Theor Appl Genet (2007) 115:1067–1074
DOI 10.1007/s00122-007-0632-z

ORIGINAL PAPER

The number of genes having different alleles between rice cultivars estimated by SNP analysis

Kenta Shirasawa · Hiroaki Maeda · Lisa Monna · Sachie Kishitani · Takeshi Nishio

Received: 23 March 2007 / Accepted: 10 August 2007 / Published online: 6 September 2007

Abstract Identification of single nucleotide polymorphisms (SNPs) in a large number of genes will enable estimation of the number of genes having different alleles in a population. In the present study, SNPs between 21 rice cultivars including 17 Japanese cultivars, one upland rice, and three *indica* cultivars were analyzed by PCR–RF–SSCP. PCR–RF–SSCP analysis was found to be a more efficient method for detecting SNPs than mismatch–cleavage analysis, though both PCR–RF–SSCP and mismatch–cleavage are useful for screening SNPs. The number of DNA fragments showing polymorphism between Japanese cultivars was 134 in the 1,036 genes analyzed. In 137 genes, 638 DNA polymorphisms were identified. Out of 52 genes having polymorphisms in the exons, one had a frame-shift mutation, three had polymorphism causing amino acid insertions or deletions, and 16 genes had missense polymorphisms. The number of genes having frame-shift mutations and missense polymorphisms between the 17 Japanese cultivars was estimated to be 41 and 677 on average, respectively, and those between *japonica* and *indica* to be 425 and 6,977, respectively. Chromosomal regions of cultivars selected in rice breeding processes were identified by SNP analysis of genes.

Communicated by Q. Zhang.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0632-z) contains supplementary material, which is available to authorized users.

K. Shirasawa · H. Maeda · S. Kishitani · T. Nishio (✉)
Graduate School of Agricultural Science,
Tohoku University, Aoba-ku, Sendai, 981-8555, Japan
e-mail: nishio@bios.tohoku.ac.jp

L. Monna
Plant Genome Center Co., Ltd,
Tsukuba 305-0856, Japan

Introduction

The nucleotide sequence of the rice genome (*Oryza sativa* L. ssp. *japonica* cv. 'Nipponbare') has been published by the International Consortium on Rice Genome Sequencing (International Rice Genome Sequencing Project 2005), and the number of genes in rice has been estimated to be about 32,000 (The Rice Annotation Project 2007). However, few genes have so far been found to be responsible for different phenotypes between rice cultivars, e.g., *Xa21* (Song et al. 1995), *Hd1* (Yano et al. 2000), *Gn1* (Ashikari et al. 2005), *qSH1* (Konishi et al. 2006), and *Sub1A* (Xu et al. 2006), some of which have been identified by quantitative trait loci (QTL) analysis and a map-based cloning strategy.

Comparison between the nucleotide sequence of the *japonica* rice genome (Goff et al. 2002; International Rice Genome Sequencing Project 2005) and that of the *indica* rice genome (Yu et al. 2002) has revealed a large number of single nucleotide polymorphisms (SNPs) (Yu et al. 2002; Feltus et al. 2004; Shen et al. 2004). SNPs between *japonica* rice cultivars have been identified by nucleotide sequencing of randomly selected intergenic regions (Nasu et al. 2002; Monna et al. 2006). Detection of SNPs in a large number of genes would enable estimation of the number of genes having different alleles in a population, and furthermore may contribute to identification of SNPs responsible for different phenotypes between cultivars. Detection of SNPs between *japonica* rice cultivars is, however, costly and time consuming because the frequency of SNPs is lower in gene regions than in the intergenic regions (Han and Xue 2003).

PCR–RF–SSCP analysis, which is SSCP analysis of DNA fragments digested by restriction enzyme, has high ability to detect SNPs (Sato and Nishio 2002). Point mutations in the *wx* gene of rice and SNPs in *japonica* rice cultivars have been detected efficiently by such analysis (Sato

and Nishio 2003; Shirasawa et al. 2004). As a method for detecting mutations in a mutagenized population, TILLING, which is based on cleavage of mismatch sites in heteroduplex DNA by CEL I endonuclease derived from celery (McCallum et al. 2000), is widely used for studies of plant and animal genetics. This technique has been applied to the screening of natural variations among ecotypes of *Arabidopsis thaliana* (L.) Heynh. (Comai et al. 2004). Mismatch–cleavage endonuclease derived from *Brassica rapa* L. has also been used for selection of radiation-induced mutations in rice (Sato et al. 2006).

