

## Development and Application of a Whole-Genome Simple Sequence Repeat Panel for High-Throughput Genotyping in Soybean

TAKASHI Sayama<sup>1,2,†</sup>, TAE-YOUNG Hwang<sup>2,†</sup>, KUNIIHIKO Komatsu<sup>2</sup>, YOSHITAKE Takada<sup>3</sup>, MASAKAZU Takahashi<sup>4</sup>, SHIN Kato<sup>5</sup>, HIROKO Sasama<sup>1,2</sup>, AYAKO Higashi<sup>2</sup>, YUMI Nakamoto<sup>2</sup>, HIDEYUKI Funatsuki<sup>2,‡</sup>, and MASAO Ishimoto<sup>1,2,\*</sup>

*National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan<sup>1</sup>; National Agricultural Research Center for Hokkaido Region, 1 Hitsujigaoka, Toyohira, Sapporo, Hokkaido 062-8555, Japan<sup>2</sup>; National Agricultural Research Center for Western Region, 1-3-1 Senyuu, Zentsuuj, Kagawa 765-8508, Japan<sup>3</sup>; National Agricultural Research Center for Kyushu Okinawa Region, 2421 Suya, Koshi, Kumamoto 861-1192, Japan<sup>4</sup> and National Agricultural Research Center for Tohoku Region, 297 Uenodai, Kariwano, Daisen, Akita 019-2112, Japan<sup>5</sup>*

\*To whom correspondence should be addressed. Tel./Fax. +81 29-838-7452. Email: ishimoto@affrc.go.jp

Edited by Kazuo Shinozaki

(Received 19 December 2010; accepted 16 February 2011)

### Abstract

**Among commonly applied molecular markers, simple sequence repeats (SSRs, or microsatellites) possess advantages such as a high level of polymorphism and codominant pattern of inheritance at individual loci. To facilitate systematic and rapid genetic mapping in soybean, we designed a genotyping panel comprised 304 SSR markers selected for allelic diversity and chromosomal location so as to provide wide coverage. Most primer pairs for the markers in the panel were redesigned to yield amplicons of 80–600 bp in multiplex polymerase chain reaction (PCR) and fluorescence-based sequencer analysis, and they were labelled with one of four different fluorescent dyes. Multiplex PCR with sets of six to eight primer pairs per reaction generated allelic data for 283 of the 304 SSR loci in three different mapping populations, with the loci mapping to the same positions as previously determined. Four SSRs on each chromosome were analysed for allelic diversity in 87 diverse soybean germplasms with four-plex PCR. These 80 loci showed an average allele number and polymorphic information content value of 14.8 and 0.78, respectively. The high level of polymorphism, ease of analysis, and high accuracy of the SSR genotyping panel should render it widely applicable to soybean genetics and breeding.**

**Key words:** SSR marker; fluorescent primer; multiplex PCR; polymorphic information content; high-throughput genotyping

### 1. Introduction

A comprehensive genetic linkage map is fundamental to modern plant genetics and breeding because it

allows the identification and utilization of agronomic trait loci, such as qualitative and quantitative trait loci, as well as the evaluation of genetic diversity and genomic structure of genetic resources. It also serves as a scaffold for construction of a physical map. In the case of soybean [*Glycine max* (L.) Merrill], a grain legume of global importance, many useful agronomic trait loci associated with growth, product quality, tolerance to biotic and abiotic stresses, and

† These authors contributed equally to this article.

‡ Present address: National Agricultural Research Center for Western Region, 6-12-1 Nishifukatsu, Fukuyama, Hiroshima 721-8514, Japan.

other characteristics have been identified in genetic resources and deposited in public databases such as Soybase (<http://soybase.org>). The first soybean linkage map was constructed on the basis of phenotypic traits,<sup>1</sup> but it did not contain sufficient information for application to the above-mentioned purposes. Since the 1990s, various types of molecular markers, including restriction fragment length polymorphism, random amplification of polymorphic DNA, amplified fragment length polymorphism, simple sequence repeat (SSR, or microsatellite), and single nucleotide polymorphism (SNP) markers, have been developed and applied to soybean, and enriched marker information has enabled genetic analyses of qualitative and quantitative traits.<sup>2-4</sup>

Among such molecular markers, SSRs in particular have contributed to the construction of a genome-wide linkage map for soybean with a converged linkage group (LG) number equal to the chromosome number. As previously described,<sup>5</sup> these microsatellite markers have several advantages, including (i) a codominant manner of inheritance at each locus, (ii) a high level of polymorphism in the form of multiple alleles, (iii) a non-biased distribution in the genome, and (iv) ease of detection of polymorphism by the polymerase chain reaction (PCR) and subsequent electrophoresis.<sup>6-9</sup> It is thus possible to detect allelic differences at highly polymorphic loci among genetic resources as well as among many segregating populations derived from the hybridization of any given genotypes.<sup>10-12</sup> Furthermore, given that most PCR primer pairs for SSR markers were designed to yield a single amplification product for each allele in spite of the complex chromosomal structure of soybean,<sup>13-15</sup> each SSR marker localizes a definite site in the genome, unlike other molecular markers. These features also allow the simultaneous detection of multiple SSR loci with the use of multiplex PCR analysis.<sup>11</sup> In addition, high-throughput and repetitive genotyping can be performed by semiautomated methods with a combination of fluorescently labelled SSR markers.<sup>11,12,16</sup> This strategy has been successfully adapted and improved for analysis of genetic diversity and high-throughput mapping in various species, including human,<sup>17,18</sup> mouse,<sup>19</sup> rat,<sup>20</sup> rice,<sup>21,22</sup> and sunflower.<sup>23</sup> However, in spite of the public availability of many SSR markers for soybean,<sup>24</sup> no comprehensive SSR genotyping panel has yet been developed for whole-genome coverage. A high-throughput genotyping system for soybean that is based on an SNP array capable of high multiplexing and which discriminates up to 384 or 1536 mapped SNPs in one reaction has been described.<sup>25,26</sup> Although the high multiplexing capacity and continued improvement of the SNP array may make this a standard technique in the foreseeable future, the

