified. If the interval between two SSRs was less than 100 bp, they were labeled as one compound SSR (cSSR).

SSR primers development and validation

Fifty base pairs of the flanking sequences of the SSR loci from the 'Weining' rye reference genome were examined. The sequence similarity among these sequences and the sequences of the CS and 'Lo7' reference genomes were analyzed by the online tools BLASTN on the NCBI and Wheatomics websites. SSR sequences with more than 80% similarity to *CS* and less than 90% similarity to 'Lo7' were excluded. To develop the SSR markers, 200 bp sequences of the flanking region of each remaining SSR locus were used for primer design through Primer3 software [27]. The parameters of the primer design were set

as follows: PCR product sizes ranging from 300 to 600 bp, primer lengths ranging from 18 to 23 bp, melting temperatures ranging from 55 °C to 65 °C, and GC contents ranging from 40 to 60%. The primers were synthesized by Beijing Tsingke Biotech Co., Ltd.

The genomic DNA of plants was isolated from young leaves by the surfactant cetyltrimethylammonium bromide (CTAB). PCR was carried out in a Bio-Rad iCycler thermal cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA). DNA was amplified with 5 μL of premix Taq, 1 μL (10 ng/μL) of each primer, and 2 μL (50 ng/μL) of DNA in a total volume of 10 μL. After initial denaturation for 5 min at 95 °C, each cycle included 30 s of denaturation at 95 °C, 30 s of annealing at 55 ~ 58 °C (determined by primer sequence), and 15 s of extension at 72 °C. A final extension for 10 min at 72 °C followed by 33 cycles. PCRs were stored at 4 °C until they were resolved by electrophoresis on 2% agarose gels stained with ethidium bromide. To determine whether the primers used were specific to the rye genome, the DNA of seven rye varieties and two wheat cultivars was used. To determine which rye chromosomes could be detected via SSR primers, seven wheat-rye monosomic addition lines (1R to 7R) were used.

Data analysis and statistics

The statistics and classification of SSR types were conducted based on the MISA output results and analysis in Excel 2019. Correlation analysis was performed using SPSS ver. 25.0 (IBM SPSS 22.0, Chicago, IL, USA). The analysis results were generated using the CNSknowall platform (https://cnsknowall.com/#/HomePage), a comprehensive web service for biomedical data analysis and visualization.

Results

Total SSRs in the *Weining* rye genome

The distribution and number of different types of SSRs on each chromosome of rye are shown in Fig. 1 and Table S1. A total of 997,027 SSR loci were identified in the 'Weining' rye genome, and 13,357 SSR loci were located in the exon region (1.34% in total). There were 129,794, 152,396, 145,836, 146,901, 138,225, 144,937, and 138,938 SSR loci distributed from the 1R to 7R chromosomes, respectively (Table S1). Among them, the number of SSRs on the 2R chromosome was the highest (15.29% of the total), while the number on the 1R chromosome was the lowest (13.02% of the total). However, there was no significant difference in the number of SSRs on each chromosome. The results showed that mononucleotide simple sequence repeats (MNRs) were the most abundant at 467,556 (46.9% of the total), followed by trinucleotide simple sequence repeats (TNRs; 225,703, 22.64% of the total) and dinucleotide simple sequence repeats (DNRs; 183,179, 18.37% of the total). The number and proportion of tetranucleotide simple sequence repeats (TtNRs; 11,921, 1.2% of the total), pentanucleotide simple sequence repeats (PNRs; 1,811, 0.18% of the total), and hexanucleotide simple sequence repeats (HNRs; 907, 0.09% of the total) were relatively low, accounting for only 1.47% of the total SSRs. In addition, there were 105,950 cSSRs in the genome, accounting for 10.63% of the total SSRs (Fig. 1, Table S1).

The different types of SSRs also exhibited different densities in the 'Weining' rye genome (Fig. 2, Table S2). The results showed that MNRs had the highest density of 64.55 per Mb, and the average distance between two MNRs was only 15.34 kb. The average densities of DNRs and TNRs were 25.24 per Mb and 31.13 per Mb, respectively, and the average distances between the two SSRs were 39.64 kb and 32.12 kb, respectively. The average density of TtNRs, PNRs, and HNRs was low (within 2 per Mb), and the average distance between the two SSRs was long, above 600 kb. The average density of the HNRs was the lowest (0.13 per Mb), and the average distance between the two HNRs was the greatest (8093.67 kb).

The results showed that the SSRs were more densely distributed in the two distal parts of the chromosomes

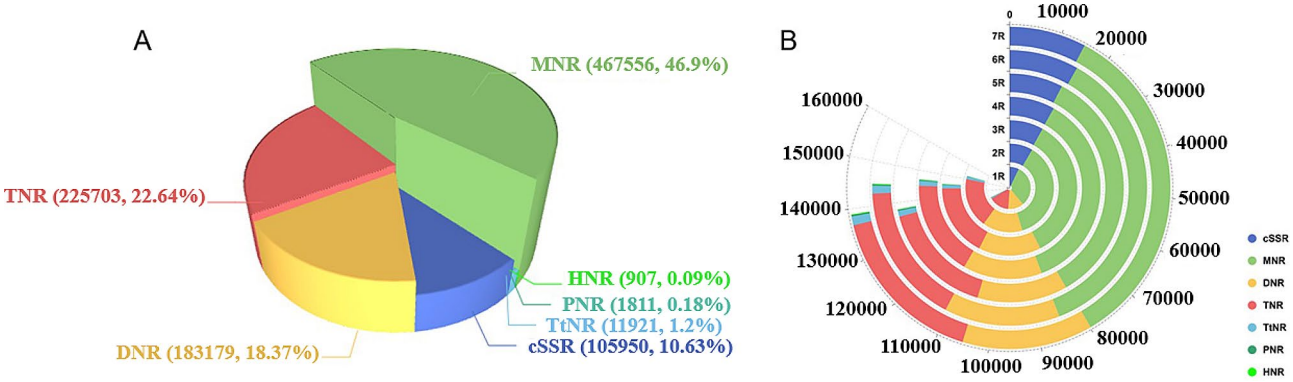

Fig. 1 SSRs distribution in the whole genome of 'Weining' rye. **A**: Proportion of different types of SSRs in the whole genome of 'Weining' rye. **B**: Number of SSRs on each chromosome of the whole genome of 'Weining' rye

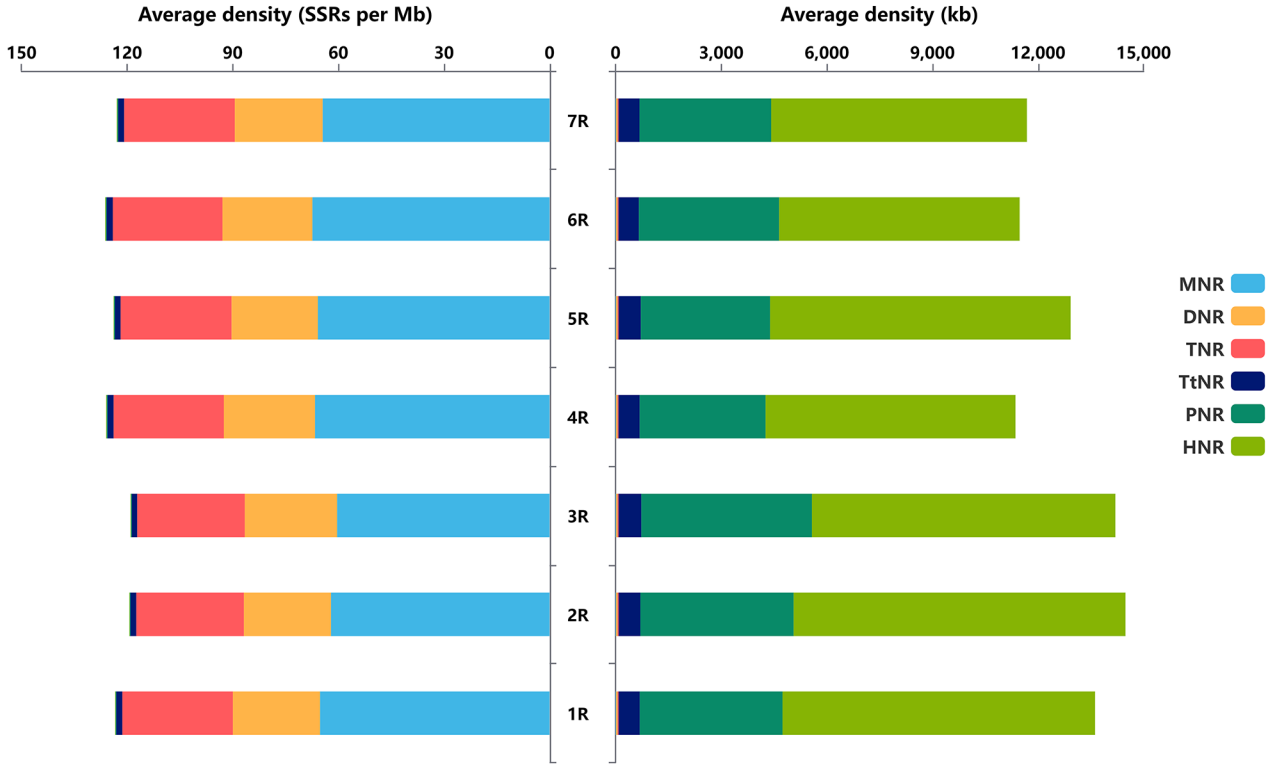

Fig. 2 Distribution density and distance of different types of SSRs in the 'Weining' rye genome