In conventional rice breeding programs, two cultivars or lines are crossed and elite lines are selected from their progeny. Population size of the progeny is determined empirically. Information of the number of genes having different alleles between the two parental cultivars or lines will enable estimation of optimum population sizes of the progenies. However, such information is not available. In the present study, we screened and identified SNPs of the genes between rice cultivars, and estimated the number of the genes having frame-shift or missense polymorphism between the cultivars, which have been used in the conventional rice breeding. For this estimation, efficiencies of DNA polymorphism detection by PCR–RF–SSCP and mismatch–cleavage analysis were investigated.

Materials and methods

Plant materials and DNA isolation

Twenty-one cultivars of rice were used as plant materials. Among them, 17 were *japonica* cultivars, i.e., 'Akihikari', 'Hatsuboshi', 'Hatsunishiki', 'Hinohikari', 'Hitomebore', 'Itadaki', 'Koganebare', 'Koshihikari', 'Kihou', 'Kirara 397', 'Nipponbare', 'Nourin 1', 'Nourin 6', 'Nourin 8', 'Nourin 22', 'Rikuu 132', and 'Sasanishiki', which are leading Japanese cultivars or their parental cultivars, one was the upland variety 'Senshou', and three were *indica* cultivars, i.e., 'Basmati 370', 'IR36', and 'Kasalath'. Genomic DNA was isolated from a leaf by the CTAB method (Murray and Thompson 1980).

PCR–RF–SSCP analysis

To amplify 2-kb genomic DNA fragments of genes, primer pairs were designed from the genomic DNA sequence with the EST sequences published in a database (Wu et al. 2002; http://rgp.dna.affrc.go.jp/publicdata/estmap2001/index.html) using Primer3 software (Rozen and Skaletsky 2000). Analyzed genes were arbitrarily selected so as to obtain an even distribution of the genes on the rice chromosomes. Half the volume of the PCR mixture was treated with *Afa*I, and the other half was treated with *Alu*I. The digested PCR products were mixed with four volumes of denaturing solution before electrophoresis. Electrophoresis and detection of DNA fragments were performed as described previously (Shirasawa et al. 2004).

Mismatch–cleavage analysis

A 1:1 mixture of genomic DNA of 'Nipponbare' and that of 'Koshihikari', a leading cultivar in Japan, or 'Kasalath', was used for DNA amplification by PCR. The amplified DNAs were reannealed to form a heteroduplex and digested by *Brassica* petiole extract (BPE) prepared from petioles of *Brassica rapa* L. var. *laciniifolia* according to Sato et al. (2006). The digested sample was subjected to electrophoresis on 2% agarose gel in TAE buffer. The resulting DNA bands were stained with ethidium bromide.

DNA sequence analysis

The DNA fragments showing polymorphisms were sequenced by direct sequencing with a CEQ2000 sequencer (Beckman Coulter, CA, USA), an ABI PRISM 310 genetic analyzer (Applied Biosystems, CA, USA), or a MegeBACE 1000 DNA Sequencing System (Amersham Biosciences, NJ, USA). Prediction of exons and introns, membrane-spanning domain, cellular localization, and motifs was carried out using Spidey (Wheelan et al. 2001: http://www.ncbi.nlm.nih.gov/spidey/), SOSUI (Hirokawa et al. 1998: http://bp.nuap.nagoya-u.ac.jp/sosui/), PSORT (Nakai and Horton 1999: http://psort.ims.u-tokyo.ac.jp/), and InterPro (Zdobnov and Apweiler 2001: http://www.ebi.ac.uk/interpro/index.html), respectively.

Results

Comparison of SNP detection ability between PCR–RF–SSCP and mismatch–cleavage analysis

Efficiencies of SNP detection were compared between PCR–RF–SSCP and mismatch–cleavage analysis. Of 120 randomly selected genes, 87 genes showed polymorphism between 'Nipponbare' and 'Kasalath' by PCR–RF–SSCP. Mismatch–cleavage analysis using agarose gel electrophoresis detected DNA polymorphisms of 64 genes, including the 50 genes whose DNA polymorphisms were also detected by PCR–RF–SSCP (Table 1). Between 'Nipponbare' and 'Koshihikari', five of the 120 genes showed polymorphism by PCR–RF–SSCP, while only three genes, whose polymorphisms were also detected by PCR–RF–SSCP, showed polymorphism by mismatch–cleavage analysis. In total, DNA polymorphisms of 106 of the 240

Table 1 Comparison of SNP detection ability between PCR–RF–SSCP and mismatch–cleavage analysis

Cultivars	No. of tested genes	No. of genes having polymorphisms detected by			Total no. of polymorphic genes
		Both PCR–RF–SSCP and mismatch–cleavage analysis	Only PCR–RF–SSCP analysis	Only mismatch–cleavage analysis	
Nipponbare vs. Kasalath	120	50	37	14	101
Nipponbare vs. Koshihikari	120	3	2	0	5
Total	240	53	39	14	106

combinations of genes and cultivars were detected by PCR–RF–SSCP and/or mismatch–cleavage analysis. PCR–RF–SSCP analysis was found to be a more efficient method for detecting SNPs than mismatch–cleavage analysis. Therefore, PCR–RF–SSCP was used for the screening of SNPs between rice cultivars.

