Breeding Science 70: 231–240 (2020)
doi: 10.1270/jsbbs.19116

Research Paper

DNA markers based on retrotransposon insertion polymorphisms can detect short DNA fragments for strawberry cultivar identification

Chiharu Hirata[1)], Takamitsu Waki[2)], Katsumi Shimomura[1)], Takuya Wada[1)], Seiya Tanaka[3)], Hidetoshi Ikegami[1)], Yousuke Uchimura[1)], Keita Hirashima[1)], Yoshiko Nakazawa[2)], Kaori Okada[2)], Kiyoshi Namai[2)], Makoto Tahara[4)] and Yuki Monden*[4)]

[1)] *Fukuoka Agriculture and Forestry Research Center*, 587 Yoshiki, Chikushino, Fukuoka 818-8549, Japan
[2)] *Tochigi Prefectural Agricultural Experiment Station*, 1080 Kawarayacho, Utsunomiya, Tochigi 320-0002, Japan
[3)] *Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University*, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
[4)] *Graduate School of Environmental and Life Science, Okayama University*, 1-1-1 Tsushima-Naka, Kita-ku, Okayama 700-8530, Japan

In this study, DNA markers were developed for discrimination of strawberry (*Fragaria* × *ananassa* L.) cultivars based on retrotransposon insertion polymorphisms. We performed a comprehensive genomic search to identify retrotransposon insertion sites and subsequently selected one retrotransposon family, designated CL3, which provided reliable discrimination among strawberry cultivars. Through analyses of 75 strawberry cultivars, we developed eight cultivar-specific markers based on CL3 retrotransposon insertion sites. Used in combination with 10 additional polymorphic markers, we differentiated 35 strawberry cultivars commonly cultivated in Japan. In addition, we demonstrated that the retrotransposon-based markers were effective for PCR detection of DNA extracted from processed food materials, whereas a SSR marker was ineffective. These results indicated that the retrotransposon-based markers are useful for cultivar discrimination for processed food products, such as jams, in which DNA may be fragmented or degraded.

Key Words: *Fragaria* × *ananassa*, high-throughput sequencing, PCR product, processed foods, retrotransposon insertion polymorphisms.

Introduction

Strawberry (*Fragaria* × *ananassa* L.), an octoploid plant, is a highly profitable fruit crop cultivated mainly in temperate and subtropical regions. With global demand increasing, the worldwide production of strawberries in 2016 was approximately 9,118,336 tonnes (Food and Agriculture Organization, http://faostat.fao.org/, 2017). The fruit is consumed raw and is also used as an ingredient in processed foods, such as purees and jams.

In recent years, strawberries have been intensively bred to develop cultivars of superior quality in a variety of countries, including Japan. Increasingly, the fruits of these cultivars are exported overseas, and the resulting unauthorized plant use and re-importation of falsely labeled fruits of popular cultivars have led to concerns over infringement of the intellectual property rights of seed companies.

Detection of DNA polymorphisms is an effective method for protection of plant breeders' rights and elimination of fraud associated with cultivar labeling. Diverse molecular marker types, including random amplified polymorphic DNA (RAPD), restriction fragment length polymorphisms (RFLP), amplified fragment length polymorphisms (AFLP), simple sequence repeats (SSR), cleaved amplified polymorphic sequences (CAPS), inter-SSR (ISSR), and single-nucleotide polymorphisms (SNP), have been used in a wide variety of crops (Yano 2010). Differentiation of cultivated strawberries using RAPD, RFLP, AFLP, SSR, CAPS, and ISSR markers has been reported previously (Arnau *et al.* 2003, Bonoli *et al.* 2005, Degani *et al.* 1998, 2001, Honjo *et al.* 2011, Isobe *et al.* 2013, Kunihisa *et al.* 2003, 2005). Especially to identify Japanese cultivars, Shimomura and co-workers developed AFLP markers able to differentiate 10 strawberry cultivars grown in Japan (Shimomura *et al.* 2005) and SSR markers that can differentiate 12 cultivars (Shimomura and Hirashima 2006). In addition, (Kunihisa *et al.* 2009a, 2009b, Kunihisa 2010) used CAPS markers to identify 125 strawberry cultivars, which currently represents a standard method to identify Japanese strawberry cultivars.

However, these methods are based on the combination of

Communicated by Sachiko Isobe
Received August 6, 2019. Accepted December 3, 2019.
First Published Online in J-STAGE on February 26, 2020.
*Corresponding author (e-mail: y_monden@okayama-u.ac.jp)

multiple DNA markers of comparatively large fragment sizes. Identification of samples of heat- and pressure-treated products in which the DNA is fragmented or degraded, such as jams, is difficult with such markers. Identification of processed samples derived from mixed cultivars is also problematic. To overcome these difficulties, promising techniques that utilize polymorphisms of retrotransposon insertion sites have been developed (Yano 2010).

Retrotransposons are genetic elements that increase their copy number via RNA intermediates. Retrotransposon DNA sequences are first transcribed into an RNA template, reverse transcribed into DNA, and then inserted into a new genomic site. Once the DNA is inserted, it is stably inherited (Kumar and Bennetzen 1999). Through the course of evolution, plant genomes have acquired multiple copies of the replicated sequences; however, the majority lose their transcriptional activity. Replicated retrotransposon sequences that are present in multiple copies within the genome are potential DNA markers (Flavell *et al.* 1998, Kumar and Hirochika 2001, Waugh *et al.* 1997).

One advantage of retrotransposon insertion polymorphisms is the potential they provide as small marker fragments. The use of small marker fragments is likely to be more effective for detection of DNA that has been degraded or fragmented during the manufacturing process of heat- and pressure-treated products, such as jams. These markers are dominant and therefore do not need to be separated by size in electrophoresis. In addition, the presence of cultivar-specific retrotransposon insertion sites raises the possibility of their utility as direct cultivar-specific markers (Tahara *et al.* 2009).

The use of retrotransposons for cultivar discrimination has been reported for rice (*Oryza sativa*) (Fukuchi *et al.* 1993), dried sweet potato (*Ipomoea batatas*), which is made by heating sweet potato (Ooe *et al.* 2004), and red bean paste, which is a heat-treated product of the adzuki bean (*Vigna angularis*) (Yamashita *et al.* 2008).

Full-length *FaRE*1, which is a transcriptionally activated long terminal repeat (LTR) retrotransposon sequence, was isolated from strawberry (He *et al.* 2010). Akitake *et al.* (2013) reported that *FaRE*1 insertion polymorphisms can be useful to differentiate strawberry cultivars. Recently, a novel LTR retrotransposon family, CL3, was discovered. This retrotransposon family exhibits high insertion polymorphism among strawberry cultivars (Monden *et al.* 2014a).

