

# Origin and evolution of twin microsatellites in the genus *Oryza*

HIROMORI AKAGI\*, YUMI YOKOZEKI, AKIKO INAGAKI & TATSUHITO FUJIMURA†

Biochemical Technology Section, Life Science Institute Laboratory, Performance Materials R & D Center, Mitsui Chemicals Inc., Togo 1144, Mobara 297, Japan

An ancestral sequence of twin microsatellites of rice was found in a wild species. Twin microsatellite loci, RM20A and RM20B, were located on separate regions of chromosomes 11 and 12, which had been duplicated during rice evolution. These twin microsatellites showed different allele diversities in A genome species of the genus *Oryza*. This difference was caused by repetition of a simple sequence consisting of (TAA)<sub>n</sub>. *Oryza longistaminata* contains a short poly(A) sequence in this region instead of the poly(TAA) found in other species. A sequence comparison of RM20-related amplicons suggested that the poly(A)-containing sequence is the ancestral sequence of the RM20A and RM20B microsatellites. A simple base substitution in the poly(A) sequence may have produced the longer microsatellite motif (TAA). This mutation may have occurred on one of the chromosomes of a hypothetical ancestor of the A genome species before duplication of the chromosome segments.

**Keywords:** microsatellites, *Oryza*, rice.

## Introduction

Microsatellites consist of tandemly arrayed di-, tri- and tetra-nucleotide repeats. They are hypervariable and are distributed randomly throughout eukaryotic genomes. The large allelic variation in microsatellite loci can be easily detected by PCR using specific primers in both of the flanking regions of microsatellites as a variation in the length of amplicons (Litt & Luty, 1989; Tautz, 1989; Weber & May, 1989). These characteristics of microsatellites make them useful for the construction of high-density genetic maps (Dib *et al.*, 1996; Dietrich *et al.*, 1996) and for evolutionary genetic studies (Bowcock *et al.*, 1994). Microsatellite DNA markers have also been identified in several plant species as aids in breeding (Akkaya *et al.*, 1992; Senior & Heun, 1993; Wu & Tanksley, 1993; Becker & Heun, 1995; Röder *et al.*, 1995; Broun & Tanksley, 1996).

The mechanisms which cause the variation in the length of microsatellites are not well understood. Polymerase slippage has been proposed as a mechanism of microsatellite mutation (Caskey *et al.*, 1992). Furthermore, knowledge regarding the evolutionary origins of microsatellites is limited. It has

been shown that a nucleotide substitution creates di- and tetra-nucleotide repeats in primates (Messier *et al.*, 1996). Such mutations may have occurred during the evolution of a given species, resulting in the present variation in microsatellite loci.

The genus *Oryza* is divided into four species complexes and two discrete species. The *O. sativa* complex contains all of the A genome species, including the two cultivated species *O. sativa* and *O. glaberrima*. *Oryza sativa* has been classified into two subspecies, *indica* and *japonica*, based on characteristic morphological and physiological traits (Oka, 1991). The evolutionary relationship among species in the genus *Oryza* has been well analysed by chloroplast DNA (Dally & Second, 1990; Chen *et al.*, 1993) and nuclear RFLP (Wang *et al.*, 1992).

Several microsatellite markers have been identified in *O. sativa* (Wu & Tanksley, 1993; Akagi *et al.*, 1996; Panaud *et al.*, 1996; Akagi *et al.*, 1997). Two pairs of microsatellite markers have been found on duplicate chromosomal segments of chromosomes 11 and 12 (Panaud *et al.*, 1996). Because these microsatellites may have originated from a single ancestral molecule in the evolution of the genus *Oryza*, they could be useful for understanding the origin and evolution of microsatellites.

In this study, allelic variations were analysed in the A genome species of the genus *Oryza* of RM20

\*Correspondence. E-mail: hiromori.akagi@mitsui-chem.co.jp

†Present address: Institute of Agricultural and Forest Engineering, University of Tsukuba, Tennoudai 1-1-1, Tsukuba, Ibaraki 305, Japan.

loci which are located on duplicate chromosomal segments. An ancestral sequence of the microsatellite was found in a wild species by analysing nucleotide sequences of RM20-related amplicons. The origin and evolution of these microsatellites is described here.

## Materials and methods

### *Plant materials*

A total of 61 accessions of cultivated rice species *O. sativa* (AA) including 59 local varieties and two modern cultivars were used. The varieties of *O. sativa* were classified into *indica* (30 lines) and *japonica* (29 lines) subspecies (Morishima & Oka, 1981). The two modern cultivars were Asominori and IR24, which are classified as *japonica* and *indica*, respectively. The 59 local varieties are stocked at the National Institute of Genetics and were kindly provided by Dr H. Morishima. One line of another cultivated species [*O. glaberrima* (A<sup>g</sup>A<sup>g</sup>)] was used. Twelve lines of wild rice, including seven lines of *O. rufipogon* (AA), one line of *O. brathii* (A<sup>g</sup>A<sup>g</sup>), two lines of *O. longistaminata* (A<sup>1</sup>A<sup>1</sup>), one line of *O. brachyantha* (FF) and one line of *O. meyeriana* (Diploid) were also used. These lines of wild species are stocked at Tohoku University, and were kindly provided by Dr M. Oka. The lines of *O. glaberrima*, *O. brathii*, *O. longistaminata* and *O. brachyantha* were collected in Africa, and those of *O. sativa* and *O. rufipogon* were collected in Asia. The accession numbers of the original collections were used in this study.