DNA polymorphisms in genes between 21 rice cultivars

Single DNA fragments were amplified from genomic DNA of 'Nipponbare' by 1,039 primer pairs (58.5%) in 1,777 pairs tested. In all 17 Japanese cultivars, 1,036 primer pairs amplified single DNA fragments, but E3928, E1141, and C10295 did not amplify single fragments in one, seven, and nine cultivars, respectively (Supplementary Table 1). In 'Basmati 370', 'IR36', 'Kasalath', and 'Senshou', 33, 41, 41, and 17 of the 1,039 primer pairs, respectively, did not amplify single fragments.

PCR–RF–SSCP analysis using the 1,039 primer pairs detected DNA polymorphism of 789 (75.9%) DNA fragments between 'Nipponbare' and 'Kasalath', but detected DNA polymorphism of only 51 (4.9%) DNA fragments between 'Nipponbare' and 'Koshihikari' (Table 2). In pairwise comparisons between the Japanese cultivars, the number of DNA fragments showing polymorphism was 46.5 (4.5%) on average, ranging from 7 (0.7%) between 'Hatsuboshi' and 'Hitomebore' to 79 (7.6%) between 'Kirara 397' and 'Nourin 22'. The number of DNA fragments having polymorphism between the three *indica* cultivars was 619.7 (59.6%) on average, ranging from 587 (56.5%) between 'IR36' and 'Kasalath' to 647 (62.3%) between 'Basmati 370' and 'Kasalath'. The average number of DNA fragments showing polymorphism between the *japonica* cultivars and the *indica* cultivars was 710.1 (68.3%).

Two agronomically important genes, i.e., *Hd1* for heading date (Yano et al. 2000) and *Sd1* for culm length (Sasaki et al. 2002), of the 21 cultivars, were also analyzed by PCR–RF–SSCP. Two DNA fragments amplified by PCR covered the entire gene regions in the case of both *Hd1* and *Sd1*. 'Nipponbare'-type leaky, 'HS66'-type null, and 'Ginbouzu'-type functional alleles of *Hd1* were found in seven, eight, and two Japanese cultivars, respectively (Supplementary Table 1). Regarding the *Sd1* gene, only 'Akihikari' was revealed to have *sd1-r* ('Reimei'-type allele) and 'Itadaki' as well as 'IR36' and 'Basmati 370' had *sd1-d* ('Dee Geo Woo Gen'-type allele) (Supplementary Table 1). The other 15 Japanese cultivars had the wild-type allele of *Sd1*.

Table 2 Number of genes with detected polymorphisms between rice cultivars

	Nipponbare	Akihikari	Hatsuboshi	Hatsunishiki	Hinohikari	Hitomebore	Itadaki	Kihou	Kirara 397	Koganebare	Koshihikari	Nourin 1	Nourin 6	Nourin 8	Nourin 22	Rikuu 132	Sasanishiki	Senshou	Basmati 370	IR36
Akihikari	63																			
Hatsuboshi	52	39																		
Hatsunishiki	56	45	40																	
Hinohikari	40	45	25	46																
Hitomebore	49	36	7	35	22															
Itadaki	56	43	29	37	38	28														
Kihou	52	54	36	66	35	43	52													
Kirara 397	74	73	61	69	64	56	54	75												
Koganebare	21	53	45	61	27	44	51	34	70											
Koshihikari	51	41	17	29	26	10	28	51	54	52										
Nourin 1	65	44	35	27	44	28	36	61	56	66	22									
Nourin 6	47	68	54	53	47	49	61	47	77	52	49	61								
Nourin 8	28	52	48	44	40	41	51	55	74	40	41	55	53							
Nourin 22	35	52	43	35	37	38	54	53	79	42	38	58	38	21						
Rikuu 132	62	45	42	45	51	35	44	69	49	64	39	33	65	51	58					
Sasanishiki	49	39	37	31	40	32	43	63	73	54	35	41	57	41	36	44				
Senshou	306	317	312	316	311	310	309	315	308	309	309	308	309	310	311	305	315			
Basmati 370	595	595	595	600	595	596	595	596	596	594	599	596	593	600	598	594	595	592		
IR36	748	746	744	744	745	744	743	747	746	749	744	746	741	747	743	745	747	706	625	
Kasalath	789	791	789	792	787	789	788	790	787	789	788	788	788	791	790	789	791	747	647	587

Number shows the genes with detected polymorphisms among 1,039 tested genes.