Based on these findings, in this study we performed a comprehensive search of retrotransposon insertion sites in the strawberry genome to develop a practical marker system for discrimination of strawberry cultivars. The selected retrotransposon family was analyzed in a diverse selection of strawberry cultivars. Subsequently, small (around 100 bp) DNA markers were developed for cultivar discrimination, which could also be applied to processed food products.

Materials and Methods

Selection of retrotransposon families for marker development

To select retrotransposon families for marker development, analysis of sequence-specific amplification polymorphism (S-SAP) was performed using 24 strawberry cultivars/lines (**Table 1**). This method detects polymorphisms in the interval between the retrotransposon insertion site and the nearest restriction enzyme site.

This analysis was implemented as described by Akitake *et al.* (2013). First, genomic DNA was extracted from leaves of each test cultivar/strain using the DNeasy Plant Mini Kit (Qiagen, Germany) and cleaved using the restriction enzymes *Mse*I and *Rsa*I. An adapter sequence was appended to the termini of the restriction-enzyme-cut fragments. The forked adapters were prepared by annealing two pairs of oligos (Forked_Type1 and *Mse*I Forked_Com or Forked_Type1 and *Rsa*I Forked_Com). The DNA sequences between the terminal sequences of five retrotransposon families (LTRs), and the *Mse*I and *Rsa*I restriction sites located near the retrotransposon insertion sites were amplified by PCR. PCR primers were designed based on the sequences of the five retrotransposon families. We performed the initial PCR with adapter-specific (LAM-AP2-HQ) and 5′ LTRs/PBS (Met_CL_11_1st, Met_CL_20_1st, Met_CL_76_1st, Met_CL_28_1st, and Met_CL_3_PBS) primer combinations using *Mse*I- or *Rsa*I-digested DNA fragments as the template. Then, nested PCR was performed with adapter-specific (LAM-AP3-HQ) and fluorescent-labeled LTR region-specific (Met_11_FAM, Met_20_ FAM, Met_76_ FAM, Met_28_ FAM, and CL3_5End_FAM) primer sets using the initial PCR product as the template. The initial PCR and the nested PCR included an initial denaturation at 94°C for 2 min, which was followed by 30 cycles at 94°C for 60 s, 75°C for 60 s, 58°C for 90 s, and 72°C for 30 s, with a final extension at 72°C for 5 min. The amplified fragments were quantified using an ABI PRISM® 3500 Genetic Analyzer (Applied Biosystems, USA). The sequences of all adapters and primers are listed in **Supplemental Table 1**.

Comprehensive analysis of retrotransposon insertion site sequences

To obtain sequence information for regions that included a large number of retrotransposon insertion sites, we constructed an Illumina MiSeq sequencing library as previously described (Monden *et al.* 2014c). First, genomic DNA was extracted from leaves of 75 strawberry cultivars/lines (**Table 1**). Genomic DNA (5 µg) was separated into fragments of ~6 kb by g-TUBE (Covaris, USA) centrifugation and purified using the QIAquick PCR Purification Kit (Qiagen). The purified products were treated with DNA polymerase I and T4 DNA polymerase to convert the heterogeneous ends via their physical fragmentation into blunt

Table 1. Strawberry cultivars or strains used in this study

No.	Cultivar Name[ab]	Parentage	Origin
1	*Fukuoka s6 (Amaou)**	Kurume 53 × 92-46	Japan
2	Kurume 53	Toyonoka × Terunoka	Japan
3	92-46	Kurume 49 × Sachinoka	Japan
4	*Aiberry**	Reiko × Houkou	Japan
5	*Akihime**	Kunowase × Nyoho	Japan
6	*Sagahonoka**	Osuzu × Toyonoka	Japan
7	*Sachinoka**	Toyonoka × Aiberry	Japan
8	Santiigo	(Aiberry × Hokowase)-s × Toyonoka-s	Japan
9	*Toyonoka**	Himiko × Harunoka	Japan
10	*Benihoppe**	Akihime × Sachinoka	Japan
11	*Redpearl**	Aiberry × Toyonoka	Japan
12	*Kumaken i 547 (Hinoshizuku)**	98-30 (Sachinoka × Tochinomine) × 98-20-3 (Kurume 54 × Tochinomine)	Japan
13	Houkouwase*	Yakumo × Tahoe	Japan
14	Harunoka	Kurume 103 × Donner	Japan
15	Fukuba*	Progeny of General Chanzy	Japan
16	Reiko	Harunoka-s × Fukuba-s	Japan
17	*Kaorino**	0028401 × 0023001	Japan
18	*Kotoka**	7-3-1 × Benihoppe	Japan
19	Kurume 55	Satinoka × 8701-18 (Toyonoka × Terunoka)	Japan
20	*Ookimi**	Satsumaotome × Strawberry Parental Line Nou - 1	Japan
21	*Himiko*	Kurume 34 × Houkouwase	Japan
22	*Miyazakinatsuharuka**	Breeding line in Miyazaki Prefectual Agricultural Experimental Station × Sweet Charmy	Japan
23	Elsanta*	Gorella × Holiday	Netherlands
24	Florida Belle*	Released in Florida 1975.	USA
25	06A-184*	Fukuoka S6 × Santiigo	Japan
26	Fukuoka S9	(Akihime × Fukuoka S6) × Santiigo	Japan
27	Fukuoka S8	Fukuoka S6 × Santiigo	Japan
28	*Megumi*	(Toyonoka × Aiberry) × Toyonoka	Japan
29	*Yumenoka**	Kurume55 × Kei531	Japan
30	*Satsumaotome**	8821-11 × Sachinoka	Japan
31	*Marihime**	Akihime × Sachinoka	Japan
32	*Amaotome**	Tochiotome × Sagahonoka	Japan
33	*Pechika*	Oishishikinari 2 × Summer Berry	Japan
34	*Natsuakari*	Summer berry × Kitanokagayaki	Japan
35	*Deco-Rouge*	Pajaro × Morioka 26	Japan
36	*Summer Candy*	(Summer berry × M26) × (Tochiotome)	Japan
37	*Summer Tiara*	Selva × Benihoppe	Japan
38	*Natsuotome*	Tochigi 24 × 0025-1	Japan
39	*Tochihitomi*	(Progeny of Celine) × Sachinoka	Japan
40	*Kitanokagayaki*	Belle Rouge × Pajaro	Japan
41	*Mouikko*	MN3 ((Nyohou × Shizutakara) × Nyohou) × Sachinoka	Japan
42	*Otomegokoro*	Sakyu S 2 × Kitanokagayaki	Japan
43	*Echigohime*	(Belle Rouge × Nyoho) × Toyonoka	Japan
44	*Yayoihime*	(Tonehoppe × Tochiotome) × Tonehoppe	Japan
45	*Nyoho*	Kei 210 × Reiko	Japan
46	*Tochihime*	Tochinomine × Kurume 49	Japan
47	*Tochigi i 27 (Skyberry)*	00-24-1 × Tochigi20	Japan
48	*Tochiotome**	Kurume-49 × Tochinomine	Japan
49	Tochigi 26	Tochiotome × 00-11-1	Japan
50	Tochigi 29	03-33-2 × Tochigi 20	Japan
51	Tochigi 33	04-1-14 × No.252	Japan
52	Tochigi 34	Natsuotome × Tochigi 26	Japan
53	Aptos	CAL 65.63 × Tufts	USA
54	Serienu	Oishishikinari × Natsuho	Japan
55	Donner*	CAL 145.52 × CAL 222	USA
56	Shin-nyoho	Mutator strain of Nyoho	Japan
57	Chandler	Douglas × Cal 72.361-105	USA
58	America	unknown	USA
59	Summer Berry	Kaho × Reiko	Japan
60	Morioka 26	Aptos × Morioka 16	Japan
61	Tochigi 24	97-77-1 × Tonehoppe	Japan
62	00-25-1	98-28-2 × No.227	Japan
63	Belle ruju	(Huxley × Albriton) × Morioka 19	Japan
64	Tonehoppe	Kei 56 × Nyoho	Japan
65	Kei 210	Donner × (Harunoka × Donner)	Japan
66	Kurume 49	Toyonoka × Nyohou	Japan
67	Tochinomine	(Florida69-266 × Reiko) × Nyoho	Japan
68	00-24-1	98-2-2 × Tonehoppe	Japan
69	Tochigi 20	97-3-4 × 94-2-8	Japan
70	Asuka wave	52-1-32 × 50-2-2	Japan
71	Tochigi 31	Tochigi 23 × 05-108-88	Japan
72	00-24-1	91-21-7 × No.227	Japan
73	03-33-2	01-18-2 × 96-14-5	Japan
74	04-01-14	Tochinomine × 00-1-1	Japan
75	No. 252	Breeding line in Tochigi Agricultural Experimental Station	Japan