### *DNA extraction*

Crude DNA was extracted from the leaves of young seedlings. After being dried at 70°C for 2 h, leaves were homogenized with small glass beads in Eppendorf tubes using a vortex mixer. Crude DNA was then extracted according to the method of Edwards *et al.* (1991). Extraction buffer (400 µL) was added to the crushed seeds and the mixture was incubated at 100°C for 10 min. DNA was precipitated from the supernatant by adding an equal volume of isopropanol.

### *Southern blots*

Total DNA, which had been prepared from rice leaves by the CTAB method, was digested with a series of restriction enzymes, then separated on a 0.8% agarose gel. After transfer onto a nylon membrane (Hybond N+, Amersham International),

positive bands were detected using a ECL direct nucleic acid labelling and detection system according to the manufacturer's protocol.

### *Microsatellites*

Microsatellites of RM20 which were located on duplicate regions of chromosomes 11 and 12 were analysed (Panaud *et al.*, 1996). The nucleotide sequences of the primer pairs were 5'-GAAACA-GAGGCACATTTTCATTG-3' and 5'-ATCTTGTC-CCTGCAGGTCAT-3' (Panaud *et al.*, 1996). PCR amplification was performed in 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.5 or 1 unit of Taq polymerase (TAKARA), 4 nmol dNTP, 10 pmol primer, and 10 ng of genomic DNA per 20 µL using a Thermal Cycler 9600 (Perkin-Elmer). Thirty-five PCR cycles, each consisting of 30 s denaturation at 94°C, 30 s annealing at 55°C, and 1 min polymerization at 72°C, were performed. For analysis with GeneScan (ABI), 0.5 µL of fluorescent dUTP [2 µL for (TAMRA)dUTP] was added to the PCR reaction mixtures. After PCR, unincorporated fluorescent dUTPs were removed with SUPRECTM-02 (TAKARA).

### *Nucleotide sequence and diversity of amplified DNA fragments*

Amplified DNA fragments were subcloned into pCR<sup>TM</sup>II using a TA cloning kit (Invitrogen) and sequenced using an ABI 373S (Applied Biosystems Inc.). DNA sequences were analysed using GENETYX-MAC software (Software Development, Tokyo).

Lengths of alleles were determined by ethidium bromide staining after electrophoresis on 3% MetaPhor Agarose gels (FMC). Exact allele lengths were also determined by GeneScan (ABI) after fractionation in 6% denatured polyacrylamide gels using a 373A DNA sequencer (ABI).

Gene diversity (*GD*) was calculated as follows:

$$GD = 1 - \sum_{i=1}^m x_i^2,$$

where  $x_i$  indicates the population frequency for a marker and the summation extends over  $m$  patterns (Nei, 1973).

## Results

### *Diversity of twin microsatellites of RM20 loci in A genome species of the genus Oryza*

RM20A and RM20B are microsatellites that are located on the highly conserved duplicated regions

of chromosomes 11 and 12, respectively (Nagamura *et al.*, 1995; Panaud *et al.*, 1996). Because both microsatellites are amplified by a single primer set (Panaud *et al.*, 1996) the RM20A and RM20B loci could be examined in 73 lines belonging to the A genome species *O. sativa*, *O. glaberrima*, *O. rufipogon*, *O. brathii* and *O. longistaminata*, and to the F genome species *O. brachyantha*. Two different lengths of amplicons were amplified by a single primer pair in all but three of the lines. The longer amplicon corresponded to RM20A and was mapped on chromosome 12, whereas the shorter one was mapped on chromosome 11 (Akagi *et al.*, 1996; Panaud *et al.*, 1996). One line of *O. longistaminata* and one line of *O. brachyantha* gave a single amplicon and no amplicon was observed in *O. meyeriana* (Fig. 1). Amplification at some microsatellite loci also failed in several wild rice species tested by Panaud *et al.* (1996), presumably because the primer sequences were not conserved.

Overall, 19 and nine alleles (including *O. rufipogon*, *O. sativa* ssp. *indica* and ssp. *japonica*) were detected at the RM20A and RM20B loci, respectively (Table 1). In both *indica* and *japonica* subspecies, twice as many alleles were found at the RM20A locus as at the RM20B locus (Table 1). A similar difference was not observed in *O. rufipogon* but this was based on a smaller sample (Table 1).