Identification of SNPs in genes

There were 134 genes showing DNA polymorphism, distributed evenly on the 12 chromosomes, between the Japanese cultivars (Fig. 1). Among the Japanese cultivars, two and three alleles were observed in 124 and 9 genes, respectively (Supplementary Table 1). Of the 134 genes shown to have DNA polymorphism by PCR–RF–SSCP analysis, the nucleotide sequences of 106 genes were determined and compared with those of 'Nipponbare', which were also determined in this study. Moreover, we determined the nucleotide sequences of the 31 genes, the DNA polymorphism of which had been detected in our previous study (Shirasawa et al. 2004). In 137 genes, 638 DNA polymorphisms were identified (Supplementary Table 2). The positions of the SNPs and indels in genes were assumed by using the published data of full-length cDNA sequences (The Rice Full-Length cDNA Consortium 2003: http://cdna01.dna.affrc.go.jp/cDNA/). There were 164 and 358 SNPs in exons and introns, respectively. Other regions including putative promoter, 5′ UTR, and downstream regions of genes had 35 DNA polymorphisms. The other 81 polymorphisms were not assigned because of the lack of published full-length cDNA sequences.

Out of 52 alleles of the genes having polymorphisms in the exons, one allele had a 1-bp deletion in the exon causing frame-shift mutation, three had a mutation causing amino acid insertions or deletions (indels), and 16 had missense polymorphisms. Thirty-two alleles had synonymous polymorphisms and/or polymorphisms in 3′ UTR, and 49 showed polymorphisms in introns, suggesting that they are probably silent polymorphisms. Proteins (The Rice Full-Length cDNA Consortium 2003) encoded by the cDNA sequences homologous to the 20 genes having missense polymorphisms are listed in Table 3. Fourteen genes showed homology to the reported genes.

A frame-shift mutation was found in R2194, which corresponds to a full-length cDNA clone J013110M15, encoding a putative integral membrane protein DUF6 having nine transmembrane helices (Table 3). The truncated protein encoded by the B allele having a frame-shift mutation had only three transmembrane helices. Polymorphisms of three, one, and five amino acid deletions were detected in B alleles of C0668, E10780, and C0556, respectively. C0556 encodes putative ribosomal protein L19. Missense polymorphisms were detected in 16 genes, in which the number of substituted amino acids ranged from 1 to 11.

Discussion

Comparison of SNP detection ability between PCR–RF–SSCP and mismatch–cleavage analysis

SNP detection ability of mismatch–cleavage analysis was lower than that of PCR–RF–SSCP. Mismatch–cleavage analysis using CEL I endonuclease has been found to detect polymorphisms in 17 of 23 ESTs of tomato (Yang et al. 2004). Although a modified EcoTILLING, in which DNA fragments were detected on agarose gel, has been reported

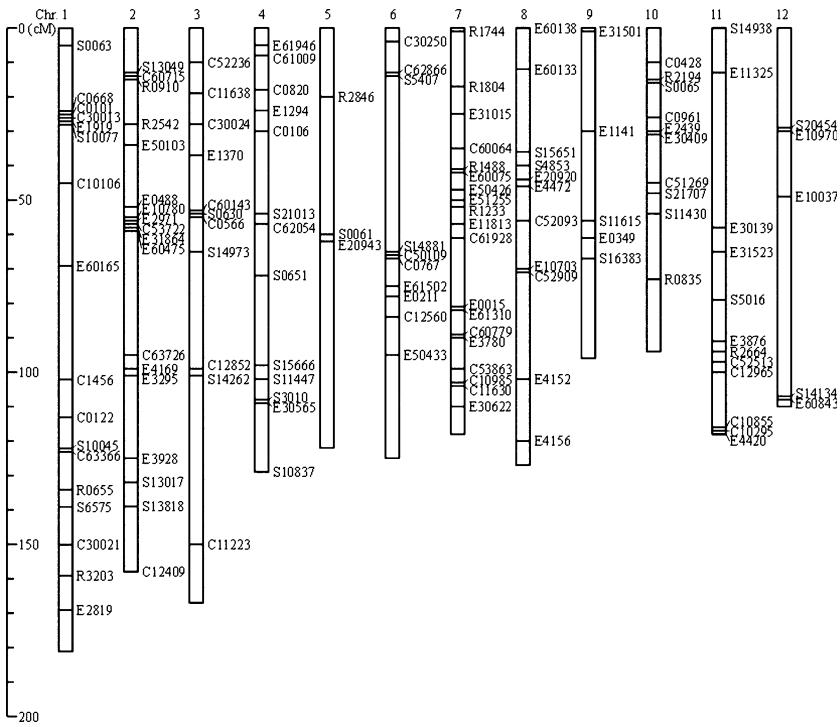

Fig. 1 Map positions of genes having polymorphisms between the 17 Japanese cultivars on the 12 chromosomes