[a] Asterisks indicate cultivars used for S-SAP analysis.
[b] Thirty-five cultivars commonly cultivated in Japan are italicized.

ends. The blunt ends were adenylated at the 3′ end using a Klenow fragment that lacked 5′ to 3′ exonuclease activity. A forked adapter with a T overhang at the 5′ double-stranded end was ligated using T4 DNA ligase. The retrotransposon insertion sites were specifically amplified using nested PCR. The adapter sequences (Forked_Type1 and Forked_Com), the primer containing internal primer-binding sites (PBS) to amplify LTR retrotransposon insertion sites (Met_CL_3_PBS), the adapter-specific primer (LAM-AP2-HQ), and the tailed PCR primers, including the cultivar tag sequences (P5_retrotransposon primer and P7_adapter sequence primer), are listed in **Supplemental Table 2**. Following the production of a library for each cultivar/strain, the molecular size (400–600 bp) and yield (500 ng) of each library were measured using an Agilent 2010 bioanalyzer and a Qubit fluorometer (Invitrogen, USA). To confirm the presence of genomic sequences adjacent to the 5′ LTR terminus, the PCR products used during library construction were randomly cloned in *Escherichia coli*, and PCR amplification of the cloning sites was performed on *E. coli* colonies using primers (M13 forward, M13 reverse, and 5′ LTR terminal sequence specific). The presence of genomic sequences adjacent to the 5′ LTR terminus was confirmed by PCR. Next, equal concentrations of all samples were pooled to prepare an Illumina MiSeq sequencing library (150 bp × 2), which was then submitted to next-generation sequencing using the Illumina MiSeq. MiSeq sequencing was performed in accordance with the manufacturer's manual. The obtained reads were analyzed using Maser, which is the pipeline execution system of the Cell Innovation Program at the National Institute of Genetics (http://cell-innovation.nig.ac.jp/index_en.html). Data were analyzed as previously described (Monden *et al.* 2014c). First, the paired-end reads were filtered for invalid dual barcodes; if a read had an erroneous or undetermined barcode at either end, the entire read was discarded. Then, the reads were trimmed to 50 bp from the retrotransposon junction, and adapter removal with cutadapt (https://code.google.com/p/cutadapt/) and quality filtering based on the quality value for all base calls ≥20 were conducted. Reads shorter than this length (50 bp) were filtered out. If there were ≥10 identical sequences, reads were collapsed into a single sequence in FASTA format, and if there were <10 identical sequences, reads were excluded from further analyses. Clustering analyses to identify insertion sites were performed with the BLAT self-alignment program (Kent 2002) using the following parameter settings: -tileSize = 8, -minMatch = 1, -minScore = 10, -repMatch = –1, and -oneOff = 2. To investigate the similarity of sequences in the FASTA file, an all-to-all comparison analysis was conducted, and clusters were built based on sequence similarities with the results of the pairwise alignments. In each cluster, a multiple sequence alignment was generated to reveal sequence similarity using the ClustalW program (Larkin *et al.* 2007). The sequence with the highest read number in each cluster was selected as the representative sequence. These analyses produced a number of clusters and non-clustered sequences with ≥10 reads, each indicating an individual insertion site where a retrotransposon copy was inserted in at least one cultivar. Detailed information on the data analysis was shown in **Supplemental Fig. 2**. We summarized the total number of reads and the number of reads for each cluster in all cultivars. However, some erroneous assignments and clustering would have resulted in a very small number of reads in some clusters. Thus, to remove the incorrectly clustered reads caused by sequencing error in the barcode, we set a critical value for determining the presence of insertions; i.e., if the number of reads at a specific insertion site comprised <0.01% of the entire reads in each cultivar, we assumed that the retrotransposon was absent from that site (**Supplemental Fig. 3**). These processes yielded genotyping information for the presence (1) or absence (0) of retrotransposon insertions in all cultivars. Cultivar-specific insertion sites and polymorphic insertion sites among cultivars were screened based on the genotyping information.