quantity and quality of SNP loci in soybean are still not sufficient for application of this system to many genotypes. Moreover, the present panel system for SNPs is not sufficiently flexible for modification of marker selection and is not a cost-effective solution.

In the latest extensive molecular linkage map, nearly 2000 SSR markers were mapped to the 20 consensus soybean LGs without any large gaps with the exception of one region in LG C1 (chromosome 4).<sup>5</sup> In addition, allelic variation at a large number of SSR loci was examined in 23 different soybean genotypes. The availability of this large amount of information prompted us to develop a whole-genome SSR panel for high-throughput genotyping in soybean. We selected SSR markers on the basis of their polymorphism and chromosomal location, and then redesigned them for adaptation to multiplex PCR. The resulting SSR panel system was applied to construct linkage maps of three segregating populations with different genetic backgrounds, and the positions of the redesigned SSR markers were confirmed.

## 2. Materials and methods

### 2.1. Design of a whole-genome SSR panel

On the basis of their positions and polymorphism information content (PIC) values previously described,<sup>5,27</sup> we initially selected 322 SSR markers from the mapped SSR loci. The selected markers in principle had a PIC value of >0.5 and were separated from each other by a distance of <20 cM. They consisted of 263 SSR markers developed by USDA-ARS (United States Department of Agriculture, Agriculture Research Service),<sup>27,28</sup> 15 SSR markers developed by Chiba University,<sup>29</sup> and 44 expressed sequence tag (EST)-derived SSR markers developed by Kazusa DNA Research Institute.<sup>30</sup> For adaptation to multiplex PCR, most of the primer pairs for the SSR markers developed by USDA were redesigned with the use of the Primer 3 program on the basis of the following parameters: a melting temperature of 50°–60°C (optimum, 55°C); a product size of 80–600 bp; and all other items set to default.<sup>31</sup> One of the four different fluorophores (6-FAM, VIC, NED, or PET) was attached to each forward primer, and a 5' tail was added to the reverse primer in order to promote full adenylation (Applied Biosystems, Foster City, CA, USA). Each SSR marker was arranged to eliminate overlap with other SSR markers labelled with the same fluorophore in each multiplex set. Details of the markers and primers are provided in Supplementary Table S1.

### 2.2. PCR conditions and SSR fragment detection

PCR was performed in a reaction mixture of 5.5  $\mu$ l containing 50 nM of each primer, 5 ng of total

genomic DNA, and 2.5  $\mu$ l of 2  $\times$  Qiagen Multiplex PCR Master Mix (Qiagen, Hilden, Germany) and with the use of a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems). Total plant DNA was extracted from 10 mg of seed powder with the use of an automated purification system (BioSprint 96 DNA Plant Kit, Qiagen). The amplification protocol comprised an initial denaturation for 15 min at 95°C; 35 cycles of denaturation for 30 s at 95°C, annealing for 90 s at 50°C, and extension for 90 s at 60°C; and a final extension for 30 min at 60°C. A portion (1  $\mu$ l) of the amplification products was then mixed with 0.3  $\mu$ l of GeneScan 600 LIZ Size Standard (Applied Biosystems) and 10  $\mu$ l of Hi-Di Formamide (Applied Biosystems), and the resulting mixture was heated for 3 min at 95°C and then cooled rapidly to 4°C in the thermal cycler. Sizing of SSR fragments was performed with a 3730 DNA Analyzer equipped with POP-7 Polymer and a 36-cm Capillary Array (Applied Biosystems), and allele calling and binning were performed with GeneMapper software v4.0 (Applied Biosystems).

### 2.3. SSR mapping

Three mapping populations were used to position fluorescently labelled SSR markers. (i) The J1 population, which consists of 94 F<sub>2</sub> individuals originating from a cross between the Japanese wild accession JP-36121 and the Japanese landrace Ibarakimame 7. (ii) The OA population, which is derived from a cross between the Japanese cultivar Osuzu and the USA cultivar Athow and comprises 225 F<sub>6</sub> recombinant inbred lines (RILs). (iii) The IF population, which consists of 143 F<sub>7</sub> RILs derived from a cross between the Japanese landrace Ippon\_sangoh and a leading Japanese cultivar, Fukuyutaka. On the basis of allelic information for the SSR markers of the whole-genome SSR panel, individual linkage maps of the three populations were constructed with the use of MAPMAKER/EXP 3.0b.<sup>32</sup> Recombination values were converted to genetic distance (cM) with the use of the Kosambi mapping function.<sup>33</sup> Each map of the 20 LGs was visualized graphically with MapChart.<sup>34</sup> The nomenclature of LGs was as described previously,<sup>27</sup> based on the common markers mapped.