Table 3 Differences of deduced amino-acid sequences between alleles

Maker name	Chr.	cM	Genes	Polymorphism
C0668	1	25.4	Hypothetical protein P0710E05.10	SSS363deletion
C30021	1	150.7	Hypothetical protein P0408G07.11	S455G, H481N
C62963	1	147.2–147.5	Membrane protein CH1-like	S127G, R515S
R3203	1	159.6	Putative EXO70-G1 protein	K349E
E10780	2	57.3–57.6	Hypothetical protein P0677G01.24	256Adeletion
E4169	2	99.6–101.2	Glycolipid transfer protein-like	A47S
C0566	3	55.8–56.3	Putative ribosomal protein L19	189AAPAAdeletion
E61946	4	5.4–6.5	Hypothetical protein OSJNBb0050O03.15	D475N, V563I
E1294	4	24.1–23.8	Facilitate chromatin transcription complex subunit SPT16	T656A
C0226	6	13.5	RRM-containing RNA-binding protein-like	D308E, M342T, Q412P, T451P, A464G, A466S, T467M, L468Q, H469Q, V473A, P476Q
C30223	6	67.4–67.7	Loricrin-like protein	C299F, T398I, A411T
R1744	7	0.8	Putative peroxidase 1	T227S, V358M
R565	7	24.2	Hypothetical protein OJ1714_H10.110	W73R, S89N
E51255	7	49.7	Granule binding starch synthase II	L523S
R3089	7	61.9	Nucleoside diphosphate kinase I	Y31F
C60779	7	89.8–91.7	Hypothetical protein OSJNBb34A20.6	A162S
E0349	9	60.8–62.4	Plastid (p)ppGpp synthase	L145F
R2194	10	15.7	Putative integral membrane protein DUF6	frame-shift
E2439	10	30.2	Putative calmodulin-binding family protein	R611P
S21168	12	108.2–109.2	Putative amine oxidase	V159I, L289F

to be useful for screening of SNPs in two genes (Raghavan et al. 2007), the efficiency of the SNP detections has not been shown. Results of the present study enable estimation of the efficiency of SNP detection by mismatch–cleavage analysis. In 92 SNPs (53 + 39) detected by PCR–RF–SSCP, 53 SNPs were detected by mismatch–cleavage analysis, indicating that the SNP-detection efficiency of mismatch–cleavage analysis was 57.6%. This low efficiency may be due to the presence of repetitive sequences or multiple SNPs in a single DNA fragment, which may be cleaved into DNA fragments being too short to be detected on an agarose gel.

Efficiency of SNP detection by PCR–RF–SSCP in the present study, in which DNA amplified by PCR was digested separately by *Afa*I and *Alu*I and mixed before electrophoresis, can be estimated to be 79.1% (= 53/(53 + 14)). Since the SNP detection ability of PCR–RF–SSCP of DNA fragments digested by a single restriction enzyme has been estimated to be about 70% (Sato and Nishio 2003), the efficiency of SNP detection by PCR–RF–SSCP using two different restriction enzymes is expected to be about 90%. Since the lengths of DNA fragments analyzed in the present study, ca. 4 kb in total, was more than twice that in the previous study by Sato and Nishio (2003), 1.4 kb on average, there is higher possibility of an overlap of bands, which makes it difficult to identify different bands, in the present study than in the previous one. The difference of efficiency of about 10% between the present study and the previous study may be attributable to the overlap of bands.

Estimation of SNP frequency in gene regions

We have previously detected polymorphisms of 671 genes by PCR–RF–SSCP using the same 21 rice cultivars as those used in the present study (Shirasawa et al. 2004). By combining the data of the previous and present studies, the average numbers of DNA fragments having polymorphism between the 17 Japanese cultivars, between the three *indica* cultivars, and between the *japonica* and *indica* cultivars were calculated to be 83.4 (4.9%), 1,022.4 (59.8%), and 1,164.0 (68.1%) in 1,710 genes, respectively. Since the efficiency of the detection of SNP by PCR–RF–SSCP was 79.1%, the polymorphism frequency in DNA fragments was estimated to be 6.2, 75.6, and 86.1% on average between the 17 Japanese cultivars, between the three *indica* cultivars, and between the *japonica* and *indica* cultivars, respectively. The average length of the transcribed region of rice genes has been estimated to be about 3 kb (Han and Xue 2003). In the present study, DNA fragments of ca. 2 kb in the downstream regions, including 3′-untranslated regions of the genes, were analyzed. Assuming that the frequency of the sequence variations in the latter half of the

transcribed region is not different from that in the former half, average frequencies of about 3-kb genes having polymorphisms between the 17 Japanese cultivars, between the three *indica* cultivars, and between the *japonica* and *indica* cultivars can be calculated to be 9.2, 87.9, and 94.8%, respectively. Because the number of rice genes has been inferred to be about 32,000 (The Rice Annotation Project 2007), sequence variations would be found in 2,944, 28,128, and 30,336 genes between the 17 Japanese cultivars, between the three *indica* cultivars, and between the *japonica* and *indica* cultivars, respectively (Table 4).

