ISSN 2079-0597, Russian Journal of Genetics: Applied Research, 2016, Vol. 6, No. 8, pp. 825–837. © Pleiades Publishing, Ltd., 2016.
Original Russian Text © M.A. Nesterov, D.A. Afonnikov, E.M. Sergeeva, L.A. Miroshnichenko, M.K. Bragina, A.O. Bragin, G.V. Vasiliev, E.A. Salina, 2015, published in Vavilovskii Zhurnal Genetiki i Selektsii, 2015, Vol. 19, No. 6, pp. 707–714.

Identification of Microsatellite Loci Based on BAC Sequencing Data and Their Physical Mapping into the Soft Wheat 5B Chromosome

M. A. Nesterov[a], D. A. Afonnikov[a], E. M. Sergeeva[a], L. A. Miroshnichenko[b], M. K. Bragina[a], A. O. Bragin[a], G. V. Vasiliev[a], and E. A. Salina[a, *]

[a] *Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
[b] *Sobolev Institute of Mathematics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
**e-mail: salina@bionet.nsc.ru*

Received September 21, 2015; in final form, October 20, 2015

Abstract—The shortage of polymorphic markers for the regions of the wheat chromosomes that encode commercially valuable traits determines the need for studying the wheat microsatellite SSR loci. In this work, SSR markers for individual regions of the short arm of soft wheat chromosome 5B (5BS) were designed based on the sequence data obtained from BAC clones, and regions of the corresponding chromosome were saturated with these markers. Totally, 130 randomly selected BAC clones from 5BS library were sequenced using the IonTorrent platform and assembled in contigs using MIRA software. The assembly characteristics (N50 = 4136 bp) are comparable to the recently obtained data for wheat and related species and are acceptable for the identification of the microsatellite loci. The algorithm utilizing the properties of complex decompositions in the sliding-window mode was used to detect DNA sequences with a repeat unit of 2–4 bp. Analysis of 17770 contigs with a total length of 25879921 bp allowed for the design of 113, 79, and 67 microsatellite SSR loci with a repeat unit of 2, 3, and 4 bp, respectively. SSR markers with a motif of 3 bp were tested using nullitetrasomic lines of Chinese Spring wheat homoeologous group 5. In total, 21 markers specific for chromosome 5B were identified. Eight of these markers were mapped into the distal region of this chromosome (bin 5BS6) using a set of Chinese Spring deletion lines for 5BS. Eight and four markers were mapped to the interstitial region (bins 5BS5 and 5BS4, respectively). One marker was mapped to a pericentromeric bin. Comparative analysis of the distribution of trinucleotide microsatellites over wheat chromosome 5B, and in different cereal species, suggests that the $(AAG)_n$ repeat proliferates and is conserved during the evolution of cereals.

Keywords: bread wheat, *Triticum aestivum*, BAC clones, Ion Torrent, MIRA, microsatellites, chromosome 5B, SSR markers, $(AAG)_n$

DOI: 10.1134/S2079059716070078

The physical mapping of chromosomes, namely, direct localization of DNA sequences on the chromosomes, was first developed with the advent of in situ hybridization methods. Physical mapping can be successfully carried out with the use of a variety of genetic resources of plants, comprising aneuploid, substitution, deletion, and introgression lines. In this respect, it is also referred to as cytogenetic mapping (Sourdille et al., 2004). In the age of genomic sequencing, the term "physical mapping" is increasingly being used to denote the localization of extended DNA contigs or BAC clones in a certain section of a chromosome. In this case, localization methods may be different, including, among others, bioinformatic approaches (Paux et al., 2008).

Soft wheat *Triticum aestivum* L. ($2n = 42$) is a naturally occurring genomic allopolyploid with a genome formula BBAADD (Feldman, 2001), which was formed, presumably, by interbreeding diploid species *Triticum* and *Aegilops*. E. Sears (Sears, 1966) showed that chromosomes of the three genomes of soft wheat can be divided into seven homoeologous groups, within which one chromosome in the extra-(tetra) dose compensates for the disturbances caused by the absence of the other. The ability of the homoeologous chromosomes to serve as a buffer during the loss of chromosomes, or their fragments, led to the creation of collections of aneuploid and deletion lines, which form the basis for a new stage in the development of genetic and molecular studies on the analysis of the wheat genome.

The first stage of such works has been associated with the creation of a series of aneuploid soft wheat lines to study the genetic contribution of each chromosome in the inheritance of the various traits of wheat, localization and distribution of genes and markers in linkage groups, and analysis of interactions of individual genes in the formation of traits (Plaschke et al., 1996).

Development of new collections, based on the stability of the genome of polyploid wheat forms during the loss of individual parts of the chromosomes, was accompanied by the creation of a series of deletion lines. Collections of these lines for each chromosome arm have

been thoroughly characterized by cytological methods (http://www.ksu.edu/wgrc/Germplasm/Deletions/). Then, the creation of overlapping deletion lines became a necessary step for the analysis of the plant genome. In recent years, this collection is used intensively by International Wheat Genome Sequencing Consortium for the construction of physical maps of individual wheat chromosomes (http://www.wheatgenome.org/Projects/IWGSC-Bread-Wheat-Projects/Physicalmapping).

Despite the fact that the number of markers for the wheat genome has increased substantially in the last five years, particularly with the advent of the first results of its sequencing and development of SNP-chips to analyze wheat genome, the task of developing new markers is still relevant (Akhunov et al., 2009; Brenchley et al., 2012). A comparison of the physical and genetic maps of wheat chromosomes points to the unequal distribution of molecular markers on the chromosomes (Sourdille et al., 2004; Timonova et al., 2013). Microsatellite loci, or tandem repeats with repeat units from 2 to 6 bp in length (Simple Sequence Repeat, SSR) can be used successfully to develop SSR-markers. The advantage of these markers is in their high level of polymorphism and stability of the results; their drawback lies in their lower frequency of occurrence relative to the SNP- and ISBP-markers (Li et al., 2002).

The short arm of chromosome 5B has a length of 290 Mb. This chromosome contains a number of genes that determine resistance to fungal pathogens, crossability with rye, the softness of grain, and resistance to the Hessian fly (http://www.shigen.nig.ac.jp/wheat/komugi/genes/symbol/ClassList.jsp). Currently, the number of markers on the physical and genetic maps of chromosome 5BS is not sufficient for gene labeling, positional cloning, or marker-assisted selection.

The development of markers using BAC clones is based on two objectives, one of which is the enlargement of the pool of markers, and the second objective is the development of markers directed to BAC clones for subsequent localization on the chromosome. BAC-end sequencing has been carried out actively in recent years to address these problems (Paux et al., 2008). However, the use of this method allows only for an extremely low efficiency of marker development due to restrictions on the length and number of decoded sequences (no more than two) for each individual BAC clone.

This work was aimed at the development of the new SSR-markers based on high-performance sequencing pools of BAC clones containing the short arm of the 5B chromosome, and their localization on the physical map of the chromosome.