### 2.4. Diversity analysis

Four SSR markers labelled with different fluorophores were chosen for even distribution on each chromosome on the basis of previous results.<sup>12</sup> Allelic diversity for the total of 80 selected markers was assessed among 87 soybean accessions for which diversity has been previously characterized on the basis of analysis by 3% agarose gel electrophoresis.<sup>12</sup> The 87 diverse genotypes consisted of 37

modern cultivars and 15 landraces collected from various regions in Japan, 27 cultivars derived from countries other than Japan (mostly those in East Asia and the USA), and 8 wild accessions. The PIC value for each SSR marker locus *i* was calculated according to the formula previously described:<sup>2</sup>

$$PIC(i) = 1 - \sum P_{ij}^2$$

where  $P_{ij}$  is the frequency of the *j*th allele of the *i*th SSR locus and the summation extends over *n* alleles. Allele number and PIC value determined with the present system were compared with those determined in the previous study.<sup>12</sup>

## 3. Results and discussion

### 3.1. Detail of the designed whole-genome SSR panel

A whole-genome SSR panel was designed predominantly on the basis of SSR markers developed by USDA-ARS, since these markers have been well documented.<sup>27</sup> The markers were selected according to their map position and polymorphic information.<sup>5</sup> However, given that the USDA markers alone did not cover the entire soybean genome with an even distribution, we selected additional SSR markers that were in the public domain. The selected markers were designed according to different strategies and could not be amplified under the same conditions.<sup>28–30</sup> In particular, primer pairs for the USDA markers were characterized by a relatively low annealing temperature (~47°C) and the presence of a GCG-clamp near the 5' end of each primer. Most (255 out of 263) of the primer pairs for the USDA SSR markers were therefore redesigned to enhance their suitability for multiplex PCR analysis and to yield PCR products within the detection range of the DNA analyser. All of the 322 fluorescent SSR markers initially constituting the whole-genome SSR panel were individually assessed with the six parents of the three mapping populations as well as with two standard cultivars, Enrei and Williams 82, that have been used for genomic studies in Japan and the USA, respectively. Eighteen SSR primer pairs yielded no amplicons by singleplex PCR with the eight test genotypes (Supplementary Table S1), likely because of the primer modification for multiplex PCR and fluorescence labelling. The remaining 304 primer pairs, which generated unambiguous and stable PCR amplicons, were therefore selected and assembled into 41 multiplex sets of six to eight markers each (Supplementary Table S1). The SSR primer pairs in each multiplex set were assigned one of the four different fluorescent labels (6-FAM, VIC, NED, or PET) in such a manner that no two markers with



the same fluorescent dye showed overlap in allele size range. The 304 SSR markers were estimated to cover 2306 cM, or 95%, of the previous integrated linkage map,<sup>5</sup> and were separated by an average interval of 7.6 cM (Table 1). Each chromosome (Chr) was covered with 9 (Chr 16) to 21 (Chr 7) SSR makers. There remained a large gap of 35.3 cM between Sat\_235 and GMES1325 on Chr 4, but no public SSR markers have been described for this region.<sup>5</sup> A gap of 33.9 cM existed between GMES2057 and GMES1693 on Chr 9, because the selected marker Satt539 gave no amplicons. An additional eight regions on different chromosomes showed a gap of >20 cM between proximal markers (Supplementary Table S1). No confident SSR markers were selected in these regions. The longest and shortest average distances between markers were on Chr 4 (9.6 cM) and Chr 7 (5.9 cM), respectively, indicating that the SSR panel was capable of genotyping with a relatively uniform marker density over the entire soybean genome, with only a few exceptions.

**Table 1.** Assignment of SSR markers designed for the whole-genome SSR panel system

Chromosome	Linkage group	No. of SSR marker loci	Coverage <sup>a</sup> (cM)	Marker density <sup>b</sup> (cM)
1	D1a	14	113	8.1
2	D1b	16	138	8.6
3	N	15	106	7.1
4	C1	12	116	9.6
5	A1	14	122	8.7
6	C2	19	140	7.4
7	M	21	124	5.9
8	A2	20	155	7.8
9	K	13	111	8.5
10	O	15	142	9.5
11	B1	18	128	7.1
12	H	15	106	7.0
13	F	15	122	8.1
14	B2	12	97	8.1
15	E	15	103	6.8
16	J	9	66	7.3
17	D2	15	96	6.4
18	G	14	120	8.5
19	L	17	102	6.0
20	I	15	100	6.7
Total		304	2306	
Average		15.2	115.3	7.6