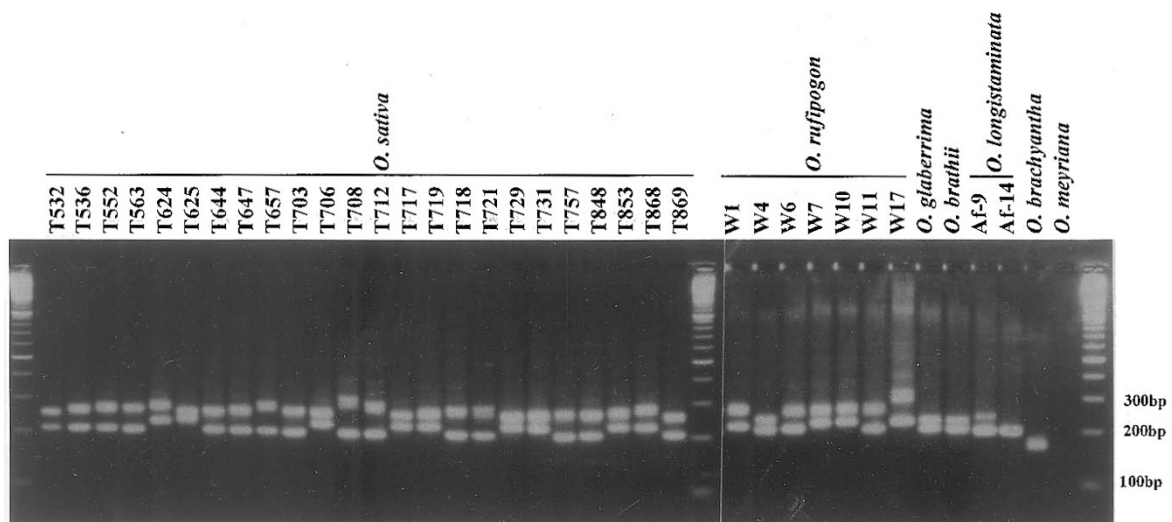
Allele lengths at the RM20A locus had a similar distribution in *japonica* and *indica* accessions (Fig. 2). In contrast, the *japonica* and *indica* subspecies showed different frequencies of allele length at

the RM20B locus (Fig. 2). Two species from Africa, *O. glaberrima* and *O. brathii*, had the shortest allele length (Fig. 2). The varieties were classified into *indica* and *japonica* subspecies according to an isozyme analysis (Morishima & Oka, 1981). Introgression of the genes between these two subspecies may indicate that *japonica* and *indica* had longer and shorter alleles at the RM20B locus, respectively (Second, 1982).

*Highly conserved nucleotide sequences between the twin microsatellites*

The nucleotide sequences of RM20A and RM20B in several lines were determined to identify the molecular basis of the difference in allelic diversity at the RM20A and RM20B loci. Amplicons from all of the lines, except *O. longistaminata* and *O. brachyantha*, contained a simple sequence trinucleotide repeat, (TAA)<sub>n</sub> (Fig. 3a). The nucleotide sequences of the flanking regions were highly conserved between RM20A and RM20B (Fig. 4). Characteristic sequences for each locus were found at the 5'-flanking region of the microsatellite motif; (T)<sub>6</sub>GCAA-(TAA)<sub>n</sub> for RM20A and (T)<sub>3</sub>A(TAA)<sub>n</sub> for RM20B (Figs 3 & 4).

The numbers of simple sequence repeats at the RM20A and RM20B loci ranged from 13–36 and 7–15, respectively (Fig. 3a, b). This indicated that the allele length at both of the RM20 loci in Fig. 2 was determined solely by the number of TAA repeats. Another microsatellite motif, (CT)<sub>n</sub>, in



**Fig. 1** Discrimination of seven alleles detected in RM20 among 59 Japanese cultivars of rice. The PCR products of RM20 were electrophoresed on 3% MetaPhor agarose gels (FMC) and then stained with ethidium bromide. DNA size-marker lanes contain 100 bp per 500 bp DNA ladders.

**Table 1** Gene diversity for twin microsatellite loci in *Oryza rufipogon* and two subspecies of *O. sativa*

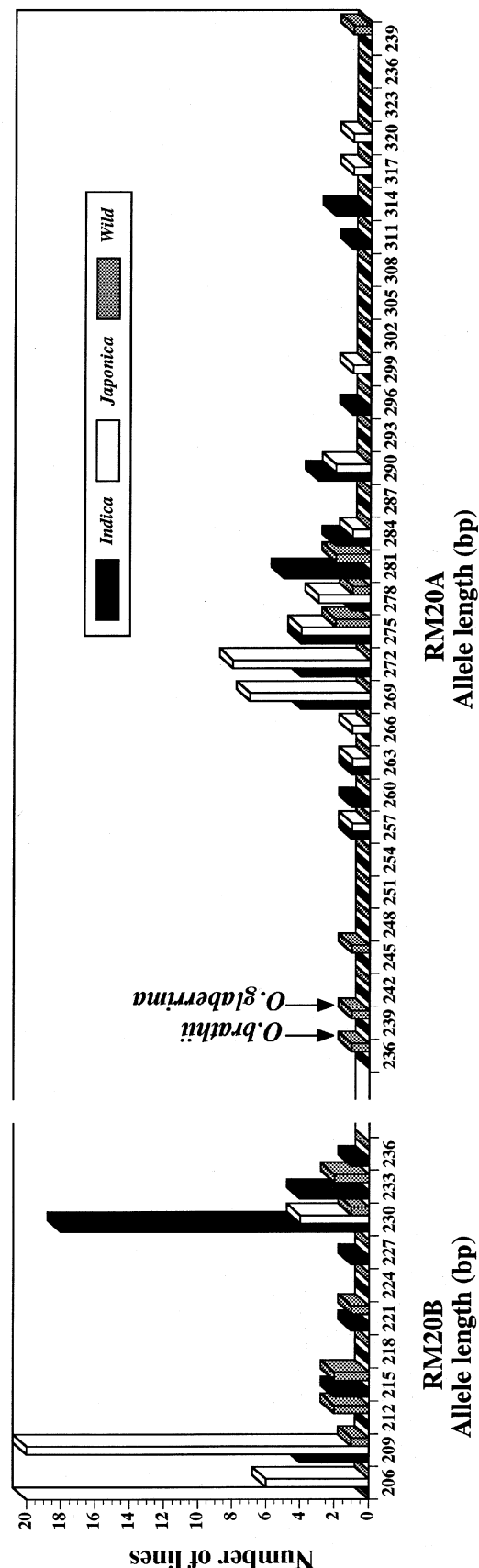
Locus	<i>indica</i>	<i>japonica</i>	<i>O. rufipogon</i>	Total
RM20A	0.890 (13)	0.845 (12)	0.776 (5)	0.893 (19)
RM20B	0.622 (7)	0.500 (3)	0.776 (5)	0.743 (9)

Numbers in parentheses indicate number of alleles detected.