and less densely distributed in the proximal part of the chromosomes near the centromere on the 1R-7R chromosomes (Fig. 3). In addition, MNRs, DNRs, and TNRs were more densely distributed in the two distal parts of the chromosomes than that in the proximal part of the chromosomes near the centromeres (Fig. 4). However, there were no obvious distribution patterns of TtNRs, PNRs, or HNRs; that is, these three types of SSRs were randomly distributed on the chromosomes of 'Weining' rye. This result may be due to the low number of SSRs of these three types.

The lengths of different types of SSRs were also analyzed. With the increase in the number of nucleotides in the repeat sequences, the total length of the SSRs of different repeat types tended to decrease (Fig. 5, Table S3). For example, the total length of the MNRs was the highest, with an average length of 808,578.71 bp on each chromosome. On the other hand, the average length of each MNR was the shortest, with an average length of 12.11 bp. The total length of the HNRs on each chromosome was the shortest, with an average of 4175.14 bp, and the average length of the HNRs was the longest, with an average of 32.19 bp. In general, for the total length of the different types of SSRs, the MNRs had the longest length, followed by the DNRs, TNRs, TtNRs, PNRs, and HNRs. The average length of the HNRs was the greatest, followed by that of the PNRs, TtNRs, DNRs, TNRs, and MNRs (Table S3). In addition, TNRs had the largest length span of 15 to 399 bp, and the length spans of DNRs, TtNRs, MNRs, and HNRs were 12 to 330 bp, 20 to 212 bp, 10 to 135 bp, and 30 to 90 bp, respectively. The PNRs had the shortest length, ranging from 25 to 80 bp (Table 1). The max repeat times of the MMRs, DNRs, TNRs, TtNRs, PNRs, and HNRs were 135, 165, 133, 53, 16, and 15, respectively (Table 1). Among the MNRs, the 10-repeat motif was the most common (31.32% of the MNRs), the 6-repeat motif was the most common (33.68% of the DNRs), and among the TNRs, TtNRs, PNRs, and HNRs, the 5-repeat motif was the most common (68.85%, 66.33%, 80.16%, and 71.72%, respectively) (Table 1).

A total of 712 types of SSR motifs in the 'Weining' rye genome were found (2 types in MNRs, 6 types in DNRs, 30 types in TNRs, 120 types in TtNRs, 283 types in PNRs, and 271 types in HNRs) (Table 2). In the MNRs, the amount of $(C/G)_n$ (261,449, 55.92%) was significantly greater than that of $(A/T)_n$ (206,107,44.08%). Among the DNRs, the (AT/AT)n (50,937, 27.81%), $(CT/AG)_n$ (47,389, 25.87%), and $(AG/CT)_n$ (45,778, 24.99%) motifs were more abundant, while the other three types were less abundant (90,012, 21.33%). The (CGC/GCG)n (25,542, 11.32%) and $(CCG/CGG)_n$ (25,400, 11.25%) motifs were more abundant in TNRs, more $(AATT/AATT)_n$ (1539, 12.91%) and (ATTA/TAAT)n (1511, 12.68%) motifs

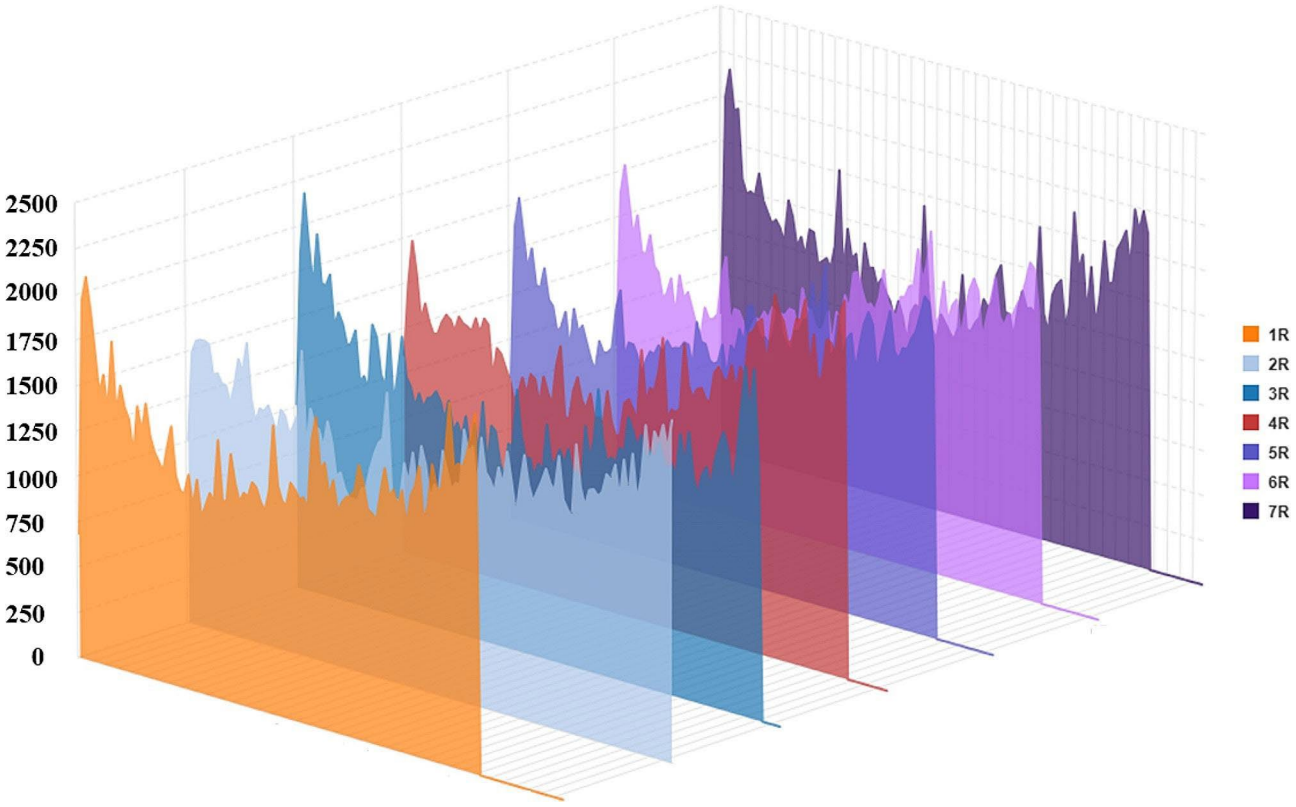

Fig. 3 SSR density distribution on each chromosome of the 'Weining' rye genome. Y-axis: the number of SSRs; X-axis: chromosome position, the region between two lines = 30 Mb

were found in TtNRs, more $(AACAA/TTGTT)_n$ (257, 14.19%) and $(AAACA/TGTTT)_n$ (232, 12.81%) motifs were found in PNRs, and more $(AAAGGC)_n$ (71, 7.83%) and $(GCCTTT)_n$ (62, 6.84%) motifs were found in HNRs (Table 2).

Distribution analysis revealed that SSRs were variably distributed on every chromosome and that the number of SSRs was positively correlated ($r=0.931$; $p<0.01$) with chromosome length (Fig. 6). The largest number of SSRs was observed on chromosome 2R, while the smallest number of SSRs was observed on chromosome 1R (Fig. 6). The number of different types of SSRs, except for PNRs and HNRs, also showed positive correlations with chromosome length (Fig. 7).

