SCIENCE CHINA
Life Sciences

High-throughput development of genome-wide locus-specific informative SSR markers in wheat

Li Lei, Sun Fayu, Wu Di, Zhen Fei, Bai Guihua, Gao Derong and Li Tao

Citation: SCIENCE CHINA Life Sciences ; doi: 10.1007/s11427-016-0252-x

View online: http://engine.scichina.com/doi/10.1007/s11427-016-0252-x

Published by the Science China Press

SCIENCE CHINA
Life Sciences

• LETTER TO THE EDITOR •

doi: 10.1007/s11427-016-0252-x

High-throughput development of genome-wide locus-specific informative SSR markers in wheat

Lei Li[1], Fayu Sun[1†], Di Wu[1†], Fei Zhen[1], Guihua Bai[2,3], Derong Gao[4] & Tao Li[1*]

[1]*Jiangsu Provincial Key Laboratory of Crop Genetics and Physiology/Co-Innovation Center for Modern Production Technology of Grain Crops; Key Laboratory of Plant Functional Genomics of Chinese Ministry of Education; Wheat Research Center; Yangzhou University, Yangzhou 225009, China;*

[2]*Department of Agronomy, Kansas State University, Manhattan KS 66506, USA;*

[3]*USDA-ARS Hard Winter Wheat Genetics Research Unit, Manhattan KS 66506, USA;*

[4]*Lixiahe Academy of Agricultural Sciences, Yangzhou 225000, China*

Received September 4, 2016; accepted October 31, 2016; published online December 16, 2016

Citation: Li, L., Sun, F., Wu, D., Zhen, F., Bai, G., Gao, D., and Li, T. High-throughput development of genome-wide locus-specific informative SSR markers in wheat. Sci China Life Sci. doi: 10.1007/s11427-016-0252-x

Dear Editor,

Although simple sequence repeat (SSR) markers are not new, they are still useful and often used markers in molecular mapping and marker-assisted breeding, particularly in developing countries. However, locus-specific SSR markers could be more useful and informative in wheat breeding and genetic studies. In the present study, 221,911 locus-specific SSR markers were designed. Verification of polymorphisms showed that the proportion of polymorphic markers increases with an increase in SSR size. Evaluation of the polymorphic information content (PIC) of the SSR markers in a panel of 12 wheat accessions showed that PIC also increases with an increase in SSR size. Alternative locus-specific SSR markers for known QTLs and the distributions of SSR markers and single-nucleotide polymorphisms (SNPs) from the Axiom 820K SNP array were also analyzed in this study.

A total of 1,037,987 SSRs were identified from a 10.2 Gb reference sequence of the hexaploid wheat cultivar Chinese Spring (CSS). SSRs were found to be considerably abundant with an average density of SSR per 9.8 kb (Table S1A in Supporting Information). Among these SSRs, 341 motif types were found on the CSS (Table S1B and Figure S1 in Supporting Information). The top 10% of the motif types account for 99% of the SSRs (Figure S2 in Supporting Information). The average sequence length of all SSRs is 16.2 bp.

A total of 735,419 primer pairs were designed using Primer 3 software (Rozen and Skaletsky, 2000). The e-PCR package (Schuler, 1997) was used to analyze the specificity of each primer pair. As expected, a proportion of these primers (236,155) had unique binding sites. A linkage map (Poland et al., 2012) was used to estimate the locations of the SSR markers. Considering that the CSS covers approximately 61% of the CS draft sequence and that the draft sequences of the wheat progenitor species *Triticum urartu* and *Aegilops tauschii* cover 94% (Ling et al., 2013) and 97% (Jia et al., 2013) of the A and D genomes, respectively, these draft sequences of A and D genomes were also included in the primer specificity analysis (Table S1C in Supporting Information). Finally, a total of 221,911 locus-specific SSR markers were identified (Table S2 in Supporting Information).

The e-PCR package was used to analyze polymorphisms of the 236,155 SSR markers between the draft sequences of cultivars CS and W7984. In the W7984 draft

†Contributed equally to this work
*Corresponding author (email: taoli@yzu.edu.cn)

sequence (http://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0582-8), 160,472 primer pairs (68%) generated locus-specific products. Among these locus-specific primers, 77,096 generated length polymorphisms (LP) between CS and W7984. Among the polymorphic markers, most have a mononucleotide motif (57.1%), but percentages of markers with LP>2 bp (17.6%) and LP>7 bp (7%) were low. There were high percentages of polymorphic markers with compound SSRs—48.3% of LP>2 and 30.3% of LP>7 (Figure 1A).