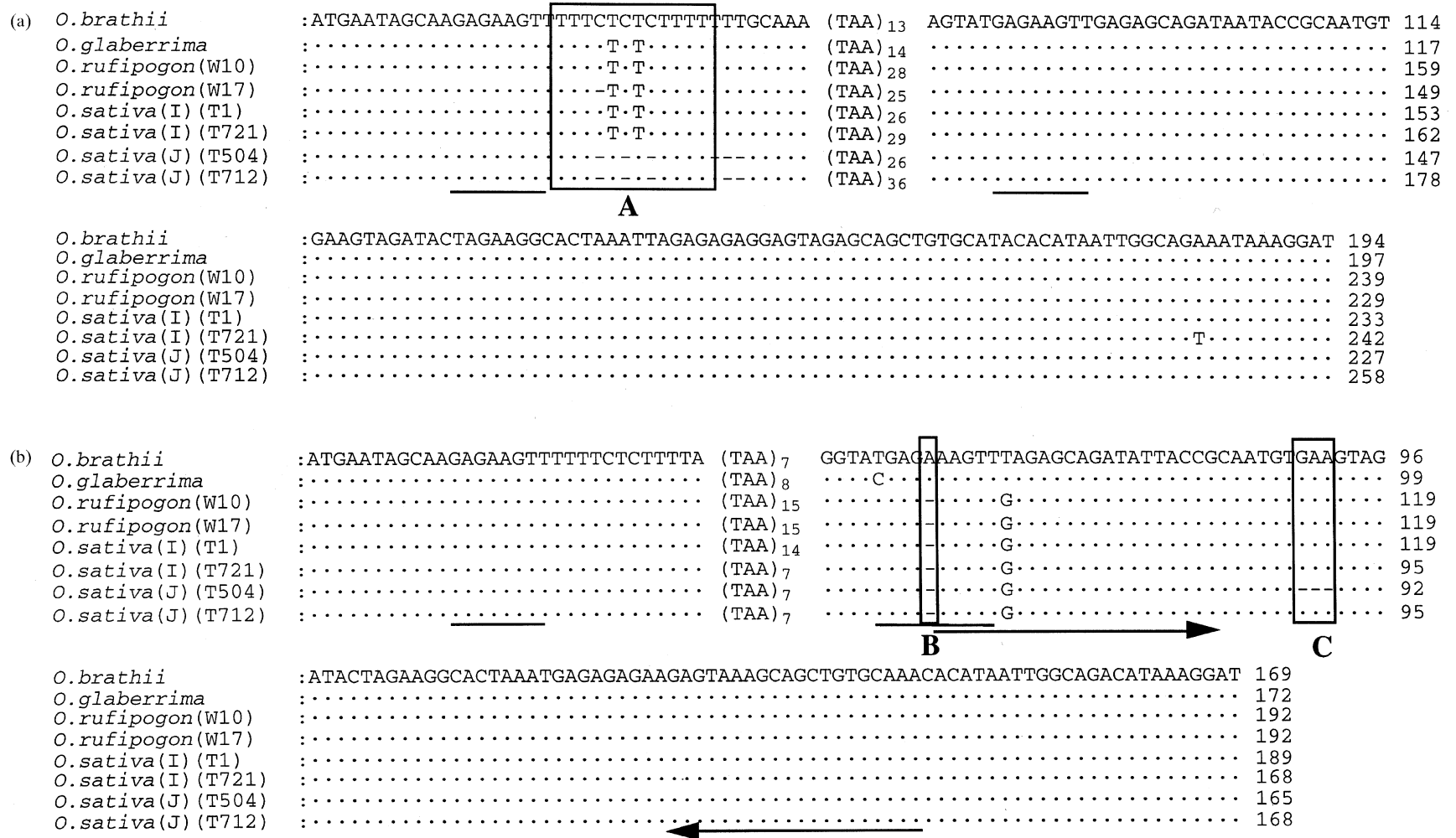
RM20A was not found in the *japonica* subspecies (Fig. 3a, box A). In some varieties of *japonica*, a three-nucleotide deletion was also found in the region 5' of RM20B microsatellite (Fig. 3b, box C). Distribution of this deletion among the A genome species was surveyed by PCR using the primers arrowed in Fig. 3(b). This deletion was found only in the *japonica* subspecies by PCR (data not shown), indicating that the deletion had occurred after the divergence of *japonica*. A single-base insertion at the RM20B locus was only found in *O. glaberrima* and *O. brathii* (Fig. 3b, box B). This agrees well with the notion that *O. brathii* is the ancestor of the cultivated species *O. glaberrima* (Morishima *et al.*, 1963; Second, 1982).