Development and validation of SSR markers

A total of 657 pairs of SSR primers were identified in this study. The PCR results showed that 119 pairs of SSR primers could distinguish between the DNA of rye and wheat. Therefore, these 119 SSR markers were considered rye-specific molecular markers that could be used to detect rye DNA in the wheat genetic background. Moreover, these markers showed different amplification patterns. fifteen markers, which are considered A-type molecular markers, showed the same specific bands in different rye genomes, and no bands in wheat genomes (Fig. 8A). 44 markers, which are considered B-type molecular markers, showed different specific bands in different rye genomes but no bands in wheat genomes (Fig. 8B). 45 markers, which are considered C-type molecular markers, showed the same specific bands in different rye genomes and bands of different sizes in the wheat genome (Fig. 8C). fifteen markers, which are considered D-type molecular markers, showed different specific bands in different rye genomes as well as bands of different sizes (different from all rye-specific bands) in the wheat genome (Fig. 8D). B-type and D-type markers were considered to be polymorphic markers for rye, accounting for 49.6% (59 in 119) of the rye-specific molecular markers. The sequences and Tm values of these 119 pairs of primers are listed in Table S4. The sizes of the products amplified by these primers in the rye and wheat genomes are listed in Table S5.

To determine which of the 119 SSR markers could be used to detect the specific rye chromosome, we conducted the following experiments. The results showed that 84 pairs of primers could be used to detect a specific chromosome or several chromosomes of rye (Fig. 9, Table S6). For example, the molecular markers WN1R110, WN2R47, WN5R36, WN6R82, and WN7R61 can be used to detect chromosomes 1R, 2R, 5R, 6R, and 7R, respectively, in the wheat genetic background (Fig. 9A).

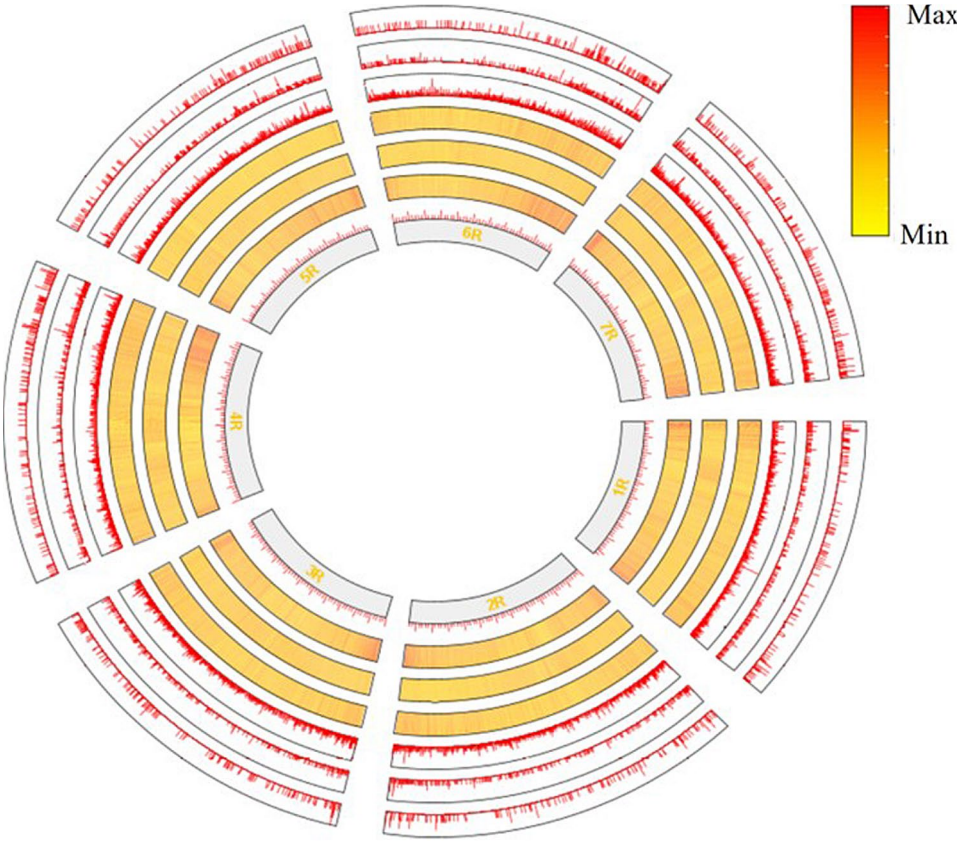

Fig. 4 Different types of SSR density distributions on each chromosome. From inner ring to outer ring: the density of MNR, DNR, TNR, TtNR, PNR, and HNR on 1R to 7R chromosomes

In addition, several primers can be used to detect more than one rye chromosome in the wheat genetic background. For example, WN6R78, WN6R24, WN1R7.8, WN4R26, WN6R35.1, and WN1R42.3 can detect two (6R and 7R), three (4R, 6R, 7R), four (1R, 4R, 6R, 7R), five (1R to 5R), six (1R to 6R), and seven (1R to 7R) rye chromosomes in the wheat genetic background (Fig. 9B). A specific chromosome can also be detected by a combination of different molecular primers. For example, the combination of the molecular marker WN6R12 and the molecular marker WN1R41.2 can be used to effectively detect the 3R chromosome (Table S6). The details of the amplification of molecular primers on different rye chromosomes are shown in Table S3.

The application of SSR markers

All SSR markers were successfully validated in cultivated rye (Table S5). Some plant materials that were developed from wild or weedy rye (including wheat-rye translocation and addition lines) were detected by these SSR markers. The results showed that the SSR markers could be used to detect weedy or wild rye chromosomes in the wheat background (Fig. 10). For example, the SSR marker WN1R24.1 amplified a specific band from the DNA of cultivated rye ('Weining' and 'Kustro' rye) and weedy rye (*S. cereale* subsp. *segetale*, etc.) and from 1R addition lines originating from the cross of CN25 × *S. dighoricum*, 6R addition lines originating from the cross of CN25 × *S. dighoricum*, 2R addition lines originating from the cross of CN27 × *S. segetale*, and 1RS·1BL translocation lines originating from the cross of CN25 × *S. dighoricum* (Fig. 10A). The SSR markers WN1R7.8 and WN7R61 amplified specific bands from the DNA of cultivated rye and wild rye (*S. vavilovii*) and from the 1R and 7R addition lines originating from the cross of CN23 x *S. vavilovii*, respectively (Fig. 10B and C). These results indicated that the SSR markers developed in this study can not only detect the chromosomes of cultivated rye in the wheat genetic background but also detect weedy and wild rye chromosomes in the wheat genetic background.

Discussion

SSRs in the rye genome

The goal of studies on SSRs is to reveal the evolution and function of the species. The distribution and organization of SSRs in the genome can also provide important information about the genome structure [20, 21]. The analysis of the number, length, type, and distribution of SSRs in