The frequencies of DNA fragments having frame-shift mutations and missense polymorphisms were 1.4 and 23.0%, respectively (Supplementary Table 2). The numbers of genes having frame-shift mutations and missense polymorphisms between the 17 Japanese cultivars were estimated to be 41 (0.1%) and 677 (2.1%) on average, respectively (Table 4). In the closest combination, 'Hatsuboshi' and 'Hitomebore', and the farthest combination, 'Kirara 397' and 'Nourin 6', 93 (0.3%) and 1,210 (3.8%) genes, respectively, are considered to have frame-shift mutations or missense polymorphisms. On the other hand, the numbers of genes having frame-shift mutations or missense polymorphisms between the three *indica* cultivars and between the *japonica* and *indica* cultivars were estimated to be 6,863 (21.4%) and 7,402 (23.1%), respectively.

Identification of chromosomal regions selected in rice breeding

'Hitomebore', the second most popular cultivar in Japan, has been developed from a cross between 'Koshihikari' and 'Hatsuboshi' (Fig. 2a, Sasaki et al. 1993). 'Hatsuboshi', which has a short culm and early heading traits, has been developed from a cross between 'Koshihikari' and 'Kihou' (Fig. 2a, Koumura et al. 1977). The excellent eating quality and cool-temperature tolerance at the booting stage of 'Hitomebore' are considered to be derived from 'Koshihikari'. On the other hand, due to 'Hatsuboshi' alleles, culm length has been shortened to prevent lodging, and earlier heading has been achieved as an adaptation to the climate of the Tohoku region (northeastern part) of Japan.

Genotypes of 'Koshihikari', 'Hatsuboshi', and 'Hitomebore' of 88 genes having polymorphisms between 'Koshihikari' and 'Kihou' are shown as graphical genotypes in Fig. 2b. Two, two, four, and six QTLs for cool-temperature tolerance at the booting stage, culm length, days-to-heading, and eating quality of 'Koshihikari', respectively, have been reported (Takeuchi et al. 2001; Wada et al. 2006). Cool-temperature tolerance at the booting stage of 'Hitomebore' and 'Hatsuboshi' may be due to two QTLs of 'Koshihikari' (*qCT-7* and *qCT-11*), both of which enhance cool-temperature tolerance at the booting

Table 4 The estimated number of genes having polymorphisms in the 32,000 rice genes between cultivars

Combinations	No. of 2-kb DNA fragments with detected polymorphisms in the 1,710 genes (*A*)	Estimated no. of genes having polymorphisms in the whole genome (*N*)[a]	Exon				Intron
				Frame-shift	Nonsense	Missense /deletion	Synonymous /3′ UTR	
Av. between Japanese cv.	83.4	2,944	1,592	41	<21	677	874	1,351
Hatsuboshi-Hitomebore	11	384	207	5	<3	88	114	176
Kirara 397-Nourin 6	144	4,960	2,683	69	<35	1,141	1,473	2,277
Av. between *indica* cv.	1,022.4	28,128	15,217	394	<197	6,469	8,354	12,911
Av. between Japanese cv. and *indica* cv.	1,164.0	30,336	16,412	425	<212	6,977	9,010	13,924
Nipponbare-Kasalath	1,305	31,776	17,190	445	<222	7,308	9,437	14,585

[a] The number of polymorphic genes in the genome (N) was calculated from the number of genes with detected polymorphism in the 1,710 genes (A) as follows

$B = (A/1{,}710)/0.791$; 0.791: efficiency of SNP detection by PCR–RF–SSCP. *B*: Frequency of polymorphic 2-kb fragments in the 1,710 genes

$N = 32{,}000(1\text{-}(1\text{-}B)^{3/2})$ Average size of the transcribed region of the genes is ca. 3 kb