*The wild species O. longistaminata contains poly(A) instead of poly(TAA)*

Different amplification patterns with the primer pair for RM20 were observed in two lines of *O. longistaminata*. Two different sizes of amplicon were observed in line Af-9, whereas only an amplicon with a lower molecular weight was observed in line Af-14 (Fig. 1). The nucleotide sequences of these higher (RM20HI) and lower (RM20LI) molecular weight amplicons from line Af-9 were determined (Fig. 4). In the nucleotide sequences of RM20LI, a short stretch of poly(A) was found in place of poly(TAA) (Fig. 4), whereas RM20HI contained a poly(TAA) sequence (Fig. 4). However, the nucleotide sequences on both sides of this microsatellite motif were diverged from those of either RM20A or RM20B. In the 3'-flanking region of the poly(TAA) of RM20HI, the sequence GTATGAGAA was deleted (Fig. 4, underline). The twin microsatellites RM20A and RM20B can be distinguished from each other by the number of repeats of the microsatellite motif and by variation in their flanking sequences. However, it was not possible to conclude that either RM20LI or RM20HI corresponded to RM20A or RM20B based on the features of their nucleotide sequences. RM20A and RM20B were previously mapped on rice chromosomes using a cross between



**Fig. 2** Distribution of allele lengths among *Oryza sativa*, *O. rufipogon*, *O. glaberrima* and *O. brathii* at the RM20B locus and the RM20A locus. *Oryza sativa* was classified into *japonica* and *indica* subspecies based on an isozyme analysis (Morishima & Oka, 1981).



**Fig. 3** Comparison of the nucleotide sequences of the microsatellites in the (a) RM20A and (b) RM20B loci, amplified from several lines of *Oryza sativa*, *O. rufipogon*, *O. glaberrima* and *O. brathii*. *Oryza sativa*(I) and *O. sativa*(J) indicate *indica* and *japonica* subspecies, respectively. Accession numbers of the lines are also indicated. Conserved nucleotides among these sequences are shown by dots. Underlining indicates the 8-bp repetitive sequences flanked by the microsatellite. Arrows indicate primer sequences for detection of a three-base deletion at the RM20B locus. Boxes A, B and C show special features of nucleotide sequences for *japonica* subspecies, the African species *O. glaberrima* and *O. brathii*, and some *japonica* subspecies, respectively.



*Origin and evolution of the twin microsatellites in the genus Oryza*

Microsatellites vary in length because of slippage amplification or reduction. However, there is little information available on the origins of microsatellites. A single-base substitution created tetranucleotide repeats and a dinucleotide repeat in the  $\eta$ -globin pseudogene in primates (Messier *et al.*, 1996). In this study, a putative ancestral sequence of the twin microsatellites was found. Poly(TAA) may have been created from poly(A) by a single-base substitution (A to T), as shown in Fig. 7, which describes the simplest routes from a hypothetical ancestral molecule to the present-day molecules.

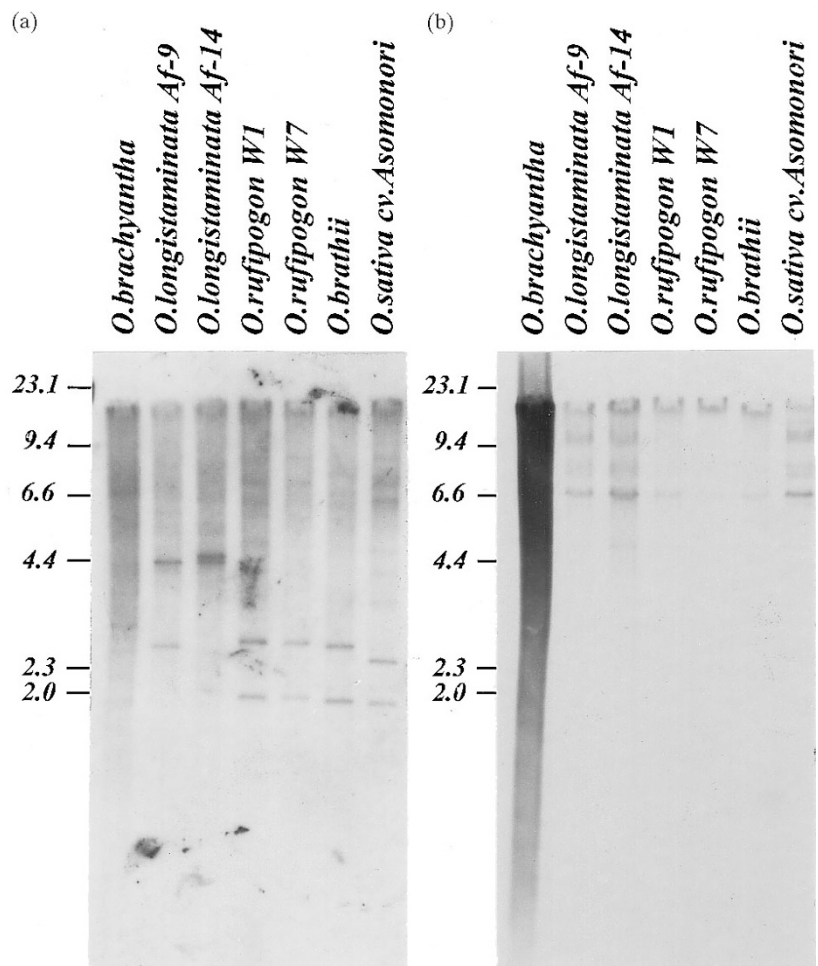
*Relationship between the evolution of the genus Oryza and twin microsatellites*

The genus *Oryza* is divided into four species complexes. All of the A genome species analysed

here belong to the *O. sativa* complex. The evolutionary relationships among species of the genus *Oryza* have been evaluated by morphological (Morishima & Oka, 1960), isozyme (Second, 1982), chloroplast DNA (Dally & Second, 1990) and nuclear RFLP (Wang *et al.*, 1992) analyses.