Further analysis showed that the markers with SSR size <20 bp are the most abundant (128,602; 80.1%) in the entire wheat genome, but that the percentage of polymorphic markers is low (44.3%), with only 15.9% of LP>2 and 6.4% of LP>7 being polymorphic. In contrast, for the SSR markers with SSR size >99 bp, the percentage of polymorphic markers is high (96.6%), with 95.9% of LP>2 and 88.3% of LP>7 being polymorphic (Figure 1B). Thus, the percentage of polymorphic markers increases linearly with an increase in size of the SSR markers. The proportion of markers with LP>2 or 7 increases rapidly with motif size below 49 bp, but there is a declining increase at motif sizes greater than 49 bp.

The Axiom 820K SNP array is one of the most widely used arrays in wheat genotyping because of its high-throughput performance. We compared the distributions of the locus-specific SSR markers and the Axiom 820K SNPs on the CSS. The 221,911 locus-specific SSR markers identified in the present study are from 175,776 contigs with 1.26 markers/contig, whereas the 794,176 SNPs of the 820K SNP array can be anchored on the CSS and are distributed only on 68,211 contigs. Although the number of 820K SNPs is 3.58 times greater than the number of SSR markers, the contig coverage of the SSR markers is 2.58-fold higher than that of the 820K SNPs (Table S3A in Supporting Information).

Eighty-seven PCR-based markers associated with known QTLs for processing quality, agronomic traits, and disease resistance were collected from a public database (http://wheat.pw.usda.gov/GG3/). Many of these markers have been recommended for the marker-assisted selection (MAS) of over 20 traits. Of these, 54 markers were anchored to the contigs of CS. An additional 74 novel SSR markers were identified to share contigs with the known set of 54 markers (Table S3B in Supporting Information). These novel SSR markers could be useful alternative markers in cases where the published markers are monomorphic between the wheat accessions of interest.

Initially, 102 SSR markers that are polymorphic between CS and W7984, generated from e-PCR with LP>3 bp, were randomly selected for marker verification. Of these, 97 primer pairs amplified distinct fragments. Eighty-six primers amplified polymorphic fragments between CS and W7984 (Table S4A in Supporting Information).

The validation experiment was extended to 1,024 markers that were randomly selected from 236,155 SSRs in a set of 12 hexaploid wheat varieties. Among these, 963 primers (94.0%) generated distinct fragments. Among the 963 primers, 676 amplified polymorphic bands across the 12 varieties with an average PIC of 0.415 (Table S4B in Supporting Information). When the SSR size was between 9 and 19 bp, the average PIC was 0.287, and the PIC increased to 0.560 when the SSR size was between 59 and 70 bp (Figure 1C). These results show that the PIC increases with an increase in the repeat size of SSR markers. A total of 3,234 markers with an SSR size greater than 59 bp are listed in Table S5 in Supporting Information.