MATERIALS AND METHODS

Plant Material

We used a soft wheat variant Chinese Spring; nullitetrasome lines N5BT5A (chromosome 5B substituted by 5A) and N5BT5D (chromosome 5B replaced by 5D); ditelosomic line Dt5BL (no short arm 5BS); and a set of ten soft wheat deletion lines for chromosome 5B of the Chinese Spring variety, which have been obtained and described previously (Endo and Gill, 1996). The deletion lines define seven deletion areas (bins) 5BS: C–5BS3–0.41; 5BS3–0.41–0.42; 5BS2–0.42–0.43; 5BS4–0.43–0.56; 5BS8–0.56–0.71; 5BS5–0.71–0.81; and 5BS6–0.81–1.00. This range of deletion lines and bins is described in detail by L. Qi and colleagues (Qi et al., 2003). Deletion lines were kindly provided by the center of genetic resources of the University of Kansas (http://wwwk-state.edu/wgrc/Germplasm/Deletions/del_index.html.). The lines were propagated and screened for the presence of deletions (Timonova et al., 2013).

BAC Clones

BAC clones containing 5BS were made at the Institute of Experimental Botany (Olomouc, Czech Republic) and kindly provided by Professor J. Doležel.

Isolation of DNA

Isolation and purification of DNA from plants was carried out according to a procedure described previously (Plaschke et al., 1995).

DNA from BAC clones was isolated using the NucleoSpin 96 Flash (Macherey-Nagel) kit, which is based on SDS-alkaline lysis. This technique has been adapted to the existing equipment with the inclusion of an additional purification step by centrifugation.

PCR Analysis

Analysis was performed based on a reaction mixture of 20 μL containing 50 ng of template DNA, 65 mM Tris-HCl (pH 8.9), 1.5 mM $MgCl_2$, 16 mM $(NH_4)_2SO_4$, 0.05% Tween 20, 0.2 mM of dNTP, 1 unit of DNA polymerase Taq, and 0.25 pM of specific forward and reverse primers (Supplemental Material 1[1]). The following cycle of reactions was carried out: denaturation for 30 s at 94°C; annealing for 30 s at 55°C; polymerization for 30 s at 72°C; and the number of cycles was 35. PCR was performed in a thermocycler BioRad T100 (United States). PCR fragments were separated by electrophoresis on a 2% agarose gel or 10% polyacrylamide gel, according to the level of polymorphism between fragments.

Sequencing on an IonTorrent Platform

DNA from 130 individually selected BAC clones was pooled in approximately equimolar amounts. To prepare the libraries using IonTorrent, 5 μm g of DNA was frag-

[1] For Supplemental Materials 1–3, see Appendix 5 online at http://www.bionet.nsc.ru/vogis/download/pict-2015-12/appx5.pdf.

mented by ultrasound using a Covaris device, fragments of 330–450 bp were isolated with the LabchipXT device, then, the DNA ends were repaired, followed by the ligation of the adapters using the Ion Plus Fragment Library Kit. The library was processed through 8 cycles of amplification, cleaned using AmPureXP kit, and checked using the Bioanalyzer2100 device. Next, emulsion PCR was carried out using the Ion PGM Template OT2 400 Kit, followed by sequencing using the Ion-Torrent device for the estimated length of fragments at 400 bp with a maximum capacity chip of 318.v2.

For the assembly of DNA sequences, we used the MIRA program (Chevreux et al., 1999). Evaluation of the depth of coverage of the genome assembly and determination of the share of duplicate PCR products was carried out based on the mapping of the initial sequenced fragments longer than 30 bp in length on the sequences of contigs using program Bowtie2 (Langmead and Salzberg, 2012) with further processing using the equalization program Samtools (Li et al., 2009).

Identification of SSR-loci, development of markers

To identify sequences with a repeating motif of 2–4 bp, we used an algorithm based on the properties of complex expansion executed in a sliding window mode (Gusev et al., 2009).

Development of primers into the repeating 2–4 bp motifs (SSR-loci) was carried out using PrimerQuest program (http://eu.idtdna.com/PrimerQuest/Home/Index) with the following preset parameters: length of the amplicon (150–450 characters); primer length (17–25 characters); (C + G) composition (40–60%); Tm (59–65°C); and G or C at the 3'-end of the primer. Other PrimerQuest parameters were used by default.

Of the several (up to five) primer pairs, proposed by the PrimerQuest program, a single pair was selected for each SSR-locus. Preference was given to highly complex primers: a linguistic or DNA-oriented measure of LZ (Gusev et al., 1999). Primers with nonequilibrium composition of nucleotides and with long stretches of the (G + C), (A + G), (A + C), (A + T), (C + T), and (G + T) regions, as well as primers with repetitive motifs, were excluded.

RESULTS

Sequencing of BAC Clones and Contig Assembly

The number of clones selected for sequencing was designed based on at least a one hundredfold reading of the DNA sequences under study, the average length of BAC clone equaling about 100 Kb, and the use Ion-Torrent chip size of the maximum capacity. Thus, the obtained number of BAC clones amounted to 130.

During sequencing of the pools of BAC clones, the results in respect to the density of application and quality of sequencing were reported to be above average when using this method. The distribution of readout fragment lengths (Fig. 1) showed that most had a length of 150–400 bp.

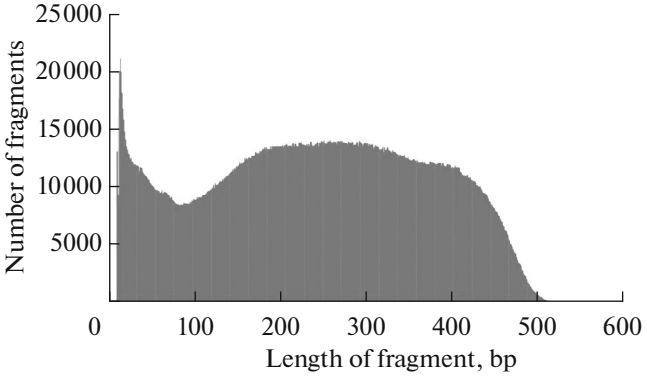


Fig. 1. Distribution of the sequenced fragments of BAC clones' library by length.

The total amount of processed data amounted to more than 5.5 million individual readouts with an average length of 239 bp and a total of more than 1.3 billion nucleotides. The proportion of sequences with $Q > 20$ was 74.6%. As result, we obtained on average a hundredfold excess of the estimated readouts of BAC clones selected for the analysis.

One of the first steps in the assembly of DNA sequences after the sequencing of BAC-clones is the removal of the sequences of the DNA vector, and the DNA of *E. coli*. Raw data revealed less than 2.2% of the *E. coli* (*Escherichia coli* str. K12 substr. DH10B) DNA, pointing to the high quality of the isolated BAC-clones.

The results of sequence assembly after removal of the *E. coli* DNA using the MIRA program are presented in the table. Note that the characteristics of contigs with a length of 500 bp (~5000 contigs) almost double for parameter N50, and more than double for parameters N90 and N95.

For the further evaluation of the quality of sequencing of the derived libraries, we evaluated the PCR products for the duplicate content and average depth of coverage of the DNA sequence of the readout BAC-clones. To this end, the collected contigs were mapped with original readouts (longer than 30 bp) and quality characteristics were calculated based on the alignment data. We found that the percentage of duplicate PCRs was around 14%, and the average depth of coverage was about 25 reads for each position in the contig. In addition, the MIRA program estimates the average depth of coverage for contigs with the average length of 500 bp and coverage values of not less than 15. The value of the average coverage for such high quality contigs was ~40 reads per position. These estimates allowed us to conclude that the libraries and contigs assembled based on the sequencing are suitable for solving the problem of identifying SSR-markers in the sequences of the BAC-clones.

Identification and Labeling of SSR-Loci

The starting material (genomic DNA sequence) contains repetitive fragments. Some sequences are repeated as whole, while others are fragments of longer sequences. As a result, part of the DNA sequence comprising the SSR-loci is also repeated. To identify repetitions, a program was developed based on the search algorithm for samples, which allows us to remove identical SSR-loci from the analysis.