The lines analysed in the present study had been collected in Africa (*O. longistaminata*, *O. brathii*, *O. glaberrima* and *O. brachyantha*) and Asia (*O. sativa* and *O. rufipogon*). The nucleotide sequences of RM20 molecules were highly conserved between African and Asian species. The features of these nucleotide sequences suggested that *O. brathii* had a closer affinity with *O. rufipogon* than with *O. longistaminata*, despite the geographical isolation of these wild species.

Figure 8 shows a dendrogram which was constructed based on a nuclear RFLP analysis (Wang *et al.*, 1992). Steps in the evolution of RM20-related molecules are indicated on this



**Fig. 5** Southern blot analysis of *Pst*I fragments. Panel (a) shows the hybridization patterns with the RM20A probe from *Oryza sativa* cv. Asononori. Panel (b) shows the hybridization patterns with the RM20b probe from *O. brachyantha*.

```

O. brachyantha (RM20b)          1: AT-AA-A-CAAGAGAAATTT-----GAGAGCAGAGGTTGCCGGAACCTGAAA 43
O. longistaminata (RM20LI)    1: ..G..T.GT.....G..TTTTCTCTCTTTTTCGCAAAAAAAGTATGAGAAGTT-----TA..A..C..TG.....G 86

O. brachyantha (RM20b)          44: TACATACCAAAAGGCACCTCAATTTACAGAGAGTAGAGCAG-TGGTGCATGTATATAATTGG-TGCAAAAATAAAGGAT 127
O. longistaminata (RM20LI)    87: ..G....T.G.....A.....G.....C.....C.C.....GA.T.G..... 164

```

**Fig. 6** Homology between nucleotide sequences of RM20b from *Oryza brachyantha* and RM20LI from *O. longistaminata*. Repetitive sequences of 8 bp are underlined. Conserved nucleotides among these sequences are shown by dots.

dendrogram. Mutation at the original molecule resulted in the intermediate molecule 'X'. Before or after the formation of the intermediate 'Y', the chromosomal segment that contained the source of the microsatellites may have duplicated, resulting in the present structures of chromosomes 11 and 12 in the A genome species. Because RM20HI and RM20LI in *O. longistaminata*, and RM20A and RM20B in other species may have evolved directly from the intermediate molecule 'Y' independently, these RM20 molecules may have formed after the divergence of *O. longistaminata* and other A genome species. This result agrees well with the previous finding that *O. longistaminata* had initially diverged from other A genome species. RM20A(I) and RM20B(I) evolved before the divergence of *O. brathii* and *O. rufipogon* at chromosomes 12 and 11, respectively (Fig. 8). However, the time of the formation of RM20B(J) and RM20A(J) during rice evolution could not be verified by a sequence comparison. There was no evidence regarding whether *indica* or *japonica* subspecies developed first in the evolution of RM20A and RM20B. However, allele diversities did not differ between *indica* and *japonica* carrying RM20A(I) and RM20A(J), respectively (Fig. 2), suggesting that RM20A(J) may have been created close to RM20A(I). Therefore, RM20A(J) may also have been created before the divergence of *O. brathii* and *O. rufipogon* (Fig. 8). These results suggest that three types of cultivated rice, *O. glaberrima*, *O. sativa* ssp. *indica* and *O. sativa* ssp. *japonica*, may have evolved in parallel from individual ancestors, although ancestors of the *japonica* subspecies were not found here by nucleotide sequence analysis. *Oryza glaberrima* was domesticated directly from *O. brathii* in Africa independently of *O. sativa* in Asia (Morishima *et al.*, 1963; Second, 1982). Previous reports have also suggested that the *japonica* and *indica* subspecies differentiated before their domestication (Second, 1982; Wang *et al.*, 1992; Chen *et al.*, 1993). Because RM20B(I) may have been formed from RM20B(J) by the duplication of poly(TAA), RM20B(J) may have been formed before the divergence of *O. brathii* and *O. rufipogon*. This suggests that an ancestor of *japonica* subspecies had already differentiated from an ancestor of *indica* and *O. glaberrima* before the divergence of *O. brathii*.

### Acknowledgements

We thank Dr H. Morishima (National Institute of Genetics) and Dr M. Oka (Miyagi Teachers' College) for kindly providing the rice strains.