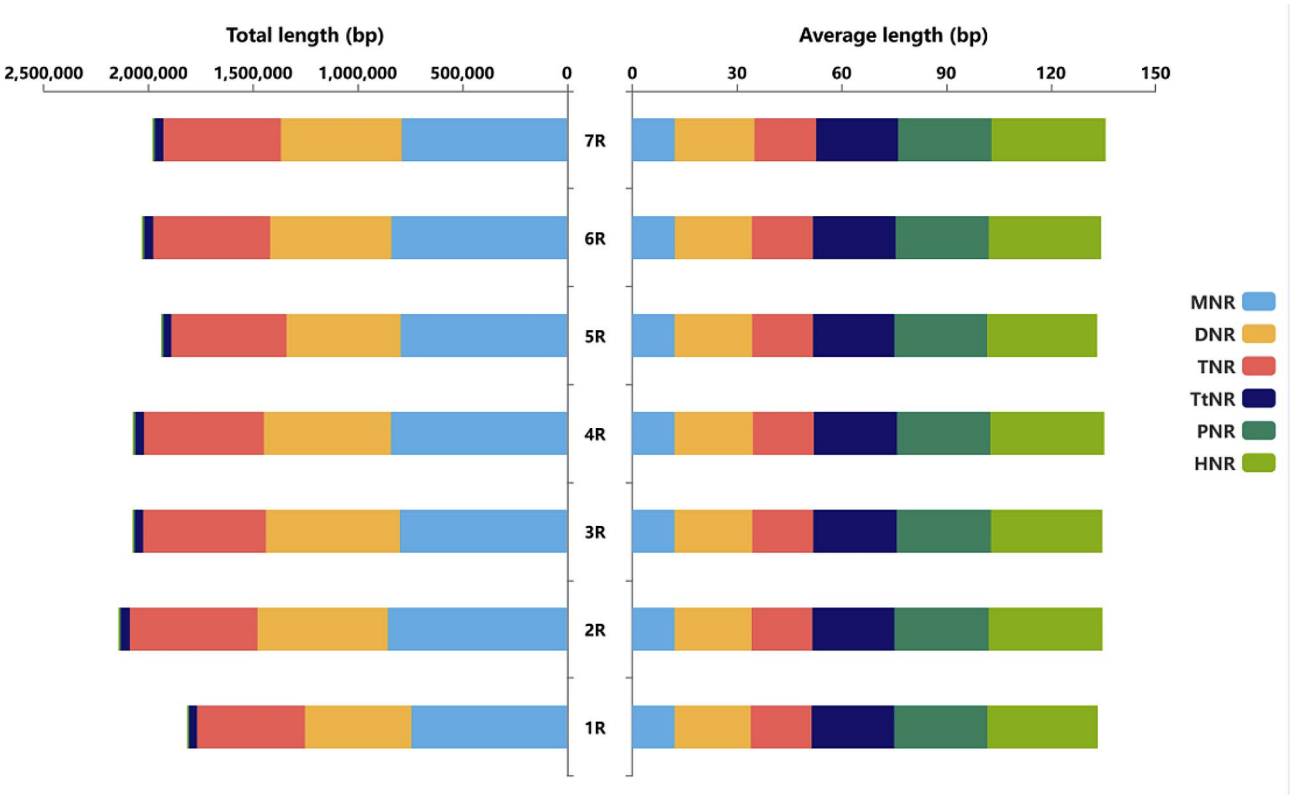

Fig. 5 Total and average lengths of different types of SSRs on each chromosome in the 'Weining' rye genome

Table 1 Spans and number of repeats of different types of SSRs in the 'Weining' rye genome

Types	Length span	Max repeat times	No. of SSR	Most SSR motif	Ratio
MNR	10~135	135	467,556	10 times repeat	31.32%
DNR	12~330	165	183,179	6 times repeat	33.68%
TNR	15~399	133	225,703	5 times repeat	68.85%
TtNR	20~212	53	11,921	5 times repeat	66.33%
PNR	25~80	16	1811	5 times repeat	80.16%
HNR	30~90	15	907	5 times repeat	71.72%

the whole genome provides insights into the organization of the genome and lays the foundation for further genome research [21]. Moreover, because SSRs are highly variable and easy to mutate and recombine, by comparing the differences in SSRs among different individuals, populations, or species, their genome evolution, population structure, and kinship can be seen [22, 28, 29]. Although SSRs cannot encode proteins, they play important roles in genome function and regulation and may participate in the maintenance of chromosome structure, the regulation of genome stability, and transcriptional regulation [29–31].

Therefore, the study of plant SSRs helps understand the evolutionary history and relationships among different plant species. At present, genome-wide SSRs in many plant species, such as *Actinidia chinensis* [32], *Akebia trifoliata* [22], *Spinacia oleracea* [33], *Cucumis sativus* [34], *Shorea robusta* [35], *Nicotiana* [36], and *Camellia sinensis* [37], have been analyzed. For example, Zhong et al. [22] analyzed the genome-wide SSRs of *A. trifoliata*, and 434,293 SSRs were identified. The major types of SSRs in the *A. trifoliata* genome were combinations of "A" and "T" repeats, such as "A/T", "AT/AT", "AAT/ATT", and "AAAT/ATTT". A total of 174 accessions and 72 SSR markers were used to systematically analyze the diversity and structure of *A. trifoliata*, and all the accessions could be divided into two groups [38]. Song et al. [39] identified 249,882 SSRs in 112 different plants, and more SSRs were found in lower plants than in higher plants, showing that lower plants need to adapt to early extreme environments. A search for SSR loci of spinach line Sp75 and a total of 42,155 loci with repeat motifs of two to six nucleotides in the Sp75 reference genome were identified [33]. Compared with those of other spinach accessions, there were 5986 polymorphic SSR loci, and they were present at a density of 12.9 SSRs/Mb on the chromosome

Table 2 SSR motifs in the 'Weining' rye genome

SSR types	No. of motifs	Major motifs	No.	Ratio in this SSR type	Ratio in total SSR
MNRs	2	C/G	261,449	55.92%	26.22%
		A/T	206,107	44.08%	20.67%
DNRs	6	AT/AT	50,937	27.81%	5.11%
		CT/AG	47,389	25.87%	4.75%
		AG/CT	45,778	24.99%	4.59%
		Others (3)	90,012	21.33%	3.92%
TNRs	30	CGC/GCG	25,542	11.32%	2.56%
		CCG/CGG	25,400	11.25%	2.55%
		CTT/AAG	10,808	4.79%	1.08%
		AAG/CTT	15,391	6.82%	1.54%
		AGT/ACT	14,896	6.60%	1.49%
		CCT/AGG	12,492	5.53%	1.25%
		CTA/TAG	11,267	4.99%	1.13%
		Others (23)	109,907	48.70%	11.02%
TtNRs	120	AATT/AATT	1539	12.91%	0.15%
		ATTA/TAAT	1511	12.68%	0.15%
		AAAT/ATTT	484	4.06%	0.05%
		Other (117)	8387	70.35%	0.84%
PNRs	283	AACAA/TTGTT	257	14.19%	0.026%
		AAACA/TGTTT	232	12.81%	0.023%
		AAAAG/CTTTT	152	8.39%	0.015%
		CTTTT/AAAAG	99	5.47%	0.010%
		Others (279)	1071	59.14%	0.107%
HNRs	271	AAAGGC	71	7.83%	0.0071%
		GCCTTT	62	6.84%	0.0062%
		AAAAAG/CTTTTT	52	5.73%	0.0052%
		CGGAGG/CCTCCG	33	3.64%	0.0033%
		Others	689	75.96%	0.0691%

on average [33]. There have also been several reports on the SSR of *Poaceae*. For example, Zhao et al. [40] demonstrated that SSRs were more densely present in telomeric regions than in centromeric regions in maize, with a median length of 14–18 bp and a genome-wide average density of 3355.77 bp/Mbp, although SSRs accounted for only 0.03% of the maize genome. Han et al. [41] identified 364,347 SSRs in the Chinese Spring wheat reference genome, which were present at a density of 36.68 SSR/Mb, and 488 types of motifs ranging from di- to hexanucleotides were identified. Zhang et al. [42] compared the distribution patterns of $(AAC)_n$, $(AAG)_n$, $(AGC)_n$, and $(AG)_n$ among wheat, *T. urartu*, *Ae. speltoides*, *Ae. tauschii*, *T. dicocoides*, and *T. dicoccum* indicated that there were specific distribution patterns in different chromosomes from different species for each SSR and that the B genome might be more sensitive and more changeable during the process of polyploidization.

Although there are a large number of SSRs in the rye genome, there are few relevant reports on this topic. In this study, the genome-wide SSRs of 'Weining' rye were analyzed. A large number (997,027) of different types of SSRs were identified in the genome of 'Weining' rye and were variably distributed on chromosomes 1R to 7R, with an average density of 127.82 per Mb (Table S1, Fig. 1). The results showed that the distribution patterns of SSRs in 'Weining' rye were very similar to those in other plant species, with more SSRs in the distal region of the chromosomes and fewer in the proximal regions near the centromere [22, 40] (Figs. 3 and 4). The types of SSRs of rye were also similar to those of other plant species. The numbers of SSRs with fewer repetitions (MNR, DNR, and TNR) were significantly greater than those with more repetitions [22, 32, 36, 40]. However, in the genome of rye, there were also a large number of cSSRs, accounting for 10.63% of all SSRs (Table S1). For many other plant species, the major types of SSR motifs in the genome are combinations of "A" and "T" repeats [22, 37]; however, in the genome of 'Weining' rye, the major types of SSR motifs are combinations of "C" and "G" repeats (Table 2). This phenomenon indicated the genetic diversity of SSRs in different plant species. The genetic diversity of SSRs in different plants is high, not only in their number and position but also in their base composition.