In the first step, according to the analysis of 17770 contigs of a total length of 25879921 bp, we found 253, 156 and 27 DNA sequences containing SSR-loci with a period of repetition of 2, 3 and 4 bp, respectively, and a minimum length of 20 nucleotides. The second step consisted in the removal of repetitive motifs, located at the ends of the contigs, repetitive SSR-loci, and loci whose environment does not allow the development of primers. As result, 113 primers were selected for SSR-loci with a repetative motif of 2 bp; 79, for SSR-loci with a 3 bp motif; and 23, for SSR-loci with a 4 bp motif.

For the 67 pairs of primers chosen for the SSR-loci with a repetitive motif of 3 bp, PCR was performed with template DNA from nullitetrasome and deletion lines, and the possibility of their use as markers at different positions of the soft wheat chromosome 5B was determined (Supplemental Materials 1).

Evaluation and Physical Mapping of New SSR-markers

PCR results using the DNA template from Chinese Spring (CS) wheat variety with primers, developed into SSR-loci with a repetitive motif of 3 bp, are shown in Supplemental Material 2. The amplification products were clear and well detectable in SDS and agarose gels for the majority of the examined markers. Only four (Xicgc740, Xicgc1738, Xicgs1755, Xicgc1931) of the 67 studied SSR-markers gave weak amplification fragments under standard PCR conditions.

The physical mapping of markers on the 5BS chromosome was performed using seven deletion lines: 5BS6, 5BS5, 5BS1, 5BS8, 5BS4, 5BS2, and 5BS3 (Fig. 2; Supplemental Material 3). As control, we used soft wheat of the CS variety and a null- tetrasome line CSN5BT5A (chromosome 5B substituted for 5A).

Figure 2 shows the location of breakpoints in seven deletion lines for the short arm of the 5B chromosome. For example, deletion line 5BS6, lacks the distal end of the chromosome up to the breakpoint at 5BS6 (deletion bin 5BS6), while in the deletion line 5BS5, the deletion is larger and extends to the breakpoint at 5BS5 (deletion bin 5BS5). Analysis of the series deletions in wheat lines allows us to localize the markers to a particular area (bin) on the chromosome, i.e., carry out their physical mapping. Examples of the localization of markers in bins 5BS6, 5BS5, 5BS4, and C-5BS3 are shown in Fig. 2 and in Supplemental Material 3.

In total, comparative analysis of PCR fragments in the CS and CSN5BT5A lines identified 21 SSR-markers on chromosome 5B. The remaining 46 markers were similar in the length of the amplification fragments for chromosomes of the 5th homoeologous group. Analysis of 21 polymorphic SSR-markers using a series of CS deletion lines for chromosome 5B allowed us to localize these markers on the physical map of 5BS. Thus, eight markers (Xicgc178, Xicgc178_2, Xicgc284, Xicgc456, Xicg14s009, Xicg15s020, Xicg16s004_2, Xicg16s041) were mapped on to the distal region of the chromosome (bin 5BS6); eight markers (Xicgc122, Xicgc131, Xicgc229, Xicgc299, Xicgc307, Xicgs342, Xicgc686, Xicgc1572), on to bin 5BS5; four markers (Xicgc495, Xicgc498, Xicgc69, Xicgc1699), on to bin 5BS4; and Xicgc1988, on to bin C-5BS3 (Fig. 2).

DISCUSSION

The Approaches and Impact of the Assembly of DNA Sequences in Cereals

To build a genome, we used the MIRA program (Chevreux et al., 1999). On the one hand, the choice was due to the fact that this program has proved itself quite well in the assembly of bacterial genomes based on the libraries obtained using IonTorrent technology (Loman et al., 2012). In addition, the program has been used previously in the assembly of a number of plant genomic fragments, in particular, for the assembly of BAC-clones for barley (Pasquariello et al., 2014) and sunflower (*Helianthus annuus* L.) (Staton et al., 2012), for the assembly of genome fragments of the Barbados nut (*Jatropha curcas*) (Sato et al., 2011), as well as the plastid genome of the bird's-nest orchid (*Neottia nidus-avis*) (Logacheva et al., 2011) (in all these cases, sequencing was performed using the Roche 454 technology). In the review by Stein and Steuernagel (Stein, Steuernagel, 2014) dedicated to the works in the field of sequencing of the barley genome, the MIRA program is described as very successful for the analysis of relatively small amounts of data received on different platforms (Illumina GAIIx, Hiseq, Roche 454) and containing a large number of repetitions.

In this study, we were able to obtain the results of genomic fragments of an assemble of sufficiently high quality (table). It is interesting to compare the characteristics of the thread assembly with the analogous data on the decoding of genomic sequences of postglacial plants, especially, related to wheat. The analysis of recent works devoted to the sequencing of the genomes of wheat and related species using shot-gun technology (shot-gun sequencing) showed that the characteristics of the resulting assembly of BAC-clones are in quite good agreement with the results of other groups. In particular, one of the most informative characteristics of the overall assembly quality is the N50 parameter (the value at which contigs of equal and greater length make up half of the total length of the assembly). In this work, it was about 4 Kb,

and in the analysis of contigs with a length of more than 500 bp this value was 7700 bp (table).

The assembly of the soft wheat genome based on low coverage sequencing (~5 ×) using 454 technology (Brenchley et al., 2012) is characterized by the N50 value for contigs of about 900 bp. To build the barley genome, sequencing by Illumina GAIIx and a 50-fold coverage (International Barley Genome Sequencing Consortium, 2012), the value of this parameter for the contigs was 1425 bp. To assemble the genome of *Aegilops tauschii* sequencing by Roche 454 and 90-fold coverage, the N50 contigs value was ~4500 bp (Jia et al., 2013). To assemble the genome of *Triticum urartu*, the contigs value of N50 was 3400 bp during sequencing by Illumina HiSequation (2000) and 91-fold coverage (Ling et al., 2013). In the study devoted to the sequencing and assembly of sequences of the soft wheat chromosomes using Illumina technology and coverage inbetween 30 × 241, the value for the N50 parameter ranged from 0.5 to ~4.3 kb for different chromosomes (International Wheat Genome Sequencing Consortium, 2014). For the recent wheat genome assembly, obtained by Chapman and colleagues using the Illumina method on multiple wheat strains with a total ~175-fold coverage, N50 for contigs with a length of more than 1 Kb was determined by the value of ~8300 bp (Chapman et al., 2015). Thus, the characteristics of the assembly performed in our study are, in general, in good agreement with those of the recent versions of the genome assemblies (Jia et al., 2013; Ling et al., 2013; Chapman et al., 2015) and are sufficient to address the designated work tasks.

It should also be noted that virtually for all primers designed based on the results of contig assembly, good amplification products were obtained on BAC-clones used for sequencing and on the wheat genomic DNA.