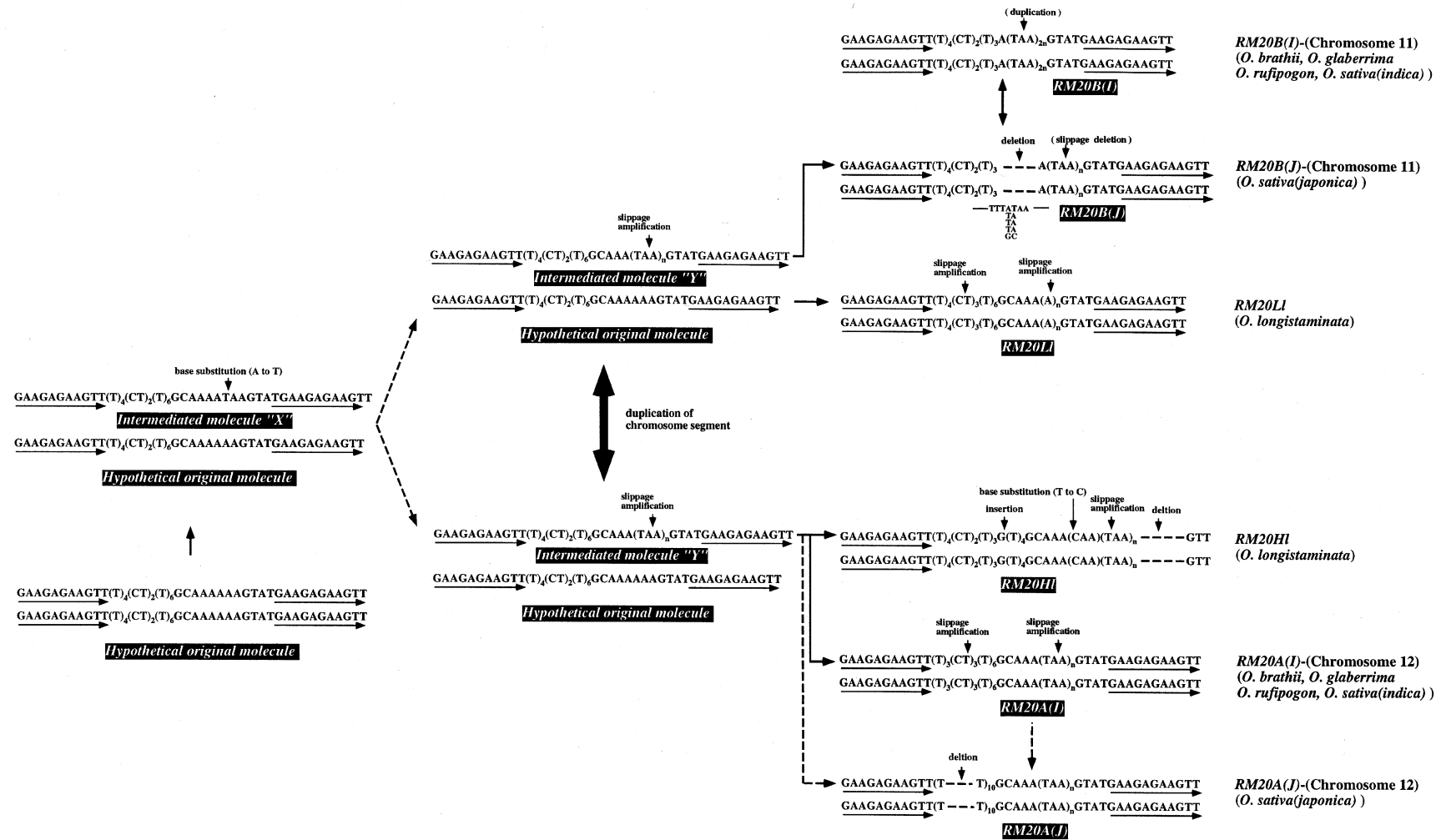
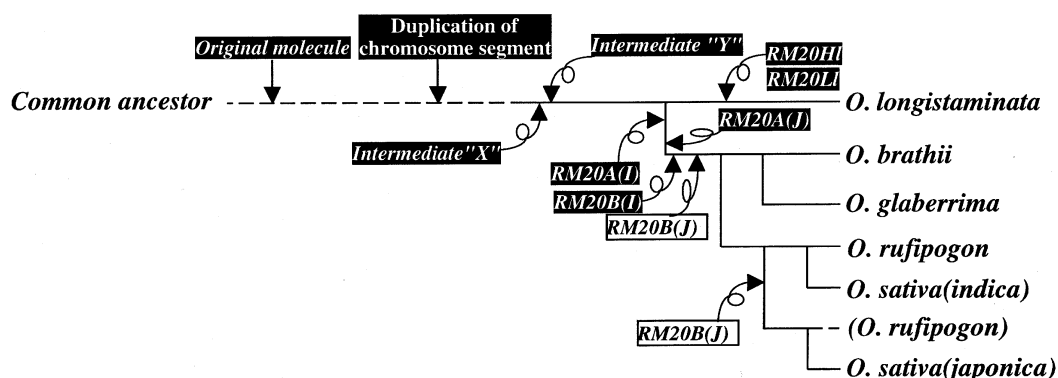


Fig. 7 Evolutionary relationships among RM20-related molecules. The simplest routes from a hypothetical ancestral sequence to the present RM20-related molecules are shown. Two possible routes are shown for the formation of RM20A(I) and RM20A(J).



**Fig. 8** Time of the formation of RM20-related molecules during the evolution of the genus *Oryza*. The dendrogram is constructed based on an RFLP analysis (Wang *et al.*, 1992). The timing of the formation of RM20B(J) could not be determined here.