SSR marker development and utilization

Several cytological methods, such as C-band, FISH, and GISH, are important methods for detecting alien DNA

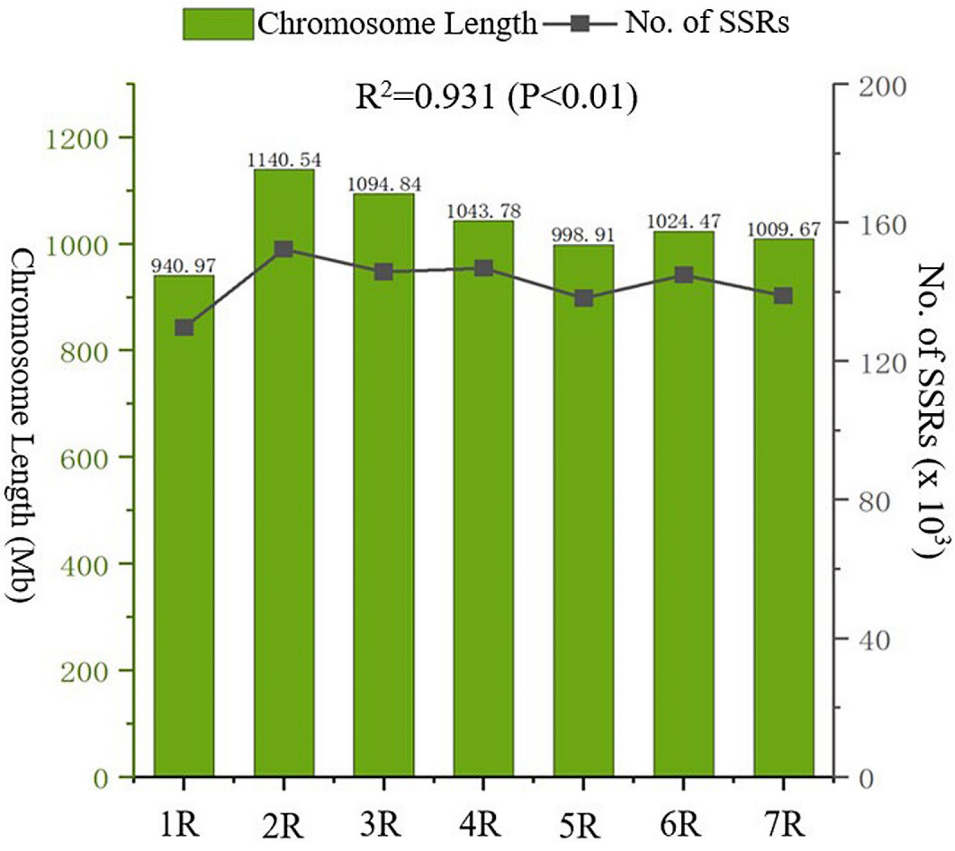

Fig. 6 Correlation between chromosome length and SSR number

in the wheat genome [5, 43]. However, these methods are complex and require expensive microscopy, and researchers are also required to have extensive experience in detecting chromosomes. Therefore, the use of SSRs, a type of molecular marker with strong repeatability, high efficiency, and low cost, requires only a common PCR instrument and still plays an important role in the identification of alien DNA in the wheat genome. Nevertheless, the genomic sequences of wheat and rye have high similarity, and rye itself has high genetic diversity [26]. The designed molecular markers often have disadvantages such as poor universality and specificity. Moreover, some rye-specific molecular markers can be used to detect rye DNA in the wheat genome, but these markers cannot be located on one or more specific rye chromosomes [44]. Several markers can amplify only rye DNA but cannot be used to detect rye DNA in the wheat genome [45]. For example, Xu et al. [44] developed 414 rye-specific molecular markers based on the EST sequence of rye, of which only 31 markers were located on the 1R-7R chromosome. Li et al. [45] reported that some primers could not amplify rye-specific PCR products when they were tested in wheat-rye addition lines and translocation lines. Hundreds of SSR markers specific for rye were also developed based on the sequence information of SLAF-Seq; however, these markers were developed from limited wheat and rye sources [9, 46, 47]. In this study, 657 pairs of new SSR primers for rye were designed. PCR validation indicated that 119 pairs of primers were rye-specific. The universality of these primers was good, and they could be used to detect the DNA of cultivated rye, weedy rye, and wild rye in the wheat genome (Figs. 8, 9 and 10). The specificity of these primers was also good, and 84 pairs of primers could be used to detect specific chromosomes of rye (Fig. 9, Table S6). Polymorphic markers have great advantages in population genetic analysis and genetic map construction. For example, Fandade et al. [48] identified 189 polymorphic SSRs and indicated that there was high genetic diversity in Indian wheat. Yet, the polymorphism of SSR markers in rye was low [9, 46, 47, 49, 50]. In this study, 59 markers (49.6%) were polymorphic when they were tested in different rye samples (Fig. 7, Table S5). These SSR markers can provide important support for marker-assisted selection (MAS) in wheat-rye distant hybridization breeding programs in the future.

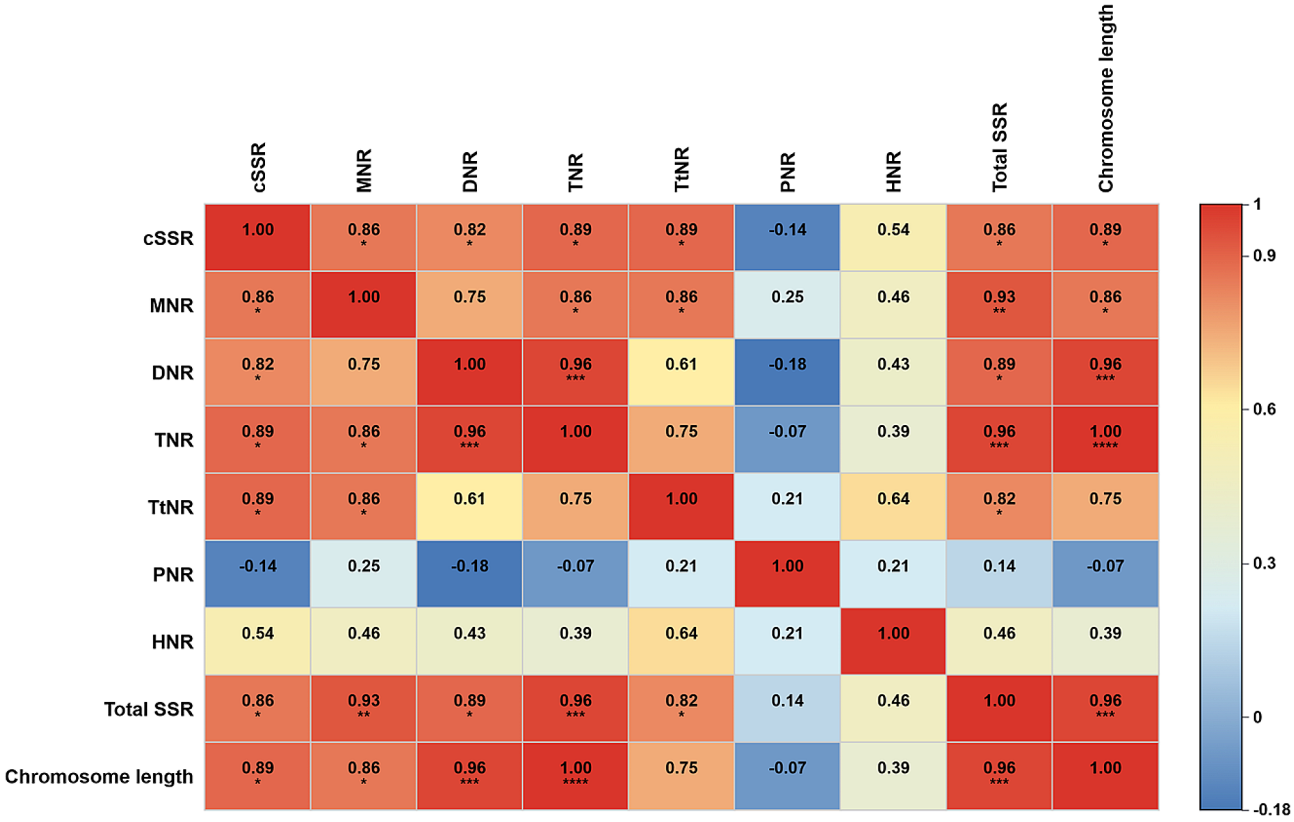

Fig. 7 The correlation coefficient between chromosome length and the number of SSRs