Features of the Organization of the Microsatellite Loci on the 5B Chromosome

Microsatellite DNA sequences are present at a high frequency in all the studied eukaryotic genomes (Tautz and Renz, 1984; Li et al., 2002). The average frequency of the occurrence of microsatellites in the eukaryote genome is assessed as 1 per 10000 bp (Brown et al., 1996). The number of SSR copies may vary up to 100 or more repetitions at one locus (Tautz and Renz, 1984; Li et al., 2002). The first estimates of the chromosome distribution of the microsatellite loci have been made as a result of mapping the plant genome, and based on the direct in situ hybridization. Studies of the distribution of microsatellites for some species, for example, sugar beet, by in situ hybridization, have shown that they are grouped in specific regions of chromosomes (Schmidt, Heslop-Harrison, 1996). Mapping tomato microsatellites revealed that they are localized in the centromeric regions of the chromosomes (Areshchenkova and Ganal, 1999). However, the majority of microsatellite loci in plant species have a more uniform distribution in the genome that contribute to the excellence of their use as markers for genotyping of the plant genome and their subsequent use in the marker-assisted selection (Mason, 2015; http://maswheat.ucdavis.edu/protocols/).

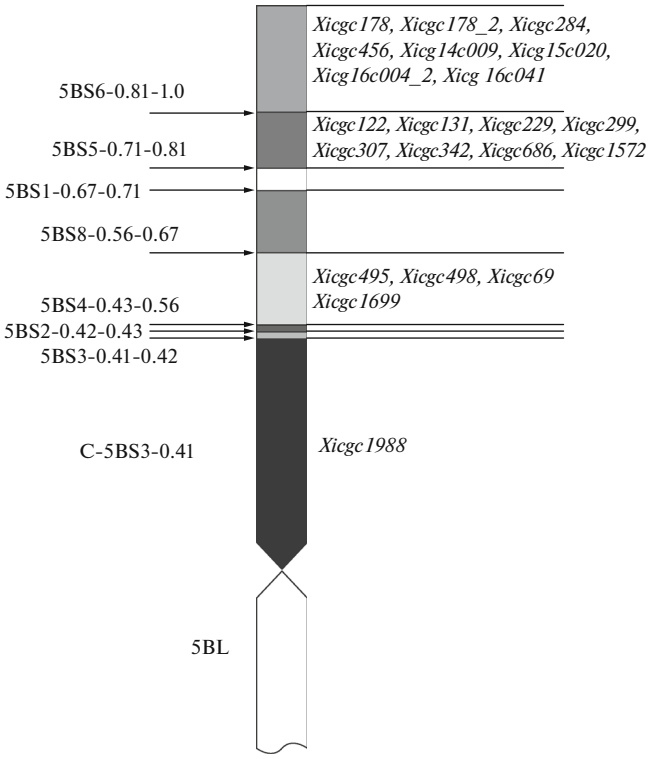


Fig. 2. Localization of SSR-markers in the deletion bins on the 5SB chromosome.

Analysis of the sequencing of BAC clone samples led to the development of an additional 67 markers for the short arm of the 5B chromosome, which contains a number of genes that influence the formation of the soft wheat quantitative traits and its resistance to abiotic and biotic environmental factors (http://www.shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp). Notably, 21 markers revealed fragments of amplification specific for the short arm of chromosome 5B, which allowed us to locate them on the physical map of chromosome 5B (in the deletion bins). The greatest num-

Main characteristics of the reassembly BAC-clones of a library

Characteristics of assembly	All contigs	Contigs longer than 500 bp
Number of contigs	17770	4780
Overall length, bp	25879921	18560927
Length of the longest contig, bp	250809	250809
N50, bp	4136	7767
N90, bp	504	1431
N95, bp	416	942

ber of markers among the ones localized on to the chromosome (76%) were present in the distal portion of the chromosome (bins 5BS6, 5BS5) and only one marker was present in the pericentromeric bin of the 5BS chromosome (Fig. 2; Additional Material 2). Analogous results of marker distribution were obtained by us earlier for the localization of SSR-markers obtained from the database on wheat chromosome 5B (Timonova et al., 2013). However, in contrast to the earlier results, in the present study, 19% of the markers were localized in the interstitial bin 5BS4. Mapping did not reveal any dependence on the structure of the microsatellite locus or its localization on the chromosome (Supplemental Material 2). At the same time, direct hybridization of different groups of microsatellites on the chromosome indicates the preference of their location to the pericentric region (Cuadrado et al., 2000, 2008). The opposite pattern was identified by us in the physical mapping of the B-specific chromosomal loci, indicating that, apparently, the centromeric regions in homoeologous chromosomes are rather conservative, including in group 5 of wheat chromosomes, and the absence of polymorphism made it impossible to locate the majority (70%) of the developed SSR-markers on chromosome 5BS.

The prevalence of different microsatellites in a the whole genome within a species depends on the structure of the repeating unit and its length. For instance, analysis of the results of sequencing the individual chromosomes, or the entire genome, of a number of cereals and maize (Zhang et al., 2007; Qu and Liu, 2013; Sergeeva et al., 2014) indicates that the dinucleotide repeats are most frequent, then, the frequency of occurrence decreases with the increasing length of the repeating motif. Our results on the analysis of the microsatellite loci based on the sequencing of a sample of 130 BAC clones also point to a higher prevalence of dinucleotide sequences (253 sequences) compared to tri- (156) and tetra-nucleotide (27) repetitive DNA sequences. The observed trend in the prevalence of the repeat motifs depending on length corresponds to the expected distribution of the motif that is inversely proportional to its length. The occurrence of di-, tri- and tetranucleotide repeats also varies depending on the portion of the genome. Assessment of the distribution of di- and trinucleotide repeats in translated and untranslated regions of the gene loci revealed that more than 87% of the microsatellites found in the coding regions of genes consist of trinucleotides, and this is characteristic for instance both for rice and corn (Zhang et al., 2007; Qu and Liu, 2013). This is due to the fact that the increase in the copy number of trinucleotide repeats does not shift the reading frame and, in some instances, may be maintained by natural selection. Species differences in cereals often consist in the spread of different trinucleotide repeats. For instance, according to the analysis of DNA sequences of the entire genome of maize, the most common trinucleotide repeat has the structure $(AGC)_n$, followed by repeats $(ACG)_n$, $(CCG)_n$, $(ATC)_n$, and $(AAG)_n$ with frequencies of occurrence ranging from 2.81 to 2.48% of the total number of di- and trinucleotide repeats (Qu and Liu, 2013). In rice, distribution of trinucleotide repeats in terms of the frequency of the occurrence e is somewhat different (Zhang et al., 2007). The most common is $(CCG)_n$, then $(AGG)_n$, $(AGC)_n$, $(ACG)_n$, and $(AAG)_n$. The major differences between species are observed in two groups of microsatellites, namely, in maize, $(ATC)_n$ are repeats common, while $(AGG)_n$ are significantly less frequent. The converse is true in rice. The common characteristic of the studied genomes of cereals (rice, maize) is a broad representation of the $(AGC)_n$, $(ACG)_n$, $(CCG)_n$, and $(AAG)_n$ repeats. It should be noted that the groups of repeats listed above do not include trinucleotide motifs representing the stop codon in the eukaryotic nuclear genome.

In wheat and its progenitors, the wide distribution in the genome is attributed to the $(AAG)_n$ repeat. This is shown by our results (Supplemental Materials 2) and by earlier studies (Cuadrado et al., 2000, 2008; Adonina et al., 2015) on the distribution of the trinucleotide microsatellites by the method of direct localization of oligonucleotides on the chromosome. Another feature of the soft wheat and the predominantly of B genome progenitor is a wide representation of the $(AGG)_n$ repeat (Supplemental Material 2; Cuadrado et al., 2008). The $(AGG)_n$ microsatellite is also widely distributed in rice (Zhang et al., 2007). This mosaic proliferation of the $(AGG)_n$ repeat in genomes of certain types of cereals in the process of evolution may have been determined by mobile elements.