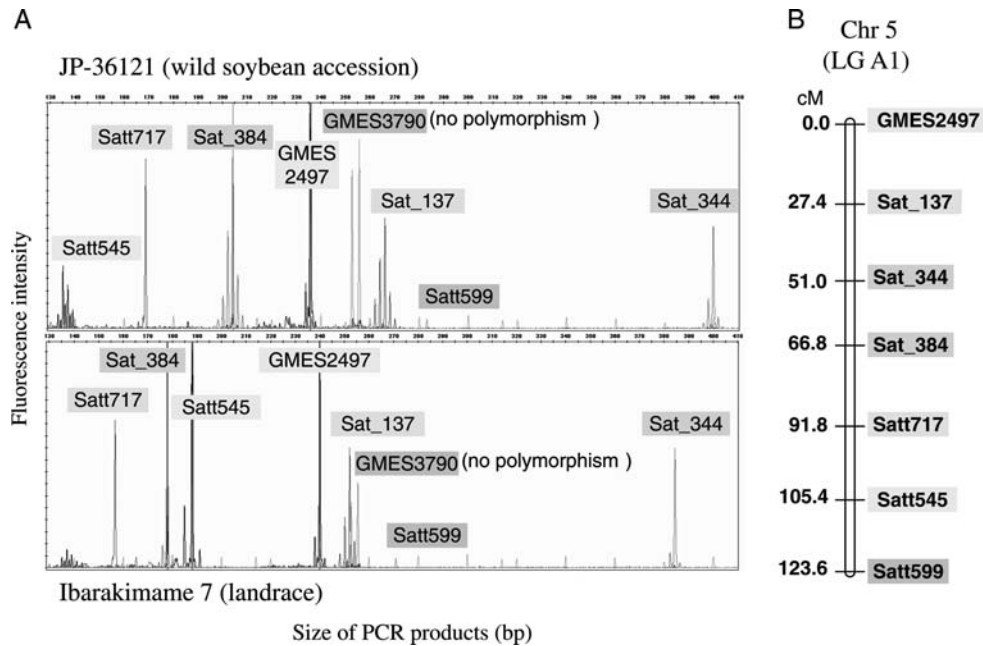
<sup>a</sup>The coverage (cM) of the SSR marker loci in each LG is based on information for the previous integrated linkage map.<sup>5</sup>

<sup>b</sup>Coverage/marker number.

### 3.2. Genetic mapping of SSR loci in three mapping populations

Linkage analyses were performed with the three mapping populations (JI, OA, and IF) to assess the efficacy of the whole-genome SSR panel. At the same time, we verified whether the fluorescent SSR markers, including those that had been redesigned, were located at the same positions as those originally determined. Among the three populations, the JI population, comprising 94 F<sub>2</sub> individuals, was assumed to be diverse at many SSR loci because it was developed from a cross between a wild soybean (JP-36121) and a Japanese landrace (Ibarakimame 7).<sup>35</sup> For each multiplex set, the allele products for each SSR locus were readily distinguished from each other as well as from those of the other loci represented (Fig. 1A). One multiplex set built a rough map of Chr 5 (Fig. 1B), whereas the two multiplex sets for Chr 5 made up a linkage frame spanning 142.1 cM with 11 polymorphic markers (Fig. 2). Given that the other three markers selected for Chr 5 in the SSR panel did not exhibit polymorphism between the parents of the JI population, the polymorphic ratio for Chr 5 was calculated at 0.786 (Table 2). This process was applied sequentially to all 20 chromosomes, with the 41 multiplex PCRs achieving genotyping of the 304 SSR markers. All the primer pairs yielded amplicons from at least one of the parents of the JI population, although 55 primer pairs showed no polymorphism between the parents. A linkage map covering 2864 cM was thus constructed from 249 markers with an average interval of 11.5 cM (Fig. 2, Table 2). All of the mapped markers were arranged in the same order and with similar genetic intervals compared with previous studies.<sup>5,27–30</sup> Our whole-genome SSR panel thus appears to achieve highly accurate and efficient genotyping with sufficient marker coverage.

Wild soybeans including JP-36121 are thought to be more genetically diverse than cultivated soybeans, with the result that the genome panel was able effectively to build a linkage map of the JI population. However, molecular linkage maps are frequently constructed from segregating populations with relatively similar genetic backgrounds in order to analyse important agricultural traits such as maturity and yield performance.<sup>36</sup> We therefore next applied the SSR genome panel to the analysis of two mapping populations, OA and IF, derived from crosses between soybean cultivars. The polymorphic ratio (0.549) of the IF population, which is derived from Japanese germplasm, was found to be much lower than that (0.819) of the JI population, but it was similar to that (0.615) of the OA population (Table 2). A total of 187 SSR markers showed



**Figure 1.** Representative genotyping and mapping of eight SSR marker loci on soybean chromosome 5 (LG A1). (A) Multiplex PCR analysis of the eight SSR marker loci with primer sets labelled with four different fluorophores: 6-FAM (Sat\_137 and Satt717), VIC (GMES3790 and Satt599), NED (GMES2497 and Satt545), and PET (Sat\_344 and Sat\_384). The separation and detection of PCR products corresponding to alleles of the eight marker loci with a fluorescence-based sequencer and GeneMapper 4.0 software are shown. (B) Genetic linkage map constructed with MAPMAKER/EXP 3.0b for the polymorphic SSR marker loci and the 94  $F_2$  individuals of the JI population derived from a cross between JP-36121 (wild soybean accession) and Ibarakimame 7 (landrace).