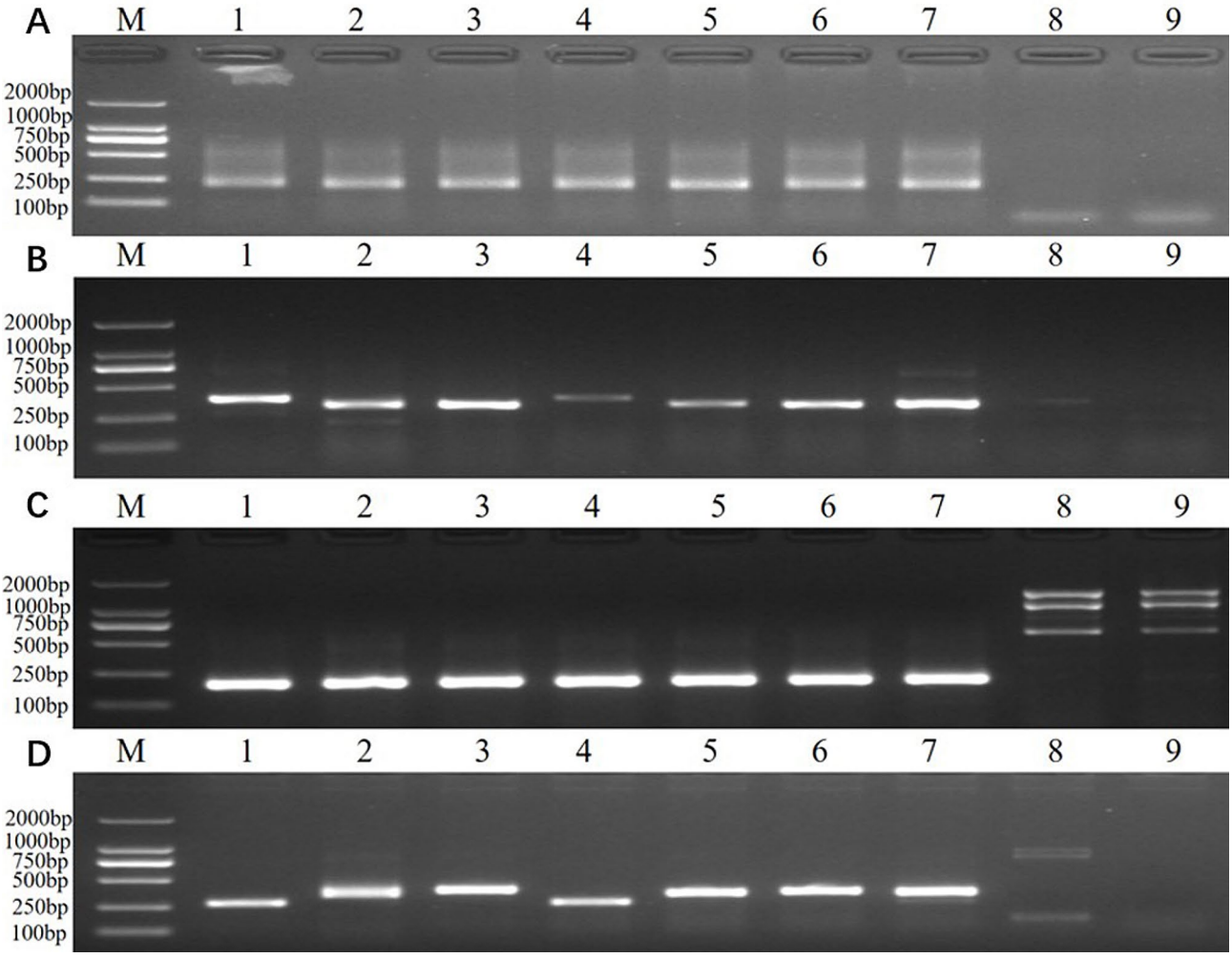

Fig. 8 PCR results of the several rye-specific primers designed in this study. **A**: PCR results of primer WN1R7.8; **B**: PCR results of primer WN3R13.2; **C**: PCR results of primer WN1R104.3; **B**: PCR results of primer WN6R7; M: marker DL2000; lanes 1–9: 'Weining', 'Kustro', 'Jingzhou, 'Qinling', 'Aigan', 'Baili', accession PI436168, CS, MY11

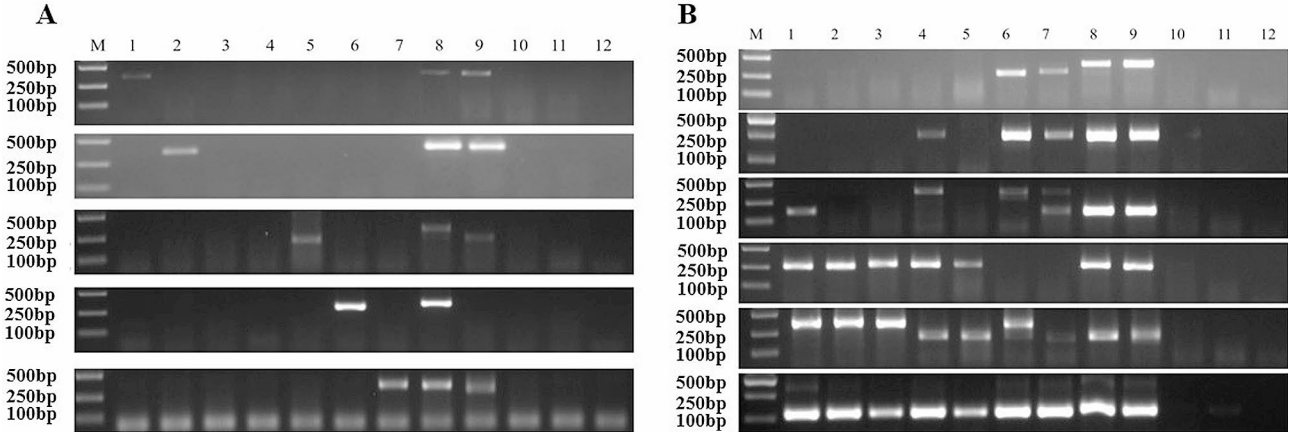

Fig. 9 PCR results for the several rye chromosome-specific primers designed in this study. **A**: From top to bottom: the PCR results for WN1R110, WN2R47, WN5R36, WN6R82, and WN7R61. **B**: From top to bottom: the PCR results for WN6R78, WN6R24, WN1R7.8, WN4R26, WN6R35.1, and WN1R42.3. M: DL2000 marker; lanes 1–7: wheat-rye 1R to 7R addition lines; lane 8: 'Weining'; lane 9: 'Kustro'; lane 10: CS; lane 11: MY11; lane 12: ddH_2O