Thus, the comparative analysis of the distribution of trinucleotide microsatellites in the 5B chromosome of wheat and other cereals in species points to the conservation of the quantitative content of the $(AAG)_n$ repeat in the evolution of cereals, while the content of other trinucleotide repeats varies significantly in the modern cereal species and in their progenitors.

ACKNOWLEDGMENTS

The authors thank the staff of the Centre for Structural and Functional Genomics of Plants, and its director J. Doležel (Institute of Experimental Botany, Olomouc, Czech Republic) for the provision of BAC-clones. This work was supported by the Russian Ministry of Education (agreement nos. 14.604.21.0106 of 07.07.2014, identification no. RFMEFI60414 X0106). These studies involved the Bioinformatics Core and the Genomic Research Core of the Siberian Branch, Russian Academy of Sciences (ICG SB RAS).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

Supplement 1. Designed primers to SSR loci

No.	Marker	Motif	Primers	SSR length, bp
1	*Xicgc69*	cca(30)	F-TTTCCTCGCTCTTCACTTCTC R-ACCAGAATGCCTGGTCTATTG	264
2	*Xicgc122*	ata(21)	F-GGACGGAGGAAGTACAAAGAAAG R-AAGTCCAGCACAGCAGAATC	311
3	*Xicgc128*	gag(32)	F-CAACCTCTCCTTTCCTCTTCTC R-TGTCTGCCATGACATCCTTC	256
4	*Xicgc131*	ttg(23)ttg(20)	F-GACCTGAAGGATTGGTGTATC R-TTTCTCACAGTAAAGTCTGGAG	260
5	*Xicgc178*	gtc(23)	F-GGGTTGGTTGGACATTTGATTC R-ACTCTGAAGCTGGTGGTTTG	317
6	*Xicgc178_2*	ttc(29)	F-AGCTGCTCAAGGACAACAG R-ATTAGTCGAGGAGGGATAGAGG	449
7	*Xicgc195*	gcc(23)	F-GGTCTTACAGTTGCGGTAGAAG R-ACACACACATTGGACCCTAAC	226
8	*Xicgc220*	atg(25)	F-GATTGGCTTCCGCTGTATTG R-CTCTTGGGAGTAGGAGAACTTG	431
9	*Xicgc229*	ttg(27)	F-GACATGGGAGGAGTCTGTTAAG R-ACTAGACCCAGAGGTAGGTAATC	427
10	*Xicgc284*	tgt(27)	F-GACATGGGAGGAATCCGTTAAG R-ACGTATGGGTGAGGGAAATAATC	279
11	*Xicgc299*	ctt(29)	F-AGTATCAGTGGGCGGTTAAAG R-CGGCCTTGACAACACTAAAC	364
12	*Xicgc307*	ttc(20)	F-AACTTACGGCACGAGGTTAG R-ACAAATTGGGAAAGATGGAAGTTAG	341
13	*Xicgc312*	tat(26)	F-TGAGCTTACAACACTCTGTCTG R-TTCAAGGTGACCACATGGATAG	267
14	*Xicgc342*	tcc(20)	F-CATGCAGGGTTCAAGTACAAAC R-GAAGACACTAGTAGCATGGTGAG	406
15	*Xicgc374*	ttg(33)	F-CGTTAAGAGAGACCTGAAGGATTG R-CAAACAGGGCCGATATCAGTAG	402
16	*Xicgc456*	gct(24)	F-TGGCGCACAGAAGAAGAC R-CTGTCTCTCAAGGTACACACAC	270
17	*Xicgc482*	ttg(21)	F-CTGCGATGAGATTGACAGGATAG R-CCGTCGAAGGATGAAGAAGAAG	277
18	*Xicgc495*	gag(29)	F-GTCTCCTCTCACAAGGAAGATG R-CTCTCCAGGTGCTAGACAATTC	365
19	*Xicgc498*	cta(91)	F-CCAATTATGCTCTTAGCCGTTTC R-ACACCAAATTCTATAGGCTCCTC	228
20	*Xicgc532*	caa(22)	F-CAGATTCCTCCGAGACATGATAC R-GGTCAGCTCCCTTCCAATC	252
21	*Xicgc619*	cga(20)	F-CTACCTCCGCGACAAGATG R-ACACGGCGCAAATTCAAC	392
22	*Xicgc642*	tcc(23)	F-GCGGCAGAAACAAAGCAAAG R-AGCGGGAGTGAGGAGAAAG	395

Supplement 1. (Contd.)

No.	Marker	Motif	Primers	SSR length, bp
23	*Xicgc667*	tca(37)	F-CAATCGACTTCCGAGAGATGAC R-GCCAGAGAGGACCAAAGATG	449
24	*Xicgc686*	ttc(23)	F-AGTATGGTGGGCGGTTAAAG R-GTAGCCGTAGTCATGTGTGTATAG	437
25	*Xicgc707*	cat(23)	F-CAACTCCAAATGGGCTGAAAC R-GTCAACCACAAGCACCAATC	397
26	*Xicgc709*	tcc(20)	F-TGGCCTTTCCTGCCTAAAG R-CTCGCTCCCACCATGATAAG	323
27	*Xicgc740*	tgc(23)	F-TTTCCTGGCGAAGTGAAGTC R-AGATGGAGGAGGAGAAGAAGAG	154
28	*Xicgc936*	atg(25)	F-CTGGGACTTGCCATACTAGAAG R-TCGATCTCCTCCTGACTTAGAG	309
29	*Xicgc967*	tcc(23)	F-ACCCAGAAACAGTCGGAATC R-CCAGGTCACTATCGAGATGTATTC	355
30	*Xicgc967_2*	tcc(20)	F-GTCGGAATCTCCATCCTCATC R-CAGGTCACTACCGAGATGTATTC	371
31	*Xicgc1382*	gag(27)	F-ACATCATCGCTACAGGATCAAC R-TTCTTCATAAGCGACGACCAC	375
32	*Xicgc1427*	ctg(20)	F-ACAAGCCAGAGAAGCAGTAAG R-GAGGACTTTGATGGAGCAGTAG	207
33	*Xicgc1453*	aca(20)	F-TGTTCGAAACAAAGCTCGATAAC R-CAACCTCATCCTCGTAATCTCC	258
34	*Xicgc1531*	atg(20)	F-CTAGTGACCAATTTGCTATGTGATG R-ACCCACGCTACTGCTAAATC	423
35	*Xicgc1550*	gaa(24)	F-CAAGGTGCTGAAGAAGGAGTATAG R-CTGGTGTTCCTCTTGGGTTTAG	444
36	*Xicgc1558*	tct(20)	F-CGTGGCAGAATAGTCCAGAAG R-GTACAAGGTGCCAGAGAAGAAG	394
37	*Xicgc1572*	aga(72)	F-ACGAGACGGTGGCATTAAG R-GGCAGGCACATAACAATACAAG	255
38	*Xicgc1699*	gag(21)	F-CATTAGCGGAGGATCTCTGTAG R-GTCGATGAAGACGGGAGTAG	438
39	*Xicgc1730*	tta(69)	F-TCGCACCGAACTCTTCATTC R-GCAGAGGTGCTCCCTTTATG	431
40	*Xicgc1738*	act(21)	F-GAGGAGTAAACGGAGGGAGTAG R-TGAGGCGAAGCTTGCTTATG	430
41	*Xicgc1755*	tta(26)	F-CCACCTGTCATTCACTGTTTAG R-CTCTGGCCCAATGACAAAC	211
42	*Xicgc1931*	atc(20)	F-TGAGGTTCTCAGAGGGTCTATC R-TCATCTTGTGGTATGTGGATGG	317
43	*Xicgc1988*	aag(23)	F-TGTGTCCATTCACAGGACATAG R-GTGGGACGTGGATGTAGTATTAG	196
44	*Xicgc2070*	ctc(30)	F-TTCTTCATAAGCGACGACACC R-ACATCATCGCTACAGGATCAAC	362
45	*Xicgc2071*	tca(34)	F-TCGATCTCCTCCTGACTTAGAG R-TTGGAAAGACTGCCCACTAC	429

Supplement 1. (Contd.)