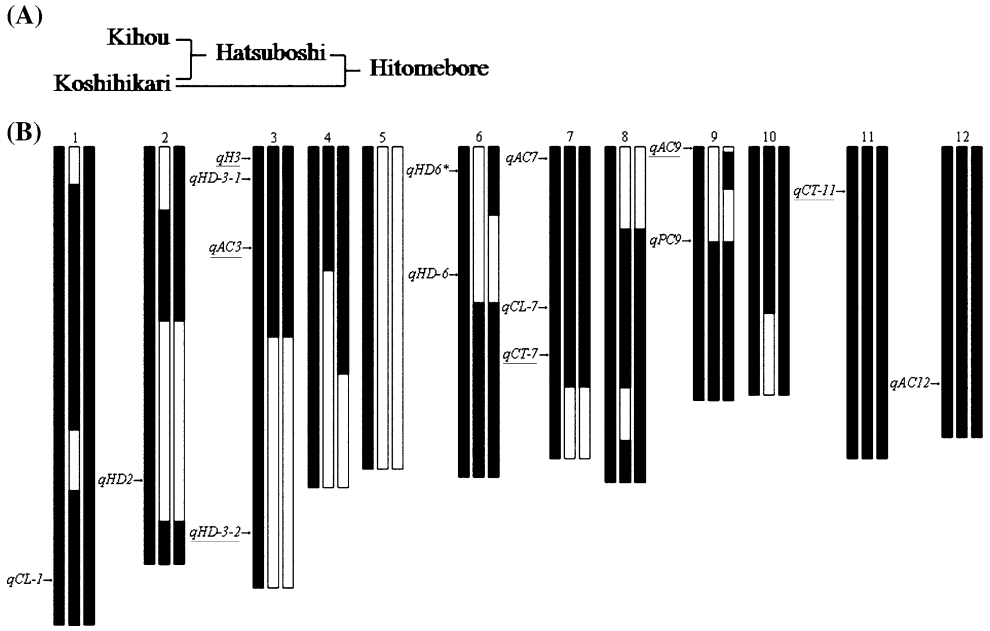

Fig. 2 Genealogy and graphical genotypes of Japanese cultivars. **a** Genealogy of four Japanese cultivars, 'Kihou', 'Koshihikari', 'Hatsuboshi', and 'Hitomebore'. **b** Graphical genotypings of 'Koshihikari' (*left bars*), 'Hatsuboshi' (*middle bars*), and 'Hitomebore' (*right bars*). *Black* and *white bars* show regions of chromosomes derived from 'Kishihikari' and 'Kihou', respectively. *Arrows* indicate positions and names of reported QTLs (Takeuchi et al. 2001; Wada et al. 2006). QTLs preferable for the development of a cultivar having cool-temperature tolerance at the booting stage, a short culm, an early heading date and excellent eating quality are underlined. *qHD6* (Wada et al. 2006) is written with an *asterisk* to distinguish it from *qHD-6* (Takeuchi et al. 2001)

stage. 'Hitomebore' and 'Hatsuboshi' have two QTLs for culm length of 'Koshihikari' (*qCL-1* and *qCL-7*), both of which elongate the culm. *qCL-1* is considered to be *sd1*, because it has been detected by using DHLs derived from a cross between 'Koshihikari' having the wild-type allele of *Sd1* and 'Akihikari' having 'Reimei'-type *sd1-r* (Supplementary Table 1). Since both 'Hitomebore' and 'Hatsuboshi' also have *Sd1*, the short culm length of these cultivars is considered to be controlled by 'Kihou' alleles in the other chromosomal regions. 'Hitomebore' has two QTLs for days-to-heading of 'Koshihikari' (*qHD-3-1* and *qHD6**; asterisk was tentatively used to distinguish it from *qHD-6*.), and 'Hatsuboshi' has one (*qHD-3-1*). However, both *qHD-3-1* and *qHD6** of 'Koshihikari' alleles have effects of late heading. *qHD-6* is considered to be *Hd1*, because it has been detected by using DHLs derived from a cross between 'Koshihikari' having a leaky allele of *Hd1* for late heading and 'Akihikari' having a null allele of *Hd1* for early heading (Yano et al. 2000, Supplementary Table 1). This null allele of *Hd1* derived from 'Kihou' is considered to contribute to early heading of 'Hatsuboshi' and 'Hitomebore'. Regarding the six QTLs for eating quality, both 'Hitomebore' and 'Hatsuboshi' are suggested to have the same combination of QTLs. Among the QTLs for eating quality, *qH3*, which has been reported to be closely related to the eating quality score of cooked rice, has been transferred to 'Hatsuboshi' and 'Hitomebore'.

As performed in the present study, chromosomal regions and genes selected in rice breeding processes can be identified by SNP analysis. Accumulation of information on SNPs in agronomically important genes will enable the prediction of the traits of individuals, lines, and cultivars without cultivation in the field.