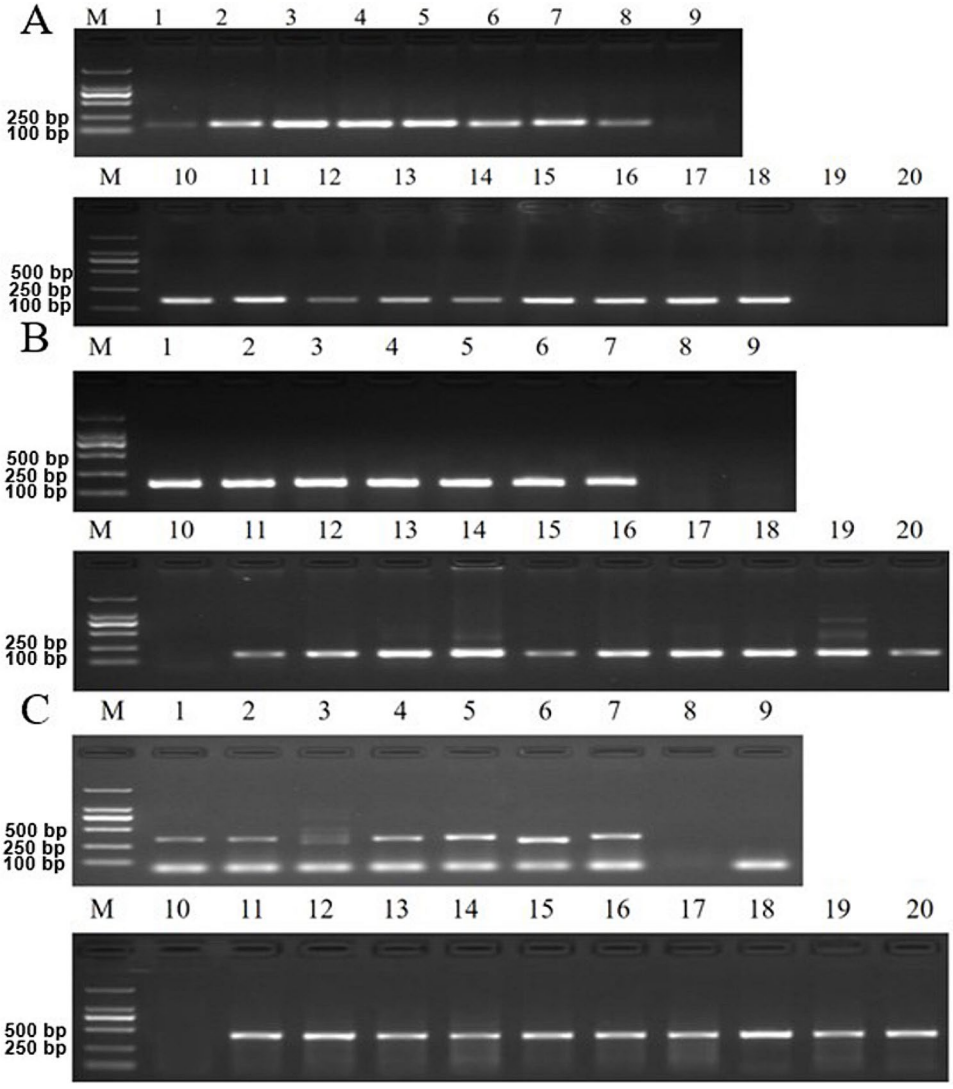

Fig. 10 The amplification results of several primers when they were tested with wheat-wild or wheat-weedy rye addition lines. **A**: The PCR products were amplified by the primer WN1R24.1. Lanes 1–10: 'Weining', 'Kustro', 'Jingzhou', 'Qinling', 'Aigan', 'Baili', accession PI436168, *S. dighoricum*, CS; *S. Segetale*; lanes 11–12: 1R addition lines developed from the cross of CN25 x *S. dighoricum*; lanes 13–14: 1RS·1BL translocation lines developed from the cross of CN25 x *S. dighoricum;* lanes 15–16: 2R addition lines developed from the cross of CN27 x *S. segetale*; lanes 17–18: 6R addition lines developed from the cross of CN25 x *S. dighoricum;* lane 19: CN25; lane 20: CN27. **B**: The PCR products were amplified by the primer WN1R7.8. Lanes 1–11: 'Weining', 'Kustro', 'Jingzhou', 'Qinling', 'Aigan', 'Baili', accession PI436168, CS, MY11, CN23, and *S. vavilovii*; 12–20: 1R addition lines developed from the cross of CN23 × *S. vavilovii*. **C**: The PCR products were amplified by the primer WN7R61. Lanes 1–11: 'Weining', 'Kustro', 'Jingzhou', 'Qinling' rye, 'Aigan', 'Baili', accession PI436168, CS, MY11, CN23, and *S. vavilovii*; lanes 12–20: the 7R addition lines developed by the cross of CN23 × *S. vavilovii*

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10689-1.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Acknowledgements

We thank the CNSknowall platform (https://cnsknowall.com/#/HomePage) for providing data analysis services.

Author contributions

TR and ZL designed the experiment; TR and ZL provided the plant materials used in this study. LZ did the SSR analysis; LZ, TY, JT, and YM did the molecular analysis; TR and ZL wrote the paper. All authors have read and approved the final version of the manuscript.

Funding
This study was funded by the Major Program of National Agricultural Science and Technology of China (NK20220607), the Foundation of State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China (SKL-ZY202219), and Foundation of Sichuan Province Science and Technology Support Program (2024NSFSC0316).

Data availability
The datasets generated and/or analyzed during the current study are available in the China National Center for Bioinformation. Genome assembly of Weining V1, https://ngdc.cncb.ac.cn/gwh/Assembly/12832/show.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 17 May 2024 / Accepted: 5 August 2024
Published online: 12 August 2024

References

1. Schlegel R, Korzun V. About the origin of 1RS.1BL wheat-rye chromosome translocations from Germany. Plant Breed. 1997;116:537–40.
2. Ren T, Chen F, Yan B, Zhang H, Ren Z. Genetic diversity of wheat–rye 1BL.1RS translocation lines derived from different wheat and rye sources. Euphytica. 2012;183:133–46.
3. Gong B, Chen L, Zhang H, Zhu W, Xu L, Cheng Y, Wang Y, Zeng J, Fan X, Sha L, Zhang H, Chen G, Zhou Y, Kang H, Wu D. Development, identification, and utilization of wheat–tetraploid Thinopyrum elongatum 4EL translocation lines resistant to stripe rust. Theor Appl Genet. 2024;137:17.
4. Han G, Liu S, Jin Y, Jia M, Ma P, Liu H, Wang J, An D. Scale development and utilization of universal PCR-based and high-throughput KASP markers specific for chromosome arms of rye (*Secale cereale* L). BMC Genomics. 2020;21:206.
5. An DG, Ma PT, Zheng Q, Fu SL, Li LH, Han FP, Han GH, Wang J, Xu YF, Jin YL, Luo QL, Zhang XT. Development and molecular cytogenetic identification of a new wheat-rye 4R chromosome disomic addition line with resistances to powdery mildew, stripe rust and sharp eyespot. Theor Appl Genet. 2019;132:257–72.
6. Ren T, Jiang Q, Sun Z, Zhao L, Peng W, Ren Z, Tan F, Luo P, Li Z. Development and molecular cytogenetic characterization of novel primary wheat-rye 1RS.1BL translocation lines from multiple rye sources with resistance to stripe rust. Plant Dis. 2022;106(8):2191–200.
7. Marais GF, Marais AS. The derivation of compensating translocations involving homoeologous group 3 chromosomes of wheat and rye. Euphytica. 1994;79:75–80.
8. Rahmatov M, Rouse MN, Nirmala J, Danilova T, Friebe B, Steffenson BJ, Johansson E. A new 2DS·2RL robertsonian translocation transfers stem rust resistance gene *Sr59* into wheat. Theor Appl Genet. 2016;129:1383–92.
9. Duan Q, Wang YY, Qiu L, Ren TH, Li Z, Fu SL, Tang ZX. Physical location of new PCR-Based markers and powdery mildew resistance gene(s) on rye (*Secale cereale* L.) chromosome 4 using 4R dissection lines. Front. Plant Sci. 2017;8:1716.
10. Xi W, Tang Z, Luo J, Fu S. Physical location of new stripe rust resistance gene(s) and PCR-based markers on rye (*Secale cereale* L.) chromosome 5 using 5R dissection lines. Agronomy. 2019;9:498.
11. Ren T, Sun Z, Ren Z, Tan F, Luo P, Tang Z, Fu S, Li Z. Molecular and cytogenetic characterization of a wheat-rye 7BS.7RL translocation line with resistance to stripe rust, powdery mildew and Fusarium head blight. Phytopathology. 2020;110(10):1713–20.
12. Ren T, Sun Z, Hu Y, Ren Z, Tan F, Luo P, Li Z. Molecular cytogenetic identification of new wheat-rye 6R, 6RS, and 6RL addition lines with resistance to stripe rust and powdery mildew. Front Plant Sci. 2022;13:992016.
13. Fu S, Chen L, Wang Y, Li M, Yang Z, Qiu L, Yan B, Ren Z, Tang Z. Oligonucleotide probes for ND-FISH analysis to identify rye and wheat chromosomes. Sci Rep. 2015;5:10552.
14. Masojć P, Myśków B, Milczarski P. Extending a RFLP-based genetic map of rye using random amplified polymorphic DNA (RAPD) and isozyme markers. Theor Appl Genet. 2001;102:1273–9.
15. Bednarek PT, Masojć P, Lewandowska R, Myśków B. Saturating rye genetic map with amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) markers. J Appl Genet. 2003;44(1):21–33.
16. Milczarski P, Banek-Tabor A. New genetic map of rye composed of PCR-based molecular markers and its alignment with the reference map of the DS2 × RXL10 intercross. J Appl Genet. 2007;48(1):11–24.
17. Bolibok-Bragoszewska H, Heller-Uszynska K, Wenzl P, Uszyński G, Kilian A, Rakoczy-Trojanowska M. DArT markers for the rye genome - genetic diversity and mapping. BMC Genomics. 2009;10:578.
18. Maria LCV, Luciane S, Augusto LD, Carla de FM. Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol. 2016;39(3):312–28.
19. Debrauwere H, Gendrel CG, Lechat S, Dutreix M. Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. Biochimie. 1997;79(9–10):577–86.
20. Kubis S, Schmidt T, Heslop-Harrison JS. Repetitive DNA elements as a major component of plant genomes. Ann Botany. 1998;82:45–55.
21. Mehrotra S, Goyal V. Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. Genomics Proteom Bioinf. 2014;12(4):164–71.
22. Zhong S, Chen W, Yang H, Shen J, Ren T, Li Z, Tan F, Luo P. Characterization of microsatellites in the *Akebia trifoliata* genome and their transferability and development of a whole set of effective, polymorphic and physically mapped SSR markers. Front. Plant Sci. 2022; 860101.
23. Li G, Wang L, Yang J, He H, Jin H, Li X, Ren T, Ren Z, Li F, Han X, Zhao X, Dong L, Li Y, Song Z, Yan Z, Zheng N, Shi C, Wang Z, Yang S, Xiong Z, Zhang M, Sun G, Zheng X, Gou M, Ji C, Du J, Zheng H, Doležel J, Deng X, Stein N, Yang Q, Zhang K, Wang D. A high quality genome assembly highlights rye genomic characteristics and agronomically important genes. Nat Genet. 2021;53:574.
24. Rabanus-Wallace MT, Hackauf B, Mascher M, Lux T, Wicker T, Gundlach H, Baez M, Houben A, Mayer KFX, Guo L, Poland J, Pozniak CJ, Walkowiak S, Melonek J, Praz CR, Schreiber M, Budak H, Heuberger M, Steuernagel B, Wulff B, Börner A, Byrns B, Čížková J, Fowler DB, Fritz A, Himmelbach A, Kaithakottil G, Keilwagen J, Keller B, Konkin D, Larsen J, Li Q, Myśków B, Padmarasu S, Rawat N, Sesiz U, Biyiklioglu-Kaya S, Sharpe A, Šimková H, Small I, Swarbreck D, Toegelová H, Tsvetkova N, Voylokov AV, Vrána J, Bauer E, Bolibok-Bragoszewska H, Doležel J, Hall A, Jia J, Korzun V, Laroche A, Ma X, Ordon F, Özkan H, Rakoczy-Trojanowska M, Scholz U, Schulman AH, Siekmann D, Stojałowski S, Tiwari VK, Spannagl M, Stein N. Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. Nat Genet. 2021;53:564–73.
25. International Wheat Genome Sequencing Consortium. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science. 2014;345:1251788.
26. Ren TH, Chen F, Zou YT, Jia YH, Zhang HQ, Yan BJ, Ren ZL. Evolutionary trends of microsatellites during the speciation process and phylogenetic relationships within the genus *Secale*. Genome. 2011;54:316–26.
27. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3-new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115.
28. Zhang Z, Deng Y, Tan J, et al. A genome-wide microsatellite polymorphism database for the indica and japonica rice. DNA Res. 2007;14(1):437–54.
29. Vinces MD, Legendre M, Caldara M, Hagihara M, Verstrepen KJ. Unstable tandem repeats in promoters confer transcriptional evolvability. Science. 2009;324(5931):1213–6.
30. Pearson CE, Nichol EK, Cleary JD. Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet. 2005;6(10):729–42.
31. Lam FH, Steger DJ, O'Shea EK. Chromatin decouples promoter threshold from dynamic range. Nature. 2008;453(7192):246–50.
32. Liu C, Zhang Q, Yao X, Zhong C, Yan C, Huang H. Characterization of genome-wide simple sequence repeats and application in interspecific genetic map integration in kiwifruit. Tree Genet Genomes. 2016;12:21.
33. Bhattarai G, Shi A, Kandel DR, Solís-Gracia N, da Silva JA, Avila CA. Genome-wide simple sequence repeats (SSR) markers discovered from whole-genome sequence comparisons of multiple spinach accessions. Sci Rep. 2021;11:9999.