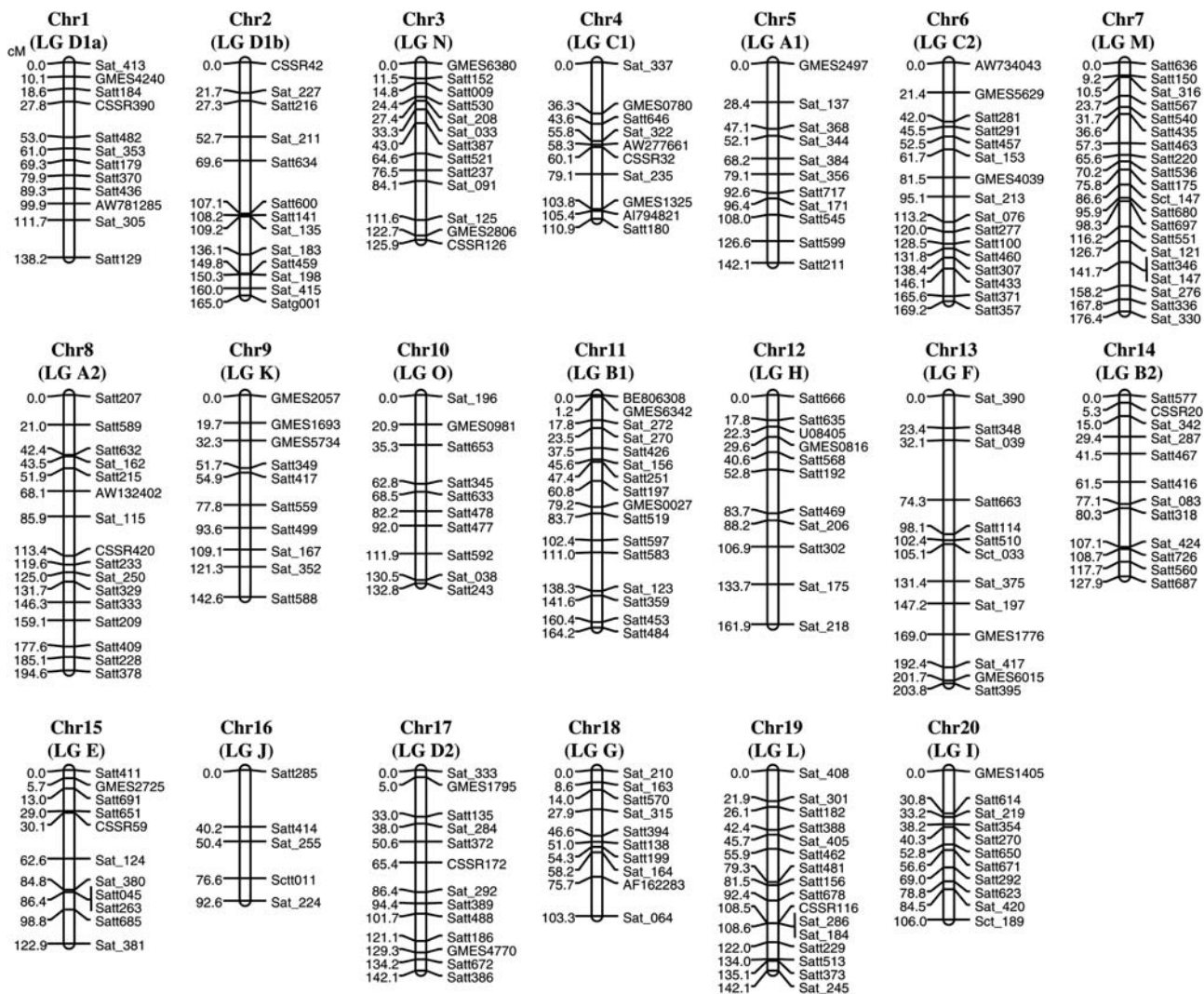
polymorphism between the Japanese cultivar Osuzu and the US cultivar Athrow, which are the parents of the OA RILs, and provided a linkage map covering 2613 cM with an average marker interval of 14.0 cM. The Japanese landrace Ippon\_sangoh and the elite Japanese cultivar Fukuyutaka, which are the parents of the IF population of 143  $F_7$  RILs, exhibited allelic variation at 167 SSR loci, and the genotype data for these RILs allowed the construction of a linkage map covering 2470 cM with an average interval of 14.8 cM. The SSR genome panel system thus achieved the construction of linkage maps with sufficient marker coverage in the three mapping populations, and it was less labour-intensive and more rapid than the precedent methods by singleplex PCRs and gel electrophoreses. The molecular linkage maps obtained provide a basis for genetic analysis of specific traits such as deficiency of group A acetylated saponins.<sup>35</sup>

The polymorphic markers were evenly distributed among the 20 LGs, with 201 (66.1%) of the 304 SSR loci being mapped in at least two populations. We thus verified that the SSR genome panel system was well designed to minimize bias of marker location and to include highly polymorphic markers. A total of 21 markers were not polymorphic in the three mapping populations; however, their locations were not confirmed. Although these markers may show polymorphic alleles in other genetic backgrounds,

they should possibly be excluded from the genome panel. Many of the markers exhibiting no polymorphism were selected from the EST-derived SSR markers, which were previously found to be less diverse, with fewer alleles, than the USDA SSR markers.<sup>5,30</sup> Large marker intervals of >30 cM were observed between Satt634 and Satt600 on Chr 2 and between Sat\_039 and Satt663 on Chr 13 in all three mapping populations. The addition of new highly polymorphic markers to these two regions and other relatively sparse regions would enhance the capability of the SSR genome panel system. We are currently in the process of arranging additional SSR markers for these genomic regions based on the recently completed whole-genome sequence of soybean and the abundant SSR motifs identified in the genome.<sup>24,37</sup>

### 3.3. Diversity analysis of 87 soybean genotypes with the whole-genome SSR panel

We next examined allelic diversity of selected SSR markers in 87 soybean genotypes, comprising 8 wild accessions and 79 cultivars, including Japanese elite cultivars, that we had analysed previously.<sup>12</sup> From the 304 SSR loci in the whole-genome SSR panel, we selected four markers for each chromosome, to give a total of 80 SSR loci, on the basis of their map position and polymorphism (Supplementary Table S2). The previous analysis by PCR and 3% agarose gel electrophoresis<sup>12</sup> revealed that the allele number for the



**Figure 2.** Molecular linkage map constructed with the whole-genome SSR panel system for the JI population. A total of 304 SSR marker loci were analysed with the 41 multiplex PCR sets of six to eight primer pairs, resulting in the generation of a linkage map consisting of 249 SSR marker loci covering 2864 cM of the soybean genome.