No.	Marker	Motif	Primers	SSR length, bp
46	*Xicgc2150*	agg(23)	F-AGCAACGTCAAGACCAAGG R-GCGGGCACCTCTTCTTC	431
47	*Xicg14c002*	aga(42)caa(25)	F-AGCAGCAGAACACCCTAATG R-TCCAATCTTCCAGAGCCAATC	188
48	*Xicg14c004*	atg(28)	F-AATCAATGGGCATCAGGTAGG R-CTGGAAGTATCTGCGTCATCAG	309
49	*Xicg14c006*	gga(32)	F-TGAAGAAGATCACGACGAAGAAG R-GAACTGCTCCCAGAACTCATC	320
50	*Xicg14c009*	ttg(20)	F-CACACAAGGTGTTCGATGAAATG R-GCACTTTGGAGACGAGATGAG	372
51	*Xicg14c010*	taa(26)	F-TTCAAGGTGACCACATGGATAG R-TGAGCTTACAACACTCTGTCTG	265
52	*Xicg14c032*	aag(26)	F-AGCAAGGTTATGATGCCATTTAC R-GGCACGGTGAGTTTGATTTG	435
53	*Xicg15c001*	gcg(20)	F-AGCAACACAGTGGCGTTATC R-TCGGATTCTTCTTCTCCCAATTC	427
54	*Xicg15c002*	ctc(24)	F-GTCTCCAAGCTCTGGTTCAC R-GCAGTACCCACCATACAACTAC	444
55	*Xicg15c003*	cat(62)	F-TCGGCACGGATCTTTGTATG R-TCTCTTACCCACTCTTCTACCC	441
56	*Xicg15c009*	gaa(23)	F-AACACCGTTATCCATCCATCC R-TCCTATAGGTTAGCTCCCAAATTAC	352
57	*Xicg15c020*	tgt(26)	F-CGCTGTTGCTTCTCAAGTTTC R-CTCTAGCTACCTCTCCTCTCAC	343
58	*Xicg16c002*	gag(20)	F-GGACACGGTATGACCACTTATC R-ATTGCTTCGGAGCCTTCTC	410
59	*Xicg16c004*	atg(25)	F-GACTTGAAATCAATGGGCATCAG R-CTGGAAGTATCTGCGTCATCAG	310
60	*Xicg16c004_2*	tta(36)tac(23)	F-ATGGCATGCCGCTAGAAG R-CACCCGAAGTTGTCATAAATTACC	243
61	*Xicg16c005*	ctt(64)	F-CGATAGGACAACGGAAATCTACTC R-ATCCAGGTAGGAGGTCAAGG	334
62	*Xicg16c005_2*	ctt(21)	F-CCACCCACCATAAGCAAATC R-CCATCTAGGGAGGATGTCAAG	440
63	*Xicg16c006*	gag(35)	F-TTCATTCCAGAGGTGGTTGAG R-CTAGATGATGGAGTCAGCTGTG	432
64	*Xicg16c008*	aat(32)	F-ATTGCAGACTCGAGAACCATAC R-CCCTGACCTCTCTTGCATTATC	329
65	*Xicg16c008_2*	taa(24)	F-AGACTTTATCGTGAAGGGAGAAAC R-AGAGAGAGACCACCCACTAAC	293
66	*Xicg16c041*	ttc(22)	F-GAATGCTAGTCGTGGTTGATTTG R-CCAAACACAAGGGTTGCTATG	450
67	*Xicg16c041_2*	cta(27)tag(35)	F-TTCCGGGCATTCAACTGTAG R-GGGCTTTGAGAATCGGATAGG	377

Supplement 2. Structure of microsatellite loci and their localization on chromosome 5B

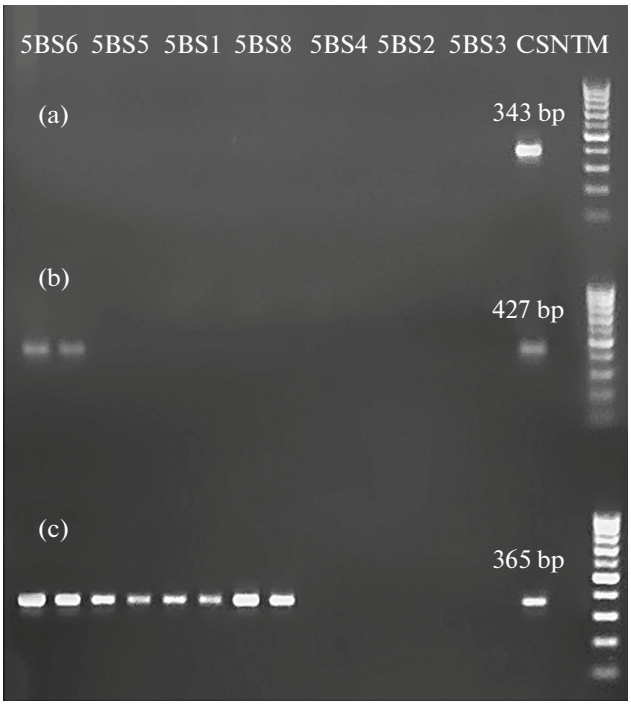
No.	Marker	Microsatellite motif	PCR fragment length	Localization on 5BS (deletion bin)
		AGGAGGAGGAGG../...CCTCCTCCTCC..		
1	*Xicgc2150*	agg(23)	275	–
2	*Xicg14c006*	gga(32)	325	–
3	*Xicg16c002*	gag(20)	450	–
4	*Xicgc1699*	gag(21)	150, 200, 350, **450**	5BS4
5	*Xicgc1382*	gag(27)	200, 250, 900	–
6	*Xicgc495*	gag(29)	**375**	5BS4
7	*Xicgc128*	gag(32)	250	–
8	*Xicg16c006*	gag(35)	450, 700	–
9	*Xicgc2070*	ctc(30)	300, 400, 900	–
10	*Xicg15c002*	ctc(24)	425	–
11	*Xicgc342*	tcc(20)	275, **400**	5BS5
12	*Xicgc709*	tcc(20)	325	–
13	*Xicgc967_2*	tcc(20)	375	–
14	*Xicgc642*	tcc(23)	425	–
15	*Xicgc967*	tcc(23)	400	–
		AAGAAGAAGAAG../..CTTCTTCTTCT..		
1	*Xicgc1988*	aag(23)	**180,** 200	C-5BS3
2	*Xicg14c032*	aag(26)	425	–
3	*Xicgc1572*	aga(72)	196, **200**	5BS5
4	*Xicg15c009*	gaa(23)	400	–
5	*Xicgc1550*	gaa(24)	375	–
6	*Xicg16c005_2*	ctt(21)	500, 800	–
7	*Xicgc299*	ctt(29)	**350**	5BS5
8	*Xicg16c005*	ctt(64)	300	–
9	*Xicgc307*	ttc(20)	**350**, 450	5BS5
10	*Xicg16c041*	ttc(22)	**500**	5BS6
11	*Xicgc686*	ttc(23)	**450**	5BS5
12	*Xicgc178_2*	ttc(29)	275, **500**	5BS6
13	*Xicgc1558*	tct(20)	400	–
14	*Xicg14c002*	aga(42)caa(25)	200	–
		CAACAACAACAA../...TTGTTGTTGTTG..		
1	*Xicg14c002*	aga(42)caa(25)	200	–
2	*Xicgc532*	caa(22)	300	–
3	*Xicgc1453*	aca(20)	275	–
4	*Xicg14c009*	ttg(20)	200, **375**	5BS6
5	*Xicgc482*	ttg(21)	300	–
6	*Xicgc131*	ttg(23)ttg(20)	**250**	5BS5
7	*Xicgc229*	ttg(27)	**450**	5BS5
8	*Xicgc374*	ttg(33)	400	–
9	*Xicg15c020*	tgt(26)	**375**	5BS6