34. Cavagnaro PF, Senalik DA, Yang L, Simon PW, Harkins TT, Kodira CD, Huang S, Weng Y. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L). BMC Genomics. 2010;11:569.
35. Mishra G, Meena RK, Kant R, Pandey S, Ginwal HS, Bhandari MS. Genome-wide characterization leading to simple sequence repeat (SSR) markers development in Shorea robusta. Funct Integr Genom. 2023;23:51.
36. Wang X, Yang S, Chen Y, Zhang S, Zhao Q, Li M, Gao Y, Yang L, Bennetzen JY. Comparative genome-wide characterization leading to simple sequence repeat marker development for *Nicotiana*. BMC Genomics. 2018;19:500.
37. Liu S, An Y, Li F, Li S, Liu L, Zhou Q, Zhao S, Wei C. Genome-wide identification of simple sequence repeats and development of polymorphic SSR markers for genetic studies in tea plant (*Camellia sinensis*). Mol Breeding. 2018;38:59.
38. Chen W, Yang H, Yang H, Zhong S, Dong Q, Chen C, Ren T, Li Z, Tan F, Luo P. Genome-wide SSR marker analysis to understand the genetic diversity and population sub-structure in *Akebia trifoliata*. Genet Resour Crop Evol. 2023;70:2741–54.
39. Song X, Yang Q, Bai Y, Gong K, Wu T, Yu T, Pei Q, Duan W, Huang Z, Wang Z, Liu Z, Kang X, Zhao W, Ma X. Comprehensive analysis of SSRs and database construction using all complete gene-coding sequences in major horticultural and representative plants. Hortic Res. 2021;8:122.
40. Zhao M, Shu G, Hu Y, Cao G, Wang Y. Pattern and variation in simple sequence repeat (SSR) at different genomic regions and its implications to maize evolution and breeding. BMC Genomics. 2023;23:136.
41. Han B, Wang C, Tang Z, Ren Y, Li Y, Zhang D, Dong Y, Zhao X. Genome-wide analysis of microsatellite markers based on sequenced database in Chinese Spring wheat (*Triticum aestivum* L). PLoS ONE. 2015;10(11):e0141540.
42. Zhang Y, Fan C, Chen Y, Wang RRC, Zhang X, Han F, Hu Z. Genome evolution during bread wheat formation unveiled by the distribution dynamics of SSR sequences on chromosomes using FISH. BMC Genomics. 2021;22:55.
43. Zhu S, Du H, Su F, Wang J, Meng Q, Liu T, Guo R, Chen Z, Li H, Liu W, Ma P, He H. Molecular cytogenetic analyses of two new wheat-rye 6RL translocation lines with resistance to wheat powdery mildew. Crop J. 2023;11:584–92.
44. Xu H, Yin D. Development and application of EST-Based markers specific for chromosome arms of rye (*Secale cereale* L). Cytogenet Genome Res. 2012;136:220–8.
45. Li JJ, Endo TR, Saito M, Ishikawa G, Nakamura T, Nasuda S. Homoeologous relationship of rye chromosome arms as detected with wheat PLUG markers. Chromosoma. 2013;122:555–64.
46. Li M, Tang Z, Qiu L, Wang Y, Tang S, Fu S. Identification and physical mapping of new PCR-Based markers specific for the long arm of rye (*Secale cereale* L.) chromosome 6. J Genet Genomics. 2016;43:209–16.
47. Qiu L, Tang ZX, Li M, Fu S. Development of new PCR-based markers specific for chromosome arms of rye (*Secale cereale* L). Genome. 2016;59(3):159–65.
48. Fandade V, Singh P, Singh D, Sharma H, Thakur G, Saini S, Kumar P, Mantri S, Bishnoi OP, Roy J. Genome-wide identification of microsatellites for mapping, genetic diversity and cross-transferability in wheat (*Triticum* spp). Gene. 2024;896:148039.
49. Bolibok H, Rakoczy-trojanowska M, Wyrzykowska M, Radecka M, Orczyk W. Identification of microsatellite markers in the rye genome. Cell Mol Biol Lett. 2006;11:291–8.
50. Schneider A, Molnár-Láng M. Polymorphism analysis using 1RS-Specific molecular markers in rye cultivars (*Secale cereale* L.) of various origin. Cereal Res Commun. 2008;36(1):11–9.

Publisher's Note