80 SSR loci ranged from 2 to 7, with an average of 4.0 (Fig. 3A). The most frequent allele number was 3, which accounted for 31.3% of all the marker loci. The PIC value, a measure of allelic diversity at a locus, ranged from 0.12 (Satt568) to 0.80 (Sat\_115), and the average PIC value was 0.55 (Fig. 3B). In contrast, in our panel system with the selected 80 SSR marker set, the number of alleles for each locus ranged from 3 (Satt307) to 32 (Satt263), with an average of 14.8, and the PIC value ranged from 0.45 (Satt568) to 0.96 (Sat\_115), with a mean of 0.78 for all the marker loci. This level of polymorphism was thus substantially higher than that obtained previously by analysis with agarose or polyacrylamide gel electrophoresis,<sup>5,12</sup> likely reflecting the fact that a capillary sequencer is able to determine the size of repetitive motifs with higher resolution and greater accuracy. The mean PIC value and the average allele number obtained in

the present study are similar to those determined in previous studies with high-resolution analysis systems, although different diverse materials were used in each study.<sup>10,11,38</sup> However, the PIC value and allele number for each SSR locus in the present study did not correlate well with our previous results,<sup>12</sup> indicating that allelic diversity at each SSR locus includes small differences that are difficult to determine by 3% agarose gel electrophoresis. The high level of diversity of the SSR loci will allow wide application of the whole-genome SSR panel to detailed genotype identification as well as to mapping phenotypic and quantitative trait loci in various segregating populations.

### 3.4. Conclusion

Thousands of microsatellite markers in soybean have been developed over the last two decades, with



**Table 2.** Summary of molecular linkage maps constructed for three mapping populations with the whole-genome SSR panel system

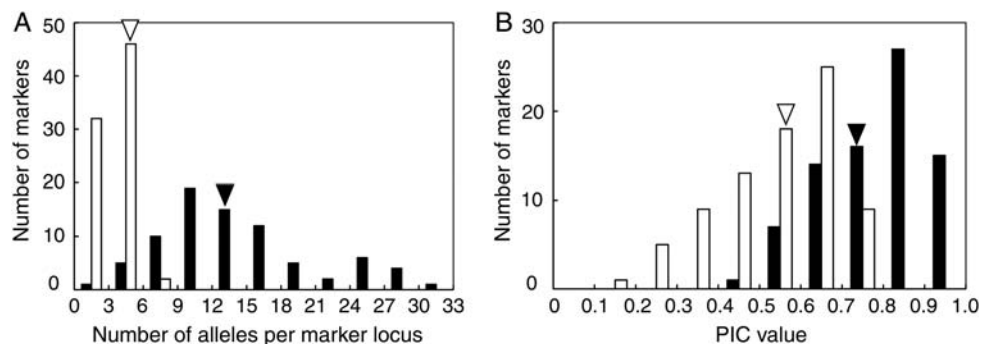
Chromosome	Linkage group	No. of SSR marker loci	Polymorphic markers in each mapping population						No. of non-polymorphic markers	
			JI (2864 cM) <sup>a</sup>		OA (2613 cM) <sup>b</sup>		IF (2470 cM) <sup>c</sup>			Total
			No.	Ratio <sup>d</sup>	No.	Ratio <sup>d</sup>	No.	Ratio <sup>d</sup>		
1	D1a	14	12	0.857	7	0.500	5	0.357	13	1
2	D1b	16	13	0.813	9	0.563	8	0.500	15	1
3	N	15	13	0.867	14	0.933	10	0.667	15	0
4	C1	12	10	0.833	9	0.750	9	0.750	12	0
5	A1	14	11	0.786	9	0.643	5	0.357	12	2
6	C2	19	16	0.842	9	0.474	7	0.368	18	1
7	M	21	20	0.952	11	0.524	13	0.619	21	0
8	A2	20	16	0.800	10	0.500	12	0.600	17	3
9	K	13	10	0.769	9	0.692	10	0.769	13	0
10	O	15	10	0.667	11	0.733	8	0.533	12	3
11	B1	18	16	0.889	8	0.444	8	0.444	16	2
12	H	15	11	0.733	13	0.867	8	0.533	13	2
13	F	15	13	0.867	9	0.600	6	0.400	15	0
14	B2	12	12	1.000	7	0.583	8	0.667	12	0
15	E	15	11	0.733	5	0.333	6	0.400	12	3
16	J	9	5	0.556	6	0.667	7	0.778	8	1
17	D2	15	13	0.867	13	0.867	11	0.733	15	0
18	G	14	10	0.714	11	0.786	9	0.643	13	1
19	L	17	16	0.941	8	0.471	7	0.412	16	1
20	I	15	11	0.733	9	0.600	10	0.667	15	0
Total		304	249	0.819	187	0.615	167	0.549	283	21
Average		15.2	12.5		9.4		8.4		14.2	1.1

<sup>a</sup>A population of 94 F<sub>2</sub> plants derived from a cross between JP-36121 and Ibarakimame 7 was used for construction of a linkage map spanning 2864 cM.