Supplement 2. (Contd.)

No.	Marker	Microsatellite motif	PCR fragment length	Localization on 5BS (deletion bin)
10	*Xicgc284*	tgt(27)	**275**	5BS6
		ATCATCATCATC../...GATGATGATGAT..		
1	*Xicgc1931*	atc(20)	300	–
2	*Xicgc2071*	tca(34)	500	–
3	*Xicgc667*	tca(37)	485	–
4	*Xicgc707*	cat(23)	400, 450, 500	–
5	*Xicg15c003*	cat(62)	400	–
6	*Xicgc1531*	atg(20)	400	–
7	*Xicgc220*	atg(25)	400	–
8	*Xicgc936*	atg(25)	350	–
9	*Xicg16c004*	atg(25)	350	–
10	*Xicg14c004*	atg(28)	325	–
		AATAATAATAAT../...ATTATTATTATT..		
1	*Xicg16c008*	aat(32)	350	–
2	*Xicgc122*	ata(21)	**300**	5BS5
3	*Xicg16c008_2*	taa(24)	275	–
4	*Xicg14c010*	taa(26)	275	–
5	*Xicgc1755*	tta(26)	250	–
6	*Xicgc1730*	tta(69)	200, 350	–
7	*Xicgc312*	tat(26)	250	–
8	*Xicg16c004_2*	tta(36)tac(23)	**250**	5BS6
		TACTACTACTAC../...GTAGTAGTAGTA..		
1	*Xicg16c004_2*	tta(36)tac(23)	**250**	5BS6
2	*Xicgc1738*	act(21)	200	–
3	*Xicgc498*	cta(91)	**250**	5BS4
4	*Xicg16c041_2*	cta(27)tag(35)	275, 400	–
		CTGCTGCTGCTG../...CAGCAGCAGCAG..		
1	*Xicgc1427*	ctg(20)	275	–
2	*Xicgc740*	tgc(23)	180	–
3	*Xicgc456*	gct(24)	**270,** 280	5BS6
		GCCGCCGCCGCC../...GGCGGCGGCGG..		
1	*Xicgc195*	gcc(23)	250	–
2	*Xicg15c001*	gcg(20)	400	–
		CGACGACGACGA../...TCGTCGTCGTCG..		
1	*Xicgc619*	cga(20)	425	–
2	*Xicgc178*	gtc(23)	300	5BS6
		CCACCACCACCA../...TGGTGGTGGTG..		
1	*Xicgc69*	cca(30)	300	5BS4

[1] Comparative analysis of Chinese Spring and CSN5BT5A was performed by electrophoresis in agarose gel (Supplement 3); when polymorphism was undetectable, further analysis was performed by 10% PAGE. The PCR fragments located on chromosome 5B are boldfaced. Gray denotes microsatellite loci composed of different repeats.

Supplement 3



Electrophoresis of the SSR fragments of deletion lines—5BS6, 5BS5, 5BS1, 5BS8, 5BS4, 5BS2, and 5BS3 (theoretically expected lengths of fragments are shown): (a) localization of marker in bin 5BS6 by the example of marker Xicg15c020_2; (b) localization of marker in bin 5BS5 by the example of marker Xicgc229; and (c) localization of marker in bin 5BS4 by the example of marker Xicgc495. Each bit is represented by two lines; Chinese Spring (CS) is used as a control and line CSN5BT5A (NT), as a negative control; (M) length marker (100 bp).

REFERENCES

Adonina, I.G., Goncharov, N.P., Badaeva, E.D., Sergeeva, E.M., Petrash, N.V., and Salina, E.A., $(GAA)_n$ microsatellite as an indicator of the A genome reorganization during wheat evolution and domestication, *CompCytogen*, 2015, vol. 9, no. 4, pp. 533–547. doi 10.3897/CompCytogen.v9i4.5120

Akhunov, E., Nicolet, C., and Dvorak, J., Single nucleotide polymorphism genotyping in polyploid wheat with the illumina goldengate assay, *Theor. Appl. Genet.*, 2009, vol. 119, pp. 507–517. doi 10.1007/s00122-009-1059-5

Areshchenkova, T. and Ganal, M.W., Long tomato microsatellites are predominantly associated with centromeric regions, *Genome*, 1999, vol. 42, pp. 536–544.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., D'Amore, R., Allen, A.M., Mckenzie, N., Kramer, M., Kerhornou, A., Bolser, D., Kay, S., Waite, D., Trick, M., Bancroft, I., Gu, Y., et al., Analysis of the bread wheat genome using wholegenome shotgun sequencing, *Nature*, 2012, vol. 491, no. 7426, pp. 705–710. doi 10.1038/nature11650

Brown, S.M., Szewc-McFadden, A.K., and Kresovich, S., Development and application of simple sequence repeat (SSR) loci for plant genome analysis, in *Methods in Genome Analysis in Plants*, Boca Raton: CRC Press, 1996.

Chapman, J.A., Mascher, M., Buluç, A., Barry, K., Georganas, E., Session, A., Strnadova, V., Jenkins, J., Sehgal, S., Oliker, L., Schmutz, J., Yelick, K.A., Scholz, U., Waugh, R., Poland, et al., A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome, *Genome Biol.*, 2015, vol. 16, no. 1, p. 26. doi 10.1186/s13059-015-0582-8

Chevreux, B., Wetter, T., and Suhai, S., Genome sequence assembly using trace signals and additional sequence information, *Computer Science and Biology: Proc. of the German Conference on Bioinformatics*, 1999, pp. 45–56.

Cuadrado, A., Schwarzacher, T., and Jouve, N., Identification of different chromatin classes in wheat using in situ hybridization with simple sequence repeat oligonucleotides, *Theor. Appl. Genet.*, 2000, vol. 101, pp. 711–717. doi 10.1007/s001220051535

Cuadrado, A., Cardoso, M., and Jouve, N., Increasing the physical markers of wheat chromosomes using SSRs as FISH probes, *Genome*, 2008, vol. 51, no. 10, pp. 809–815. doi 10.1139/G08-065

Endo, T.R. and Gill, B.S., The deletion stocks of common wheat, *J. Hered.*, 1996, vol. 87, no. 4, pp. 295–307.

Feldman, M., The origin of cultivated wheat, in *The World Wheat Book*, Paris: Lavoisier Publishing, 2001.