Development and verification of cultivar-specific markers

Discriminatory markers were developed for the test cultivars/strains for which cultivar-specific retrotransposon insertion sites were obtained, and also for 35 important commercial cultivars (**Table 1**). To achieve this, insertion site sequences that exhibited polymorphisms among cultivars and numerous read counts (>100) among 75 cultivars were selected based on the genotyping data. Primers were designed based on the selected sequences (**Supplemental Table 3**), and the sequences were confirmed by PCR amplification. The primers were designed such that one annealed across the junction between the genomic sequence and the retrotransposon 5′ LTR region, had a GC content of 45%–55%, and produced amplification products of about 100 bp. PCR amplification was carried out in 20 µL reaction mixtures containing 1–10 ng genomic DNA, 10 µL AmpliTaq Gold® 360 Master Mix (Applied Biosystems), and 1 µM of each primer. PCR products were amplified in a thermal cycler (PC-818A, ASTEC Co. Ltd., Japan), and the reaction conditions were 94°C for 4 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s, and a final extension at 72°C for 1 min before cooling to 4°C. Amplification products were electrophoresed in 2.0% agarose gel to verify the presence of fragments. For amplified markers, the presence or absence of amplification for each cultivar was scored using a MultiNA microchip electrophoresis apparatus (Shimadzu Corporation, Japan), and the smallest set of discriminatory markers was selected using Minimal Marker (Fujii *et al.* 2013).

Marker applicability to processed products

Four commercially available strawberry jams were selected as a heat- and pressure-treated product and used as templates for marker detection. Two jams (jams A and B) labeled 'Made of "Fukuoka S6 (Amaou)" fruits' were used

to investigate the applicability of cultivar-discrimination markers. The other two jams (jams C and D), made from fruits of unknown strawberry cultivars, were used to compare detectability because of differences in marker size. About 100 mg of each jam (with achenes excluded) were measured in 2.0 ml safe-lock tubes and dried using a freeze dryer (FDU-1200, Tokyo Rikakikai, Japan). Total genomic DNA was extracted from the samples using the DNeasy Plant Mini Kit (Qiagen). Eleven retrotransposon-based markers (CL3_Cl19, CL3_P524, CL3_Cl124, CL3_Cl322, CL3_Cl261, CL3_P320, CL3_Cl115, CL3_Cl242, CL3_Cl258, CL3_Cl76, and CL3_Cl214), and one SSR marker described in the Kazusa Marker Database (http://marker.kazusa.or.jp) as a positive control marker (FVES3384), were selected and tested for validation using jam A, jam B, and "Fukuoka S6 (Amaou)" leaves. Three retrotransposon-based control markers designed for different product sizes (Cl214_1, Cl214_2, Cl214_3) and a SSR marker (FVES3384) were tested using jams C and D. The experimental procedure, which included PCR, electrophoresis, and detection operation, followed the aforementioned protocol used for leaf samples. The primer sequences are listed in **Supplemental Table 3**. DNA qualities from the four jams and "Fukuoka S6 (Amaou)" leaves are listed in **Supplemental Table 6**.

Results

Selection of retrotransposon families for marker development

To select the most suitable retrotransposon family for this study, we investigated the insertion polymorphisms among strawberry cultivars for five LTR retrotransposon families, as reported by Monden *et al.* (2014a). Two different restriction enzymes (*Mse*I: TTAA and *Rsa*I: GATC), which recognize four bases, were used for S-SAP analysis. Five LTR retrotransposon families exhibited high insertion polymorphism among cultivars, as well as cultivar-specific fragments (**Tables 2**, **3**). Of the candidate LTR retrotransposon families, CL76 showed the highest number of cultivars with cultivar-specific fragments (24), and CL3 showed the highest number of detected fragments (2106).

As the LTR-terminal sequence of CL3 was determined easily and was identical to the sequence of the iMet_CL3 retrotransposon family, which showed high insertion polymorphism among strawberry cultivars in a previous study (Monden *et al.* 2014a), the CL3 retrotransposon family was selected for comprehensive analysis of insertion sites using a high-throughput sequencing platform.

Comprehensive analysis of retrotransposon insertion site sequences

A total of 19,532,224 reads (min: 99,808, average: 260,429.7, max: 637,946 reads per cultivar) were obtained by MiSeq sequencing analysis. After the preprocessing steps, a total of 13,502 collapsed sequences were generated for clustering analyses (**Supplemental Table 4**). The insertion polymorphism of the CL3 retrotransposon is summarized in **Table 4**. A total of 807 independent insertion sites were identified in 75 cultivars; 478 insertion sites with the average of 146.6 insertion sites per cultivar were determined to show high reliability because 100 or more sequence reads for 75 cultivars per site were obtained. Of 478 insertion sites, 474 sites (99.1%) were polymorphic and 63 sites (13.4%) were specific to one of the tested cultivars. Twenty cultivars possessed specific sites. Insertion

Table 2. Summary of sequence-specific amplification polymorphisms showing multiple polymorphic fragments detected among 24 strawberry cultivars in the five selected LTR retrotransposon families

LTR family	Restriction enzyme	Number of detected fragments[a]		Number of cultivar-specific fragments			Number of cultivars with specific fragment
		Total (A)	Average of cultivars	Total (B)	B/A (%)	Average of cultivars	
CL11	*Mse*I	143	6.0	8	5.6	0.3	7
	*Rsa*I	152	6.3	34	22.4	1.4	11
	Total	295	12.3	42	14.2	1.8	14
CL20	*Mse*I	382	15.9	31	8.1	1.3	15
	*Rsa*I	135	5.6	17	12.6	0.7	8
	Total	517	21.5	48	9.3	2.0	17
CL28	*Mse*I	1353	56.4	25	1.8	1.0	10
	*Rsa*I	235	9.8	20	8.5	0.8	8
	Total	1588	66.2	45	2.8	1.9	15
CL76	*Mse*I	438	18.2	38	8.7	1.6	18
	*Rsa*I	364	15.2	50	13.7	2.1	19
	Total	802	33.4	88	11.0	3.7	24
CL3	*Mse*I	1362	56.7	15	1.1	0.6	8
	*Rsa*I	744	31.0	23	3.1	1.0	9
	Total	2106	87.7	38	1.8	1.6	13

[a] Fragments with a peak height of 100 or more and a peak size of 40–700 bp were counted; however, fragments with a peak height of 50 or more from *Rsa*I treatment of CL11, CL20, CL28, and CL76 were also counted.

Table 3. Number of cultivar-specific fragments detected among 24 strawberry cultivars by sequence-specific amplification polymorphism analysis

			Sample No.[a]																								
			1	4	5	6	7	9	10	11	12	13	15	17	18	20	22	23	24	25	29	30	31	32	48	55
LTR	Restriction enzyme	Number of cultivers with specific fragments	FukuokaS6	Aiberry	Akihime	Sagahonoka	Sachinoka	Toyonoka	Benihoppe	Redpearl	Kumaken I 548	Houkouwase	Fukuba	Kaorino	Kotoka	Ookimi	Miyazakinatsuharuka	Elsanta	Floridabell	06A-184	Yumenoka	Satsumaotome	Marihime	Amaotome	Tochiotome	Donner
CL11	*Mse*I	7	1	1										1	2			1				1	1			
	*Rsa*I	11	3		4		1		1	5			3	4								1	4	1		7
CL20	*Mse*I	15	1		1				3			2	1	4	1	2	5		2	4		1	1	2	1	
	*Rsa*I	8			2		1				1	3	2									4	3		1	
CL28	*Mse*I	10		1	2		2					1	7		3	4		1	2						2	
	*Rsa*I	8			2		1		3												2		5	4	2	1
CL76	*Mse*I	18	1	2	1	1	1	1				3	4	3	2	2	1	2	3	5	2	3			1	
	*Rsa*I	19	5	6	7	2	2		5	2	1		2	2	1		1		1	1	2		4	1	2	3
CL3	*Mse*I	8	1										1	1				1	3	2	4					2
	*Rsa*I	9				4			6		1		3				2			1	2		3			1
		Total	12	10	19	7	8	1	18	7	3	9	23	15	9	8	9	5	11	13	12	10	21	8	9	14

[a] No. of each cultivar corresponds to those listed in **Table 1**.

Table 4. Summary of CL3 retrotransposon insertion sites in 75 strawberry cultivars and lines

	Number of insertion sites	Number of insertion sites per cultivar	Number of polymorphic insertion sites	Proportion of sites polymorphic (%)	Number of cultivar-specific sites	Number of cultivars possessing specific sites	Proportion of sites cultivar-specific (%)
Total	807	150.9	771	95.5	338	63	41.9
Number of sites with ≥100 reads in 75 cultivars	478	146.6	474	99.1	63	20	13.4

sites that were cultivar-specific, or showed polymorphism among cultivars, were screened for cultivar assessment, and DNA markers were designed based on the presumption that they could be used for identification of individual cultivars.

Development and verification of cultivar-specific markers

PCR analysis was performed using primers designed based on the identified putative cultivar-specific CL3 insertion sites (**Fig. 1**). Testing 36 candidate markers derived from 20 cultivars possessing putative cultivar-specific CL3 insertion sites, cultivar-specific bands among the 75 cultivars were detected for the following eight cultivars: 'Harunoka', 'Fukuba', 'Ookimi', 'Miyazakinatsuharuka', 'Elsanta', 'Deco-Rouge', 'Summer Tiara', and 'Otome-gokoro' (**Fig. 2**).

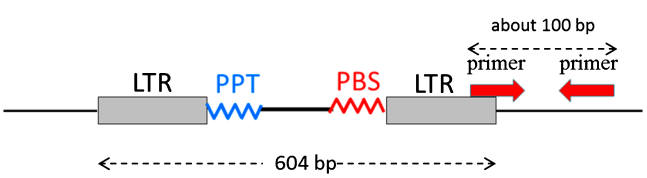

Fig. 1. Schematic illustration of the genomic structure of the CL3 retrotransposon and the location of PCR primers designed to detect putative cultivar-specific sites. The LTR sequence is shown in **Supplemental Fig. 1**. PPT, polypurine tract; PBS, primer-binding site.

To distinguish strawberry cultivars for which no cultivar-specific CL3 insertion sites were detected, we investigated markers that could differentiate cultivars based on the presence/absence pattern of CL3 insertions. PCR revealed that, of the 46 candidate markers, 29 markers were polymorphic among the 75 cultivars and were able to differentiate 72 cultivars (data not shown). Based on the calculation results using Minimal Marker, we were able to differentiate 35 common cultivars in Japan using a combination of at least 10 markers (**Table 5**). The predicted insertion sites and amplification results of the 75 cultivars using these markers are shown in **Supplemental Table 5**. In addition, the markers designed based on the insertion sequence CL3_Cl 214 could be detected in all 75 cultivars. Therefore, this marker was used as a positive control marker for identification of strawberry cultivars.

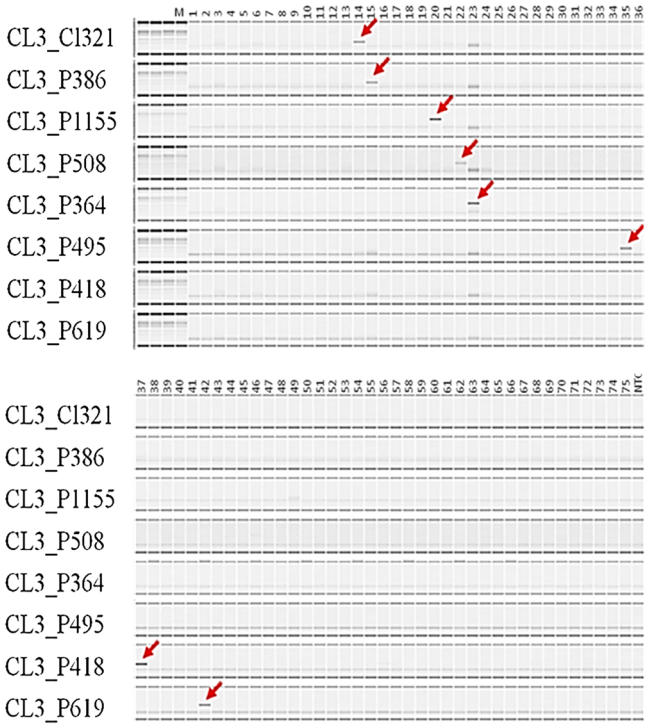

Fig. 2. PCR amplification of cultivar-specific CL3 retrotransposon-based markers. MultiNA gel image showing the markers developed in this study. CL3_Cl321, CL3_P386, CL3_P1155, CL3_P508, CL3_P364, CL3_P495, CL3_P418, and CL3_P619 are cultivar-specific markers developed to discriminate the strawberry cultivars 'Harunoka', 'Fukuba', 'Ookimi', 'Miyazakinatsuharuka', 'Elsanta', 'Deco-Rouge', 'Summer Tiara', and 'Otomegokoro', respectively. Lane numbers correspond to the cultivar numbers listed in **Table 1** and NTC indicates no template control. Primer sequences for each marker are listed in **Supplemental Table 3**.

Marker applicability to processed products

To verify the applicability of retrotransposon-based DNA markers in processed materials, we tested the retrotransposon-based markers and one SSR marker using DNA extracted from commercially available jams labeled 'Made of "Fukuoka S6 (Amaou)" fruits' (jams A and B). We were able to reproduce the expected PCR band pattern using the retrotransposon-based markers and jam samples. PCR bands amplified by 10 retrotransposon-based markers for differentiation of cultivars (**Table 5**) and a positive control marker CL3_Cl214 were detected in jam and leaf samples of "Fukuoka S6 (Amaou)" with the same pattern (**Fig. 3**). By contrast, PCR bands with the size of 249-bp for the FVES3384 marker (SSR) were not detected in the jam samples. These results indicated that the retrotransposon-based markers were more effective than the SSR marker to discern cultivars in samples containing fragmented DNA.

We designed three markers with different PCR product sizes from the retrotransposon-insertion region CL3_Cl214, and investigated their applicability to two strawberry jams (jams C and D) and "Fukuoka S6 (Amaou)" leaf samples (**Fig. 4**). The concentration and quality of jam DNAs differed depending on the product (**Supplemental Table 6**), and the sharpness of the PCR amplification band also differed among samples. If the PCR product was as small as 100 bp, the product was amplified from jam DNA extracts, but for 200–300 bp PCR products the amplification was unstable or failed.

Discussion

Currently, markers for cultivar discrimination based on retrotransposon insertion polymorphisms have been developed for several crop species, such as sweet potato, apple

Table 5. Retrotransposon-based markers developed for discrimination of 35 commercially important strawberry cultivars grown in Japan

Sample No.[a]		1	4	5	6	7	9	10	11	12	17	18	20	21	22	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Marker Name	Product size (bp)	Fukuoka S6	Aiberry	Akihime	Sagahonoka	Sachinoka	Toyonoka	Benihoppe	Redpearl	Kumaken I 548	Kaorino	Kotoka	Ookimi	Himiko	MiyazakiNastsuharuka	Megumi	Yumenoka	Satsumaotome	Marihime	Amaotome	Pechika	Natsuakari	Deco-Rouge	Sammer-Candy	Sammer-Tiara	Natsuotome	Tochihitomi	Kitanokagayaki	Mouikko	Otomegokoro	Echigohime	Yayoihime	Nyoho	Tochihime	Skyberry	Tochiotome
CL3_Cl19	99	1	1	1	0	0	0	1	1	0	0	1	0	0	0	1	0	0	1	1	1	1	0	1	0	1	0	0	1	0	1	0	1	1	0	1
CL3_P524	114	1	0	0	0	1	1	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	0	0	0	1	0	0	0	1	1	1	1	1	1
CL3_Cl1124	78	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
CL3_Cl322	83	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0
CL3_Cl261	73	1	0	0	0	1	1	0	1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CL3_P320	89	0	1	0	0	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
CL3_Cl1115	73	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	1	1	0	1	0	1	0	0
CL3_Cl242	102	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CL3_Cl258	93	1	1	0	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0
CL3_Cl76	108	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0

Primer sequences are listed in **Supplemental Table 3**.
The characters (0 and 1) represent absence and presence, respectively, of the PCR band.
[a] No. of each cultivar corresponds with those listed in **Table 1**.

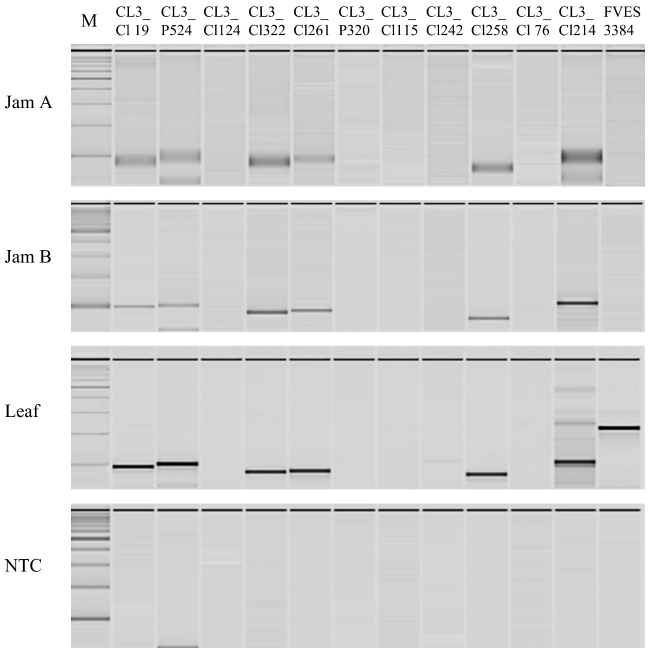

Fig. 3. DNA signal detection of retrotransposon-based and SSR markers for strawberry jam and leaf samples. MultiNA gel image showing markers developed in this study. CL3_Cl19, CL3_P524, CL3_Cl124, CL3_Cl322, CL3_Cl261, CL3_P320, CL3_Cl115, CL3_Cl242, CL3_Cl258, and CL3_Cl76 are a set of markers for discrimination of 35 commercial strawberry cultivars grown in Japan. Retrotransposon-based marker CL3_Cl214 and SSR marker FVES3384 are positive control markers. Jams A and B were labeled 'Made of "Fukuoka S6 (Amaou)" fruits' in the marketplace. NTC indicates no template control. M: 100 bp DNA ladder marker. Primer sequences for the markers are listed in **Supplemental Table 3**.

(*Malus domestica*), and citrus (*Citrus clementina* and *C. sinensis*) (Monden *et al.* 2014c, Nishitani *et al.* 2016, Tanaka *et al.* 2015). As retrotransposons are a major component of most plant genomes, this marker system should be applicable to a broad range of plant species.

In the present study, we performed a comprehensive search for retrotransposon insertion sites in strawberry cultivars via a high-throughput sequencing approach and developed eight cultivar-specific markers. The data demonstrated that markers derived from insertion sites exhibiting polymorphisms across cultivars were useful for differentiation of major domestic cultivars.

Using these markers, an unambiguous assessment based on the presence or absence of PCR amplification products derived from CL3 retrotransposon insertions was possible. The retrotransposon-based markers exhibit three important characteristics: they are chromosome-specificity, occasional cultivar-specificity and low molecular weight. In the present investigation, these features were considered to be advantageous for the following reasons.

First, the high sequence specificity of retrotransposon-based markers could overcome partial autoploidy. Given that strawberry is assumed to have an octaploid genome derived by partial autoploidy, a single marker may detect multiple genomic regions, which would result in increased complexity for cultivar identification. In the present study, however, most markers generated a single band because the primer sites were designed to bridge the boundary of LTR and genomic regions, which may generate chromosome-specific sites.

The results demonstrated the presence of cultivar-specific retrotransposon insertions in the genome. Although Akitake *et al.* (2013) used the retrotransposon *FaRE*1 to develop DNA markers for discrimination of strawberry cultivars, cultivar-specific *FaRE*1 insertion sites were not identified. It is possible that *FaRE*1 is rarely transposed during the cultivar breeding process or in derivative strains of a cultivar. In contrast, the detection of unique insertion sequences for the CL3 retrotransposon insertion in a number of cultivars in the present study suggests that CL3 may be transposed during the modern breeding process and within a cultivar's genealogy. Consequently, the CL3 retrotransposon insertion may have occurred recently after evolution of the genome was considerably advanced. Given that recent retrotransposon insertion could provide valuable information to identify cultivars, investigation of the timing

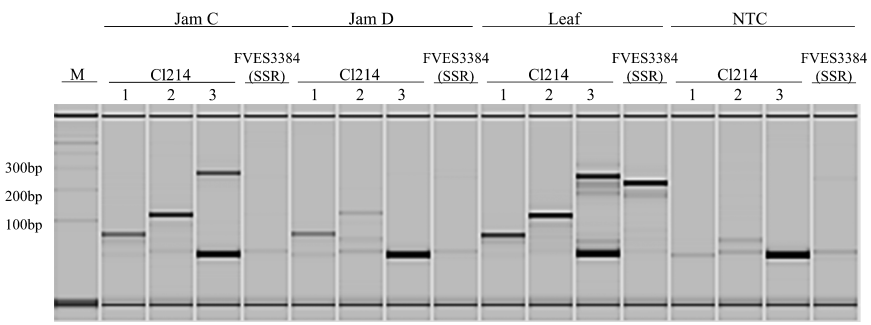

Fig. 4. DNA signal detection of control markers (retrotransposon-based and SSR) in strawberry jam and leaf samples. MultiNA gel image showing the markers developed in this study. Retrotransposon-based markers Cl214_1, Cl214_2 and Cl214_3 were designed as markers to yield PCR products of different sizes. Jams C and D were derived from an unknown strawberry cultivar. A PCR band for the SSR marker (FVES3384) was only detected in leaf samples, while a PCR band for the retrotransposon-based marker (CL3_Cl214) was detected in leaf and jam samples. NTC indicates no template control. M: 100 bp DNA ladder marker. Primer sequences for the markers are listed in **Supplemental Table 3**.

and background of the insertion may be effective for enhancing this identification technology.

Second, because the retrotransposon-based markers produce low-molecular-weight PCR products (approximately 100 bp), they are advantageous for discrimination of cultivars in processed products, such as jam, in which DNA is readily fragmented by heat and pressure. In addition, because PCR products specifically amplified in each cultivar were not detected in other cultivars, it is likely that these markers are useful for processed food samples that contain DNA from multiple cultivars.

In general, the quality of extracted DNA is supposed to have a large effect on genetic discrimination by PCR. In the test using 10 retrotransposon markers and DNA of jams A and B, the detection patterns of jams A and B were the same as that of leaf DNA. However, jam A showed slightly smeared bands compared with jam B. In the test using control markers and DNA of jams C and D, all of the markers were detected in jam C; however, the signals were very weak in jam D, and the marker Cl214_3 (300 bp) was not detected. From these results using the retrotransposon markers, detection seems to be possible even with DNA quality of O.D.260/O.D.280 > 1.15. However, DNA qualities likely influenced the detection sensitivity.

The markers developed in the present study are compatible with the single-tag hybridization (STH) chromatographic printed array strip (PAS) method, which is a rapid and convenient DNA marker detection method recently used for strawberries (Monden *et al.* 2014b, Monden and Tahara 2015), because this method can produce multiplex signals simultaneously in a single reaction using several independent PCR amplicons. It is expected that retrotransposon-based markers, in combination with the STH chromatographic PAS method, will facilitate rapid, efficient, and highly reliable cultivar discrimination, even for processed food products, in the future.

Author Contribution Statement

M.T. and Y.M. constructed the study concept. C.H. designed the study and wrote the initial draft of the manuscript. C.H., T.W., K.S., and Y.M. contributed to analysis and interpretation of data, and assisted in manuscript preparation. T.W., S.T., H.I., Y.U., K.H., Y.N., K.O., and K.N. have contributed to data collection and interpretation, and reviewed the manuscript. All authors approved the final version of the manuscript.

Acknowledgments

This study was supported by a Research and Development Projects for Application in Promoting New Policy of Agriculture Forestry and Fisheries grant from the Ministry of Agriculture, Forestry and Fisheries of Japan. We thank Edanz Group (www.edanzediting.com/ac) for reviewing the English in a draft of this manuscript.

Literature Cited

Akitake, H., M. Tahara, Y. Monden, K. Takasaki and S. Futo (2013) Strawberry cultivar identification by retrotransposon insertion polymorphisms. DNA Polymorph. 21: 64–72.

Arnau, G., J. Lallemand and M. Bourgoin (2003) Fast and reliable strawberry cultivar identification using inter simple sequence repeat (ISSR) amplification. Euphytica 129: 69–79.

Bonoli, M., G. Cipriani, L. Manzecchi and W. Faedi (2005) Utilization of microsatellite molecular markers for characterization of varieties in strawberry [*Fragaria Spp.*]. Rivista Di Frutticoltura E Di Ortofloricoltura 67.

Degani, C., L.J. Rowland, A. Levi, J.A. Hortynski and G.J. Galletta (1998) DNA fingerprinting of strawberry (*Fragaria* × *ananassa*) cultivars using randomly amplified polymorphic DNA (RAPD) markers. Euphytica 102: 247–253.

Degani, C., L.J. Rowland, J.A. Saunders, S.C. Hokanson, E.L. Ogden, A. Golan-Goldhirsh and G.J. Galletta (2001) A comparison of genetic relationship measures in strawberry (*Fragaria* × *ananassa* Duch.) based on AFLPs, RAPDs, and pedigree data. Euphytica 117: 1–12.

Flavell, A.J., M.R. Knox, S.R. Pearce and T.H.N. Ellis (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. Plant J. 16: 643–650.

Fujii, H., T. Ogata, T. Shimada, T. Endo, H. Iketani, T. Shimizu, T. Yamamoto and M. Omura (2013) Minimal Marker: An algorithm and computer program for the identification of minimal sets of discriminating DNA markers for efficient variety identification. J. Bioinform. Comput. Biol. 11: 1250022.

Fukuchi, A., F. Kikuchi and H. Hirochika (1993) DNA fingerprinting of cultivated rice with rice retrotransposon probes. Jpn. J. Genet. 68: 195–204.

He, P., Y. Ma, G. Zhao, H. Dai, H. Li, L. Chang and Z. Zhang (2010) *FaRE1*: a transcriptionally active Ty1-*copia* retrotransposon in strawberry. J. Plant Res. 123: 707–714.

Honjo, M., T. Nunome, S. Kataoka, T. Yano, H. Yamazaki, M. Hamano, S. Yui and M. Morishita (2011) Strawberry cultivar identification based on hypervariable SSR markers. Breed. Sci. 61: 420–425.

Isobe, S.N., H. Hirakawa, S. Sato, F. Maeda, M. Ishikawa, T. Mori, Y. Yamamoto, K. Shirasawa, M. Kimura, M. Fukami *et al.* (2013) Construction of an integrated high density simple sequence repeat linkage map in cultivated strawberry (*Fragaria* × *ananassa*) and its applicability. DNA Res. 20: 79–92.

Kent, W.J. (2002) BLAT—the BLAST-like alignment tool. Genome Res. 12: 656–664.

Kumar, A. and J.L. Bennetzen (1999) Plant retrotransposons. Annu. Rev. Genet. 33: 479–532.

Kumar, A. and H. Hirochika (2001) Applications of retrotransposons as genetic tools in plant biology. Trends Plant Sci. 6: 127–134.

Kunihisa, M., N. Fukino and S. Matsumoto (2003) Development of cleavage amplified polymorphic sequence (CAPS) markers for identification of strawberry cultivars. Euphytica 134: 209–215.

Kunihisa, M., N. Fukino and S. Matsumoto (2005) CAPS markers improved by cluster-specific amplification for identification of octoploid strawberry (*Fragaria* × *ananassa* Duch) cultivars, and their disomic inheritance. Theor. Appl. Genet. 110: 1410–1418.

Kunihisa, M., H. Ueda, N. Fukino and S. Matsumoto (2009a) DNA markers for identification of strawberry (*Fragaria* × *ananassa* Duch.) cultivars based on probability theory. J. Japan. Soc. Hort.

Sci. 78: 211–217.

Kunihisa, M., H. Ueda, N. Fukino and S. Matsumoto (2009b) Genotyping of strawberry (*Fragaria* × *ananassa* Duch.) cultivars by DNA markers: Interlaboratory study. J. AOAC Int. 92: 896–906.

Kunihisa, M. (2010) Development of genome-specific DNA markers in strawberry (*Fragaria* × *ananassa* Duch.) and their use for cultivar identification. Bull. Nat. Inst. Veg. Tea Sci. 9: 7–56.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez *et al.* (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.

Monden, Y., N. Fujii, K. Yamaguchi, K. Ikeo, Y. Nakazawa, T. Waki, K. Hirashima, Y. Uchimura and M. Tahara (2014a) Efficient screening of long terminal repeat retrotransposons that show high insertion polymorphism via high-throughput sequencing of the primer binding site. Genome 57: 245–252.

Monden, Y., K. Takasaki, S. Futo, K. Niwa, M. Kawase, H. Akitake and M. Tahara (2014b) A rapid and enhanced DNA detection method for crop cultivar discrimination. J. Biotechnol. 185: 57–62.

Monden, Y., A. Yamamoto, A. Shindo and M. Tahara (2014c) Efficient DNA fingerprinting based on the targeted sequencing of active retrotransposon insertion sites using a bench-top high-throughput sequencing platform. DNA Res. 21: 491–498.

Monden, Y. and M. Tahara (2015) Plant transposable elements and their application to genetic analysis via high-throughput sequencing platform. Hort. J. 84: 283–294.

Nishitani, C., T. Yamamoto, H. Fujii, K. Okada, Y. Monden and M. Tahara (2016) Development of retrotransposon-based insertion polymorphism marker in apple. DNA Polymorph. 24: 101–107.

Ooe, N., M. Tahara, H. Yamashita, Y. Marutani and T. Kuranouchi (2004) Cultivar identification technique applied for Hoshi-imo, steamed and dried sweetpotato slices, based on retrotransposon insertion polymorphism. Breed. Res. 6: 169–177.

Shimomura, K., H. Mitsui, K. Fujita and K. Satoh (2005) Identification of strawberry cultivar 'Fukuoka S6 (*Fragaria* × *ananassa* Duch.)' by amplified fragment length polymorphism. Bull. Fukuoka Agric. Res. Cent. 24: 43–47.

Shimomura, K. and K. Hirashima (2006) Development and characterization of simple sequence repeats (SSR) as markers to identify strawberry cultivars (*Fragaria* × *ananassa* Duch.). J. Japan. Soc. Hort. Sci. 75: 399–402.

Tahara, M., H. Yamashita and A. Nakagawa (2009) Katsudogata no teniinshi niyoru hinsyukoyu ma-ka- no kaihatsu. DNA-Kantei 1: 51–65.

Tanaka, M., Y. Monden, M. Tahara, Y. Kai, Y. Okada and Y. Takahata (2015) Screening of retrotransposon *Rtsp-1* insertion site useful as DNA markers for identification of sweetpotato cultivar "Beniharuka". Rep. Kyushu Br. Crop Sci. Soc. 81: 43–45.

Waugh, R., K. McLean, A.J. Flavell, S.R. Pearce, A. Kumar, B.B. Thomas and W. Powell (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol. Gen. Genet. 253: 687–694.

Yamashita, H., M. Tahara and Y. Ooyama (2008) Retrotransposon DNA marker for azuki cultivar identification. DNA Polymorph. 16: 82–87.

Yano, H. (2010) DNA hinsyushikibetsu gijutsu no genjo to syoraitenbo (tokusyu kajurui no DNA hinsyuhanbetsu gijutsu (5, 6)). J. Agric. Sci. (Nogyo-Gijutsu) 65: 478–483.