<sup>b</sup>A population of 225 RILs derived from a cross between Osuzu and Athrow was used for construction of a linkage map spanning 2613 cM.

<sup>c</sup>A population of 143 RILs derived from a cross between Ippon\_sangoh and Fukuyutaka was used for construction of a linkage map spanning 2470 cM.

<sup>d</sup>Polymorphic ratio = number of polymorphic marker loci/number of analysed marker loci.



**Figure 3.** Distribution of allele numbers (A) and PIC values (B) for 80 SSR marker loci in 87 soybean cultivars or wild accessions as determined by multiplex PCR and a fluorescence-based sequencer in the present study (solid bars) or by singleplex PCR and 3% agarose gel electrophoresis in a previous study<sup>12</sup> (open bars). Arrowheads indicate the corresponding averages for allele number and PIC value determined by the two methods.

>2000 SSR loci having been integrated into a common linkage map.<sup>5,27–30</sup> The abundant and well-documented SSR loci allow the relatively unlimited selection of SSR markers on the basis of their diverse length polymorphism and map location. We selected 304 SSR markers and assembled them into 41 multiplex PCR sets to give a whole-genome SSR panel. This genome panel system has been applied to linkage analysis of >10 segregating populations in addition to the three populations analysed in the present study, and it has successfully built the framework of each linkage map (data not shown). The results indicate the wide applicability of the genome panel system to various combinations of soybean genotypes. The combination of multiplex PCR and the high-resolution detection system is also applicable to other genotyping studies such as marker-assisted selection and the fine mapping of qualitative and quantitative traits. The flexibility of our genome panel system should improve the efficiency of molecular marker genotyping by reducing the labour intensity and time required for this process, although this system requires a fluorescent sequencer and hundreds of custom-labelled primers with fluorescent dye.

The recent breakthrough of high multiplexing has moved the SNP array into the mainstream of genotyping technology.<sup>39</sup> Many thousands of SNPs have been identified in soybean and have been used to design the GoldenGate assay, which discriminates up to 1536 SNPs in one reaction.<sup>4,25,26</sup> On the basis of its high multiplexing capacity, automation, and continued improvement, the SNP array is becoming one of most reliable methods for whole-genome genotyping. However, the quantity and quality of SNP loci in the entire soybean genome are currently not sufficient for application of this system to many genotypes. In addition, the GoldenGate assay is not flexible for modification of SNP marker selection and is expensive for each reaction. Although the multiplex PCR system coupled with high-resolution detection is capable of analysing at most 20 SSR loci at a time, it is possible to combine in a flexible manner the mapped SSR markers and SSR motifs that are scattered throughout the soybean genome.<sup>24</sup>

**Supplementary data:** Supplementary Data are available at [www.dnaresearch.oxfordjournals.org](http://www.dnaresearch.oxfordjournals.org).

## Funding

This research was supported by the Ministry of Agriculture, Forestry, and Fisheries of Japan [Genomics for Agricultural Innovation (DD-3260)].

## References

- Palmer, R.G. and Kilen, T.C. 1987, Qualitative genetics and cytogenetics. In: Wilcox, J.R. (ed.), *Soybeans: Improvement, Production, and Uses*, 2nd edition. Agronomy Monograph, Madison, Wisconsin, pp. 135–209.
- Keim, P., Beavis, W., Schupp, J. and Freestone, R. 1992, Evaluation of soybean RFLP marker diversity in adapted germplasm, *Theor. Appl. Genet.*, **85**, 205–12.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990, DNA polymorphisms amplified by arbitrary primers are useful as genetic markers, *Nucleic Acids Res.*, **18**, 6531–5.
- Choi, I.Y., Hyten, D.L., Matukumalli, L.K., et al. 2007, A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis, *Genetics*, **176**, 685–96.
- Hwang, T.Y., Sayama, T., Takahashi, M., et al. 2009, High-density integrated linkage map based on SSR markers in soybean, *DNA Res.*, **16**, 213–25.
- Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. 1992, Length polymorphisms of simple sequence repeat DNA in soybean, *Genetics*, **132**, 1131–9.
- Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A. and Cregan, P.B. 1995, Integration of simple sequence repeat DNA markers into a soybean linkage map, *Crop Sci.*, **35**, 1439–45.
- Morgante, M. and Olivieri, A.M. 2008, PCR-amplified microsatellites as markers in plant genetics, *Plant J.*, **3**, 175–82.
- Maughan, P.J., Saghi Maroof, M.A. and Buss, G.R. 1995, Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean, *Genome*, **38**, 715–23.
- Chen, Y., Wang, D., Arelli, P., Ebrahimi, M. and Nelson, R.L. 2006, Molecular marker diversity of SCN-resistant sources in soybean, *Genome*, **49**, 938–49.
- Diwan, N. and Cregan, P.B. 1997, Automated sizing of fluorescent labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean, *Theor. Appl. Genet.*, **95**, 723–33.
- Hwang, T.Y., Nakamoto, Y., Kono, I., et al. 2008, Genetic diversity of cultivated and wild soybean lines including Japanese elite cultivars as revealed by length polymorphism of SSR markers, *Breed. Sci.*, **58**, 315–23.
- Cregan, P.B., Bhagwat, A.A., Akkaya, M.S. and Rongwen, J. 1994, Microsatellite fingerprinting and mapping of soybean, *Methods Mol. Cell Biol.*, **5**, 49–61.
- Shoemaker, R.C., Polzin, K., Labate, J., et al. 1996, Genome duplication in soybean (*Glycine subgenus soja*), *Genetics*, **144**, 329–38.
- Gill, N., Findley, S., Walling, J.G., et al. 2009, Molecular and chromosomal evidence for allopolyploidy in soybean, *Plant Physiol.*, **151**, 1167–74.
- Coburn, J.R., Temnykh, S.V., Paul, E.M. and McCouch, S.R. 2002, Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L.), *Crop Sci.*, **42**, 2092–9.
- Doxiadis, G.G.M., De Groot, N., Claas, F.H.J., Doxiadis, I.I.N., Van Rood, J.J. and Bontrop, R.E. 2007,