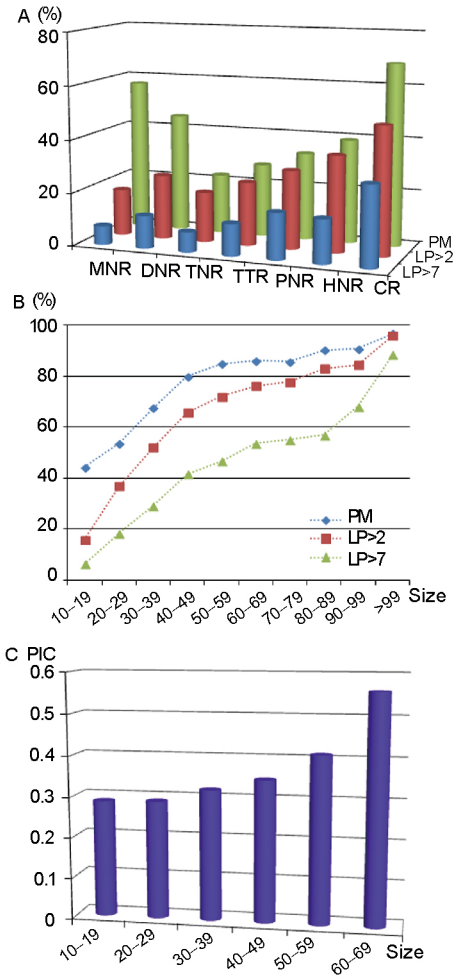

Figure 1 Polymorphism analysis of SSR markers between Chinese Spring and W7984. A, The percentage of polymorphic markers in different SSR types. B, The percentage of polymorphic markers in different SSR size. MNR, DNR, TNR, TTR, PNR, HNR and CR indicate mono-, di-, tri-, tetra-, penta-, hexa-nucleotide SSR and compound SSR, respectively. PM indicates polymorphic markers. LP>2 and LP>7 indicated that the length polymorphism is more than 2 and 7 bp, respectively. C, The relationship between SSR size and average PIC value.

The current study identified 221,911 genome-wide SSR markers that have unique binding sites on the CSS and the progenitor species of the A and D genomes of wheat. These locus-specific SSR markers are evenly distributed on 175,776

CS contigs. The identification of 74 additional markers could be useful in MAS. A selected set of these markers was validated by PCR in a panel of 12 wheat varieties. These locus-specific, evenly distributed, and highly polymorphic SSR markers will further enrich the SSR databases, and will be extremely useful for QTL mapping, diversity analysis, and gene cloning.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *We thank Dr. Aili Li at Chinese Academy of Agricultural Sciences for providing the synthetic wheat cultivar W7984. This work was supported by the National Natural Science Foundation of China (31270704, 31600997), the Yangzhou Key Research and Development Program (YZ2016035), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).*

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., Spannagl, M., Mayer, K.F.X., Li, D., Pan, S., Zheng, F., Hu, Q., Xia, X., Li, J., Liang, Q., Chen, J., Wicker, T., Gou, C., Kuang, H., He, G., Luo, Y., Keller, B., Xia, Q., Lu, P., Wang, J., Zou, H., Zhang, R., Xu, J., Gao, J., Middleton, C., Quan, Z., Liu, G., Wang, J., Wang, J., Yang, H., Liu, X., He, Z., Mao, L., and Wang, J. (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature 496, 91–95.

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, X., Zheng, S., Wang, B., Yu, K., Liang, Q., Yang, W., Lou, X., Chen, J., Feng, M., Jian, J., Zhang, X., Luo, G., Jiang, Y., Liu, J., Wang, Z., Sha, Y., Zhang, B., Wu, H., Tang, D., Shen, Q., Xue, P., Zou, S., Wang, X., Liu, X., Wang, F., Yang, Y., An, X., Dong, Z., Zhang, K., Zhang, X., Luo, M.C., Dvorak, J., Tong, Y., Wang, J., Yang, H., Li, Z., Wang, D., Zhang, A., and Wang, J. (2013). Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496, 87–90.

Poland, J.A., Brown, P.J., Sorrells, M.E., and Jannink, J.L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7, e32253.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the www for general users and for biologist programmers. Methods Mol Biol 132, 365−386.

Schuler, G.D. (1997). Sequence mapping by electronic PCR. Genome Res 7, 541−550.

SUPPORTING INFORMATION

Figure S1 Number and percentage of different repeat types of SSRs in hexaploid bread wheat. TTR, PNR, and HNR indicate tetra-, penta-, and hexanucleotide SSRs, respectively.

Figure S2 Distributions of SSR sizes in bread wheat genomes.

Table S1 Distribution of SSRs in wheat genomes

Table S2 SSR primers that had locus-specific binding sites in the Chinese Spring reference genome sequences and specificity analysis based on the e-PCR analysis of *Triticum urartu* (A genome) and *Aegilops tauschii* (D genome) and the hexaploid wheat W7984

Table S3 The distribution of simple sequence repeat (SSR) markers and 820K single-nucleotide polymorphisms (SNPs) (A). Additional markers for known quantitative trait loci (QTLs) (B)

Table S4 Validation of SSR markers

Table S5 SSR markers with repeats larger than 59 bp

The supporting information is available online at life.scichina.com and www.springerlink.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors