Development of SSR markers for identification and discrimination of tobacco cultivars

SATO S.; TAJIMA T.; KASAOKA K.; NOGUCHI S.

Japan Tobacco Inc., Leaf Tobacco Research Center, 1900 Idei, Oyama, Tochigi 323-0808, Japan

Summary

Knowledge of the genetic background in tobacco cultivars is very important not only for genetic evaluation in breeding but also for management of germplasms and protection of cultivars. DNA marker technology is one of the powerful tools for such purposes.

Simple Sequence Repeat (SSR) markers were developed for identification and discrimination of tobacco cultivars. It is known that SSR markers are abundant in plant genomes, and variations of the repeat number among different lines within a species occur at high frequency. These characteristics, along with an evenly dispersed genomic distribution, are properties which make SSR markers ideal genetic markers.

The SSR-enriched genomic libraries for AG-, CT- and TAA-repeat motifs were constructed using the magnetic beads method based on the hybridization between the biotin-labeled probes and the target nucleic acid molecules. Following determination of the DNA sequences of clones in each library, primer pairs specific for SSR loci were synthesized. After the primers were tested for the presence of polymorphism among eight tobacco cultivars, the resulting functional and polymorphic SSR markers were tested for 84 tobacco cultivars consisting of flue-cured, burley, oriental and other types.

Twenty six out of 36 SSR markers were identified for use in discriminating the 84 cultivars. To facilitate the genotyping of cultivars, a multiplex PCR system was employed in which the 26 SSR markers were sorted and bulked into 5 groups.

These SSR markers were confirmed to be useful the genotyping of DNA prepared not only from fresh materials but also from cured leaves.

Introduction

The analysis of genetic diversity and the relationships among different cultivars of tobacco has great importance for genetic resource conservation and breeding potential. The DNA marker is a powerful tool for us to understand the genetic diversity and relationships of tobacco easily and accurately.

A simple sequence repeat (SSR) is a valuable DNA marker in plants because of its abundance and high degree of polymorphism. The repeated sequence often consists of a unit with 2 to 4 nucleotides and the unit repeats 3 to 100 times. The SSR region is amplified by the polymerase chain reaction (PCR) using the unique sequences in the flanking regions as DNA primers. The difference in the molecular size of the PCR products based on the difference in the number of repeats gives a polymorphism to distinguish germplasms. SSR markers have been used for cultivar identification in crop species such as pear, quince, soybean, and durum wheat (Kimura et al. 2002, Osono et al. 2003, Yamamoto et al. 2004, and Perry et al. 2004).

In this study, we successfully developed SSR markers for discrimination of 84 tobacco cultivars using DNA from fresh leaves. We also confirmed that these markers can be applied to the analysis of cured leaf samples.

Materials and Methods

DNA extraction

Total DNA was extracted from leaf samples using a Puregene® Genomic DNA Purification Kit (Gentra Systems) according to the manufacturer's directions.

Screening of candidate SSR markers

The total DNA was extracted from a leaf of a seedling of Nicotiana tabacum cv. "Bright Yellow 4." The genomic DNA fragments containing SSR sequences were enriched using a biotin-labeled bait oligo-sequence attached to streptavidin-coated magnetic beads as described by Kijas et al. (1994). Three kinds of beads attaching bait sequences of (CT) 12, (AC) 12, and (TAA) 8 repeats were respectively used. The DNA fragments hybridized with these beads were recovered and cloned into plasmids to construct an SSR-enriched plasmid library. The 1,152 clones were randomly selected from the library, and DNA sequences of their inserts were determined. After identification of SSRs, PCR primer pairs for each SSR locus were designed in the flanking regions. Eight typical cultivars such as burley, flue-cured, domestic, or others were used to check whether or not these SSR loci are polymorphic.

Cultivar identification by SSR markers

For the analysis of fresh leaf samples, DNA was extracted from the fresh leaves of 84 cultivars developed in Japan and foreign countries (22 from burley, 28 from flue-cured, 12 from oriental, and 22 from other types). For analysis of the cured samples, commercial cured leaves were collected from different harvesting areas in Japan (Coker 319 and Virginia 115 as flue-cured, and Michinoku 1 and Taihei as burley). The SSR loci were amplified by multiplex PCR and separated by electrophoresis using a capillary sequencer, the 3100 Genetic Analyzer (Applied Biosystems).

Results and Discussion

Development of SSR markers

A total of 46 candidate SSR markers were obtained from randomly selected 1,152 clones in the SSR enrichment library. The second selection was carried out to select a valuable SSR marker set. First, the markers were sorted into groups for multiplex PCR reaction based on the optimal PCR condition and PCR product size. Multiplex PCR is a high-throughput method and makes it possible to amplify plural SSR loci at the same time in the same reaction mixture. Next, markers that showed good reproducibility of PCR and showed a different typing pattern with minimized redundancy were selected. Finally, 26 SSR markers sorted into 5 groups were selected.

Identification of the 84 tobacco cultivars using flesh leaf materials

All of the 84 cultivars were successfully discriminated by the 26 SSR markers. The DNA typing data of 22 burley cultivars are shown in Figure 1 as an example. This suggests that the 26 SSR markers will be useful in detecting contamination or segregation of seeds for management of the purity of germplasms.

Identification of tobacco cultivars using cured leaf samples

The DNA typing of cured leaf samples was carried out, and their originating cultivars were examined. The results are shown in Figure 2. The typing data of cured leaf samples originated from Coker 319 perfectly coincided with those of Coker 319. Although the typing data of the other 3 samples did not always coincide with those of their originating cultivars, the best matching cultivar was always the originating one. As a result, all of the cured leaf samples were estimated correctly. These results suggest that the developed 26 SSR markers are useful for cultivar estimation of cured leaf samples.

Cultivars	Group1					Group2							Group3					Group4				Group5				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
B21	153	170	151	219	274	155	183	154	92	88	141	213	163	125	187	117	140	159	175	139	99	187	128	189	203	203
KY10	153	168	155	219	272	155	181	154	96	88	143	213	163	125	187	117	140	159	175	139	99	177	130	189	213	203
White Burley	153	170	155	219	274	155	183	154	92	88	141	199	163	125	187	117	138	157	175	141	111	185	128	189	201	203
Michinoku 1	153	170	153	219	274	155	183	154	92	88	173	213	163	125	187	121	140	159	175	139	111	185	128	191	203	203
TN86	153	170	153	219	274	155	183	154	92	88	141	213	163	125	187	117	140	167	175	139	99	185	128	189	203	203
TN90	153	170	151	219	270	153	183	154	92	0	141	213	163	125	187	117	140	159	175	139	99	185	128	189	203	203
KY14	153	170	155	219	274	157	181	154	92	88	143	213	163	127	187	117	140	159	175	137	99	173	130	189	203	203
BANKET A1	153	170	153	219	274	155	183	154	92	88	165	205	163	123	187	121	140	159	175	139	99	183	128	189	201	203
KY17	153	170	157	219	274	157	183	154	94	88	139	213	163	125	187	117	140	167	175	139	99	185	130	189	213	203
MBN2	153	170	153	219	274	157	181	154	92	88	141	213	171	125	187	117	140	159	175	139	99	177	128	189	203	203
KY907	153	170	153	219	274	157	179	154	92	0	141	213	163	125	187	117	140	167	175	139	99	185	130	189	215	203
Taihei	153	170	153	219	274	157	183	154	92	88	141	213	163	125	187	117	140	159	175	139	99	185	128	191	203	203
WWH155	153	170	153	219	274	155	183	154	92	88	141	213	163	125	187	117	140	159	175	139	99	187	128	189	203	203
LAB21	153	170	151	219	274	155	183	154	92	88	141	213	163	125	187	117	140	159	175	139	99	185	128	189	203	203
Alta	149	172	155	219	274	157	191	164	94	88	143	205	165	127	187	123	142	159	175	143	99	179	126	187	203	197
L8	149	168	153	219	274	155	183	154	92	88	141	205	163	125	187	117	138	159	175	139	142	185	130	189	203	203
Michinoku 2	153	170	151	219	274	155	183	154	92	88	141	213	163	125	187	121	140	159	175	141	99	167	128	191	193	199
BY509	155	168	151	219	274	155	191	156	92	0	141	213	163	125	187	117	140	167	175	139	99	169	128	189	201	203
CSC200	155	168	151	219	274	155	193	156	92	0	141	213	163	125	187	117	140	167	175	139	99	167	128	189	201	203
KY16	151	170	153	219	274	157	183	154	92	88	141	199	163	125	187	117	140	159	175	137	99	185	130	189	205	203
KY26	151	170	153	219	274	157	183	154	92	88	141	199	163	125	187	117	140	159	175	139	99	183	130	187	205	203
KY56	151	170	153	219	274	157	183	154	92	88	141	199	163	125	187	117	140	159	175	139	99	185	130	189	205	203

Figure 1. Identification of 22 burley tobacco cultivars by 26 SSR markers

The different colors of the columns for each SSR marker mean the existence of polymorphism between germplasms. All cultivars were identified by each of them.

Cured leaf samples			Estimated cultivar			
			The best match		The 2nd-best match	
Cultivar		No. of samples	Cultivar	No. of 100% matches (% of average)	Cultivar	No. of 100% matches (% of average)
Flue-cured	Coker 319	12	Coker 319	12/12 (100)	MC1	0/12 (85)
	Virginia 115	12	Virginia 115	9/12 (98)	Delgold	0/12 (88)
Air-cured	Michionoku 1	11	Michionoku 1	7/11 (98)	Taihei	0/11 (83)
	Taihei	10	Taihei	9/10 (99)	B21	0/10 (95)

Figure 2. Tobacco cultivar identification in cured leaf samples with 26 SSR markers

The left-hand columns indicate the originating cultivars of the cured leaf samples and the number of samples. On the other hand, the right-hand columns indicate the estimated cultivars, the best matches, and the second-best matches among the 84 cultivars. The figures in the parentheses mean the percentile of matching of DNA type.

References

Kimura, T., Y. Z. Shi, M. Shoda, K. Kotobuki, N. Matsuta, T. Hayashi, Y. Ban and T. Yamamoto (2002) Identification of Asian pear varieties by SSR analysis. Breed. Sci. 52: 115-121.

Osono, M., T. Kimura and Y. Ban (2003) Discrimination of soybean varieties (registered varieties as a vegetable and imported frozen products) with Satt markers. Jpn. J. Breed. 5 (Suppl. 2): 230.

Yamamoto, T., T. Kimura, J. Soejima, T. Sanada, Y. Ban and T. Hayashi (2004) Identification of quince varieties using SSR markers developed from pear and apple. Breed. Sci. 54: 239-244.

Perry, D. J. (2004) Identification of Canadian durum wheat varieties using a single PCR. Theor. Appl. Genet. 109: 55-61.

Kijas, J.M.H., J.C.S. Fowler, C.A. Garbett, and M.R. Thomas (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques 16: 657-662