- A highly divergent microsatellite facilitating fast and accurate DRB haplotyping in humans and rhesus macaques, *Proc. Natl Acad. Sci. USA*, **104**, 8907–12.
18. Lindqvist, A.K., Magnusson, P.K., Balciuniene, J., et al. 1996, Chromosome-specific panels of tri- and tetranucleotide microsatellite markers for multiplex fluorescent detection and automated genotyping: evaluation of their utility in pathology and forensics, *Genome Res.*, **6**, 1170–6.
  19. Schalkwyk, L.C., Jung, M., Daser, A., et al. 1999, Panel of microsatellite markers for whole-genome scans and radiation hybrid mapping and a mouse family tree, *Genome Res.*, **9**, 878–87.
  20. Bryda, E.C. and Riley, L.K. 2008, Multiplex microsatellite marker panels for genetic monitoring of common rat strains, *J. Am. Assoc. Lab. Anim. Sci.*, **47**, 37–41.
  21. Pessoa-Filho, M., Belo, A., Alcochete, A.A.N., Rangel, P.H.N. and Ferreira, M.E. 2007, A set of multiplex panels of microsatellite markers for rapid molecular characterization of rice accessions, *BMC Plant Biol.*, **7**, 23.
  22. de Oliveira Borba, T.C., Brondani, R.P.V., Rangel, P.H.N. and Brondani, C. 2009, Microsatellite marker-mediated analysis of the EMBRAPA Rice Core Collection genetic diversity, *Genetica*, **137**, 293–304.
  23. Tang, S., Kishore, V.K. and Knapp, S.J. 2003, PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower, *Theor. Appl. Genet.*, **107**, 6–19.
  24. Song, Q., Jia, G., Zhu, Y., et al. 2010, Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR\_1.0) in soybean, *Crop Sci.*, **50**, 1950–60.
  25. Hyten, D.L., Song, Q., Choi, I., et al. 2008, High-throughput genotyping with the GoldenGate assay in the complex genome of soybean, *Theor. Appl. Genet.*, **116**, 945–52.
  26. Hyten, D.L., Choi, I.-Y., Song, Q., et al. 2010, A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping, *Crop Sci.*, **50**, 960–8.
  27. Song, Q.J., Marek, L.F., Shoemaker, R.C., et al. 2004, A new integrated genetic linkage map of the soybean, *Theor. Appl. Genet.*, **109**, 122–8.
  28. Cregan, P.B., Jarvik, T., Bush, A.L., et al. 1999, An integrated genetic linkage map of the soybean genome, *Crop Sci.*, **39**, 1464–90.
  29. Xia, Z., Tsubokura, Y., Hoshi, M., et al. 2007, An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using a single F2 population, *DNA Res.*, **14**, 257–69.
  30. Hisano, H., Sato, S., Isobe, S., et al. 2007, Characterization of the soybean genome using EST-derived microsatellite markers, *DNA Res.*, **14**, 271–81.
  31. Rozen, S. and Skaletsky, H. 2000, Primer3 on the WWW for general users and for biologist programmers, *Methods Mol. Biol.*, **132**, 365–86.
  32. Lincoln, S.E., Daly, M.J. and Lander, E.S. 1992, *Constructing Genetic Maps with MAPMAKER/EXP 3.0*, Whitehead Institute Technical Report, 3rd edition. Whitehead Institute, Cambridge, MA.
  33. Kosambi, D.D. 1944, The estimation of map distance from recombination values, *Ann. Eugen.*, **12**, 172–5.
  34. Voorrips, R.E. 2002, MapChart: software for the graphical presentation of linkage maps and QTLs, *J. Hered.*, **93**, 77–8.
  35. Takada, Y., Sayama, T., Kikuchi, A., et al. 2010, Genetic analysis of variation in sugar chain composition at the C-22 position of group A saponins in soybean, *Glycine max* (L.) Merrill, *Breed. Sci.*, **60**, 3–8.
  36. Sayama, T., Hwang, T.Y., Yamazaki, H., et al. 2010, Mapping and comparison of quantitative trait loci for soybean branching phenotype in two locations, *Breed. Sci.*, **60**, 380–9.
  37. Schmutz, J., Cannon, S.B., Schlueter, J., et al. 2010, Genome sequence of the palaeopolyploid soybean, *Nature*, **463**, 178–83.
  38. Wang, L., Guan, R., Zhangxiong, L., Chang, R. and Qiu, L. 2006, Genetic diversity of Chinese cultivated soybean revealed by SSR markers, *Crop Sci.*, **46**, 1032–8.
  39. Fan, J.B., Gunderson, K.L., Bibikova, M., et al. 2006, Illumina universal bead arrays, *Methods Enzymol.*, **410**, 57–73.