Acknowledgments We are grateful to Ms. Y. Sugiyama for her technical assistance. This work was supported in part by Research project for utilizing advanced technologies in agriculture, forestry, and fisheries. K. Shirasawa is a recipient of a research fellowship from the Japan Society for the Promotion of Science for Young Scientists.

References

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309:741–745

Comai L, Young K, Till BJ, Reynolds SH, Greene EA et al. (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. Plant J 37:778–786

Feltus FA, Wan J, Schulze SR, Estill JC, Jiang N, Paterson AH (2004) An SNP resource for rice genetics and breeding based on subspecies *indica* and *japonica* genome alignments. Genome Res 14:1812–1819

Goff SA, Ricke D, Lan TH, Presting G, Wang R et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92–100

Han B, Xue Y (2003) Genome-wide intraspecific DNA-sequence variations in rice. Curr Opin Plant Biol 6:134–138

Hirokawa T, Boon-Chieng S, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378–379

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312:1392–1396

Koumura T, Takamatsu M, Haga T, Syumiya A, Syaku I, Taniguchi M, Tanabe K, Ito T (1977) A new rice variety "Hatsu-boshi". Res Bull Aichi Agric Res Cent A 9:1–11

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeted screening for induced mutations. Nat Biotechnol 18:455–457

Monna L, Ohta R, Masuda H, Koike A, Minobe Y (2006) Genome-wide searching of single-nucleotide polymorphisms among eight distantly and closely related rice cultivars (*Oryza sativa* L.) and a wild accession (*Oryza rufipogon* Griff.). DNA Res 13:43–51

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325

Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24:34–36

Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L, Minobe Y (2002) Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. DNA Res 9:163–171

Raghavan C, Naredo MEB, Wang H, Atienza G, Liu B, Qiu F, McNally KL, Leung H (2007) Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. Mol Breed 19:87–101

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

Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A et al. (2002) Green revolution: a mutant gibberellin-synthesis gene in rice. Nature 416:701–702

Sasaki T, Abe S, Matsunaga K, Okamoto E, Nagano K, Tanno K, Chiba Y, Kano A, Uematsu K (1993) A new rice cultivar "Hitomebore". Bull Miyagi Pref Furukawa Agric Exp Sta 2:1–17

Sato Y, Nishio T (2002) Efficient detection of DNA polymorphisms in cabbage and rice cultivars by PCR-RF-SSCP (PRS). Plant Cell Rep 21:276–281

Sato Y, Nishio T (2003) Mutation detection in rice waxy mutants by PCR-RF-SSCP. Theor Appl Genet 107:560–567

Sato Y, Shirasawa K, Takahashi Y, Nishimura M, Nishio T (2006) Mutant selection of progeny of gamma-ray-irradicated rice by DNA heteroduplex cleavage using *Brassica* petiole extract. Breed Sci 56:179–183

Shen YJ, Jiang H, Jin JP, Zhang ZB, Xi B et al. (2004) Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. Plant Physiol 135:1198–1205

Shirasawa K, Monna L, Kishitani S, Nishio T (2004) Single nucleotide polymorphisms in randomly selected genes among *japonica* rice (*Oryza sativa* L.) varieties identified by PCR-RF-SSCP. DNA Res 11:275–283

Song WY, Wang GL, Chen LL, Kim HS, Pi LY et al (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804–1806

Takeuchi Y, Hayasaka H, Chiba B, Tanaka I, Shimano T, Yamagishi M, Nagano K, Sasaki T, Yano M (2001) Mapping quantitative trait loci controlling cool-temperature tolerance at booting stage in temperate *japonica* rice. Breed Sci 51:191–197

The Rice Annotation Project (2007) Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana*. Genome Res 17:175–183

The Rice Full-Length cDNA Consortium (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from *japonica* rice. Science 301:376–379

Wada T, Uchimura Y, Ogata T, Tsubone M, Matsue Y (2006) Mapping of QTLs for physicochemical properties in *japonica* rice. Breed Sci 56:253–260

Wheelan SJ, Church DM, Ostell JM (2001) Spidey: a tool for mRNA-to-genomic alignments. Genome Res 11:1952–1957

Wu J, Maehara T, Shimokawa T, Yamamoto S, Harada C et al. (2002) A comprehensive rice transcript map containing 6591 expressed sequence tag sites. Plant Cell 14:525–535

Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ (2006) *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442:705–708

Yang W, Bai X, Kabelka E, Eaton C, Kamoun S, van der Knaap E, Francis D (2004) Discovery of single nucleotide polymorphisms in *Lycopersicon esculentum* by computer aided analysis of expressed sequence tags. Mol Breed 14:21–34

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L et al (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. Plant Cell 12:2473–2484

Yu J, Hu S, Wang J, Wong GK, Li S et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92

Zdobnov EM, Apweiler R (2001) InterProScan-an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847–848