Gusev, V.D., Miroshnichenko, L.A., and Chuzhanova, N.A., The detection of fractal-like structures in DNA sequences, in *Information Science and Computing. Int. Book Series, No. 8: Classification, Forecasting, Data Mining*, Sofia: ITHEA, 2009.

Gusev, V.D., Nemytikova, L.A., and Chuzhanova, N.A., On the complexity measures of genetic sequences, *Bioinformatics*, 1999, vol. 15, no. 12, pp. 994–999. doi 10.1093/bioinformatics/15.12.994

International Barley Genome Sequencing Consortium, A physical, genetic and functional sequence assembly of the barley genome, *Nature*, 2012, vol. 491, no. 7426, pp. 711–716. doi 10.1038/nature11543

International Wheat Genome Sequencing Consortium, A chromosomebased draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome, *Science*, 2014, vol. 345, no. 6194, p. 1251788. doi 10.1126/science.1251788

Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., Zhang, X., Jing, R., Zhang, C., Ma, Y., Gao, L., Gao, C., et al., *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation, *Nature*, 2013, vol. 496, no. 7443, pp. 91–95. doi 10.1038/nature12028

Langmead, B. and Salzberg, S., Fast gapped-read alignment with Bowtie 2, *Nat. Methods*, 2012, vol. 9, pp. 357–359.

Li, Y.-C., Korol, A.B., Beiles, A., and Nevo, E., Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review, *Mol. Ecol.*, 2002, vol. 11, pp. 2453–2465. doi 10.1046/j.1365-294X.2002.01643.x

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R., and 1000 Genome Project Data Processing Subgroup, The sequence alignment/map (SAM) format and SAMtools, *Bioinformatics*, 2009, vol. 25, pp. 2078–2079.

Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D., Dong, L., Tao, Y., Gao, C., Wu, H., Li, Y., Cui, Y., Guo, et al., Draft genome of the wheat A-genome progenitor *Triticum urartu*, *Nature*, 2013, vol. 496, no. 7443, pp. 87–90. doi 10.1038/nature11997

Logacheva, M.D., Schelkunov, M.I., and Penin, A.A., Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis*, *Genome Biol. Evol.*, 2011, vol. 3, pp. 1296–1303. doi 10.1093/gbe/evr102

Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J., and Pallen, M.J., Performance comparison of benchtop high-throughput sequencing platforms, *Nat. Biotechnol.*, 2012, vol. 30, no. 5, pp. 434–439. doi 10.1038/nbt.2198

Mason, A.S., SSR genotyping, *Methods Mol. Biol.*, 2015, vol. 1245, pp. 77–89. doi 10.1007/978-1-4939-1966-6_6

Pasquariello, M., Barabaschi, D., Himmelbach, A., Steuernagel, B., Ariyadasa, R., Stein, N., Gandolfi, F., Tenedini, E., Bernardis, I., Tagliafico, E., Pecchioni, N., and Francia, E., The barley Frost resistance-H2 locus, *Funct. Integr. Genomic*, 2014, vol. 14, no. 1, pp. 85–100. doi 10.1007/s10142-014-0360-9

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., Lagudah, E., Somers, D., Kilian, A., Alaux, M., Vautrin, S., et al., Physical map of the 1-Gigabase bread wheat chromosome 3B, *Science*, 2008, vol. 322, pp. 101–104. doi 10.1126/science.1161847

Plaschke, J., Ganal, M.W., and Röder, M.S., Detection of genetic diversity in closely related bread wheat using microsatellite markers, *Theor. Appl. Genet.*, 1995, vol. 91, pp. 1001–1007. doi 10.1007/BF00223912

Plaschke, J., Börner, A., Wendehake, K., Ganal, M.W., and Röder, M.S., The use of wheat aneuploids for the assignment of microsatellite loci, *Euphytica*, 1996, vol. 89, pp. 33–40. doi 10.1007/BF00015716

Qi, L., Echalier, B., Friebe, B., and Gill, B.S., Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ests, *Funct. Integr. Genomics*, 2003, vol. 3, pp. 39–55. doi 10.1007/s10142-002-0063-5

Qu, J. and Liu, J., A genome-wide analysis of simple sequence repeats in maize and the development of polymorphism markers from nextgeneration sequence data, *BMC Res. Notes*, 2013, vol. 6, p. 403. doi 10.1186/1756-0500-6-403

Sato, S., Hirakawa, H., Isobe, S., Fukai, E., Watanabe, A., Kato, M., Kawashima, K., Minami, C., Muraki, A., Nakazaki, N., Takahashi, C., Nakayama, S., Kishida, Y., Kohara, M., Yamada, et al., Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L., *DNA Res.*, 2011, vol. 18, no. 1, pp. 65–76. doi 10.1093/dnares/dsq030

Schmidt, T. and Heslop-Harrison, J.S., The physical and genomic organization of microsatellites in sugar beet, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, vol. 93, pp. 8761–8765.

Sears, E.R., Nullisomic-tetrasomic combinations in hexaploid wheat, in *Chromosome Manipulations and Plant Genetics*, London: Oliver and Boyd, 1966. doi 10.1007/978-1-4899-6561-5_4

Sergeeva, E.M., Afonnikov, D.A., Koltunova, M.K., Gusev, V.D., Miroshnichenko, L.A., Vrána, J., Kubaláková, M., Poncet, C., Sourdille, P., Feuillet, C., Doležel, J., and Salina, E.A., Common wheat chromosome 5B composition analysis using low-coverage 454 sequencing, *Plant Genome*, 2014, vol. 7, no. 2, pp. 1–16. doi 10.3835/plantgenome2013.10.0031

Sourdille, P., Singh, S., Cadalen, T., Brown-Guedira, G.L., Gay, G., Qi, L., Gill, B.S., Dufour, P., Murigneux, A., and Bernard, M., Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.), *Funct. Integr. Genomics*, 2004, vol. 4, pp. 12–25. doi 10.1007/s10142-004-0106-1

Staton, S.E., Bakken, B.H., Blackman, B.K., Chapman, M.A., Kane, N.C., Tang, S., Ungerer, M.C., Knapp, S.J., Rieseberg, L.H., and Burke, J.M., The sunflower (*Helianthus annuus* L.) genome reflects a recent history of biased accumulation of transposable elements, *Plant J.*, 2012, vol. 72, no. 1, pp. 142–153. doi 10.1111/j.1365-313X.2012.05072.x

Stein, N. and Steuernagel, B., Advances in sequencing the barley genome, in *Genomics of Plant Genetic Resources*, Springer Netherlands, 2014. doi 10.1007/978-94-007-7572-5_16

Tautz, D. and Renz, M., Simple sequences are ubiquitious repetitive component of eukaryotic genomes, *Nucleic Acids Res.*, 1984, vol. 12, pp. 4127–4138. doi 10.1093/nar/12.10.4127

Timonova, E.M., Dobrovol'skaya, O.B., Sergeeva, E.M., Bildanova, L.L., Sourdille, P., Feuillet, K., and Salina, E.A., A comparative genetic and cytogenetic mapping of wheat chromosome 5B using introgression lines, *Russ. J. Genet.*, 2013, vol. 49, no. 12, pp. 1200–1206.

Zhang, Z., Deng, Y., Tan, J., Hu, S., Yu, J., and Xue, Q., A genome-wide microsatellite polymorphism database for the indica and japonica rice, *DNA Res.*, 2007, vol. 14, pp. 37–45. doi 10.1093/dnares/dsm005

Translated by I. Grishina