## References

- AKAGI, H., YOKOZEKI, Y., INAGAKI, A. AND FUJIMURA, T. 1996. Microsatellite DNA markers for the rice chromosomes. *Theor. Appl. Genet.*, **93**, 1071–1077.
- AKAGI, H., YOKOZEKI, Y., INAGAKI, A. AND FUJIMURA, T. 1997. Highly polymorphic microsatellites of rice consist of AT repeats, and classification of closely related cultivars with these microsatellite loci. *Theor. Appl. Genet.*, **94**, 61–67.
- AKKAYA, M. S., BHAGWAT, A. A. AND CREGAN, P. B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*, **132**, 1131–1139.
- BECKER, J. AND HEUN, M. 1995. Barley microsatellites, allele variation and mapping. *Plant Mol. Biol.*, **27**, 835–845.
- BOWCOCK, A. M., RUIZ-LINARES, A., TOMFOHRDE, J., MINCH, E., KIDD, J. R. AND CAVALLI-SFORZA, L. L. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455–457.
- BROUN, P. AND TANKSLEY, S. D. 1996. Characterization and genetic mapping of simple repeat sequences in the tomato genome. *Mol. Gen. Genet.*, **250**, 39–49.
- CASKEY, C. T., PIZZUTI, A., HU, Y.-H., FENWICK, R. G. AND NELSON, D. L. 1992. Triplet repeat mutations in human disease. *Science*, **256**, 784–789.
- CHEN, W.-B., NAKAMURA, I., SATO, Y.-I. AND NAKAI, H. 1993. Distribution of deletion type in cpDNA of cultivated and wild rice. *Jap. J. Genet.*, **68**, 597–603.
- DALLY, A. M. AND SECOND, G. 1990. Chloroplast DNA diversity in wild and cultivated species of rice (Genus *Oryza*, section *Oryza*). Cladistic-mutation and genetic-distance analysis. *Theor. Appl. Genet.*, **80**, 209–222.
- DIB, C., FAURÉ, S., FIZAMES, C., SAMSON, D., DROUOT, N., VIGNAL, A. ET AL. 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*, **380**, 152–154.
- DIETRICH, W. F., MILLER, J., STEEN, R., MERCHANT, M. A., DAMRON-BOLES, D., HUSAIN, Z. ET AL. 1996. A comprehensive genetic map of the mouse genome. *Nature*, **380**, 149–152.
- EDWARDS, K., JOHNSTONE, C. AND THOMPSON, C. 1991. A simple and rapid method for the preparation of plant genome DNA for PCR analysis. *Nucl. Acids Res.*, **19**, 1349.
- LITT, M. AND LUTY, J. A. 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.*, **44**, 397–401.
- MESSIER, W., LI, S.-H. AND STEWART, C.-B. 1996. The birth of microsatellites. *Nature*, **381**, 483.
- MORISHIMA, H. AND OKA, H. I. 1960. The pattern of interspecific variation in the genus *Oryza*, its quantitative representation by statistical methods. *Evolution*, **14**, 153–165.
- MORISHIMA, H. AND OKA, H. I. 1981. Phylogenetic differentiation of cultivated rice, XXII. Numerical evaluation of the *indica-japonica* differentiation. *Jap. J. Breed.*, **31**, 402–413.
- MORISHIMA, H., HINATA, K. AND OKA, H. I. 1963. Comparison of modes of evolution of cultivated forms from two wild rice species, *Oryza brathii* and *O. perennis*. *Evolution*, **15**, 326–339.
- NAGAMURA, Y., INOUE, T., ANTONIO, B. A., SHIMANO, T., KAJIYA, H., SHOMURA, A. ET AL. 1995. Conservation of duplicated segments between rice chromosomes 11 and 12. *Breed. Sci.*, **45**, 373–376.
- NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 3321–3323.
- OKA, H. I. 1991. Genetic diversity of wild and cultivated rice. In: Khush, G.S. & Toenniessen, G.N (eds) *Rice Biotechnology*, pp. 55–81. CAB International, Wallingford.
- PANAUD, O., CHEN, X. AND MCCOUCH, S. R. 1996. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Mol. Gen. Genet.*, **252**, 597–607.
- RÖDER, M. S., PLASCHKE, J., KÖIG, S. U., BÖNER, A., SORRELLS, M. E., TANKSLEY, S. D. AND GANAL, M. W. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. *Mol. Gen. Genet.*, **246**, 327–333.

- RUBINSZTEIN, D. C., AMOS, W., LEGGO, J., GOODBURN, S., JAIN, S., LI, S.-H. *ET AL.* 1995. Microsatellite evolution - evidence for directionality and variation in rate between species. *Nature Genet.*, **10**, 337-343.
- SECOND, G. 1982. Origin of genetic diversity of cultivated rice (*Oryza* spp.), study of the polymorphism scored at 40 isozyme loci. *Jap. J. Genet.*, **57**, 25-57.
- SENIOR, M. L. AND HEUN, M. 1993. Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome*, **36**, 884-889.
- TAUTZ, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucl. Acids Res.*, **17**, 6463-6471.
- VALDES, A. M., SLATKIN, M. AND FREIMER, N. B. 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*, **133**, 737-749.
- WANG, Z. Y., SECOND, G. AND TANKSLEY, S. D. 1992. Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theor. Appl. Genet.*, **83**, 565-581.
- WEBER, J. L. AND MAY, P. E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.*, **44**, 388-396.
- WU, K.-S. AND TANKSLEY, S. D. 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. *Mol. Gen. Genet.*, **241**, 225-235.