

GENE 03881

## Rapid isolation of a rice *waxy* sequence: a simple PCR method for the analysis of recombinant plasmids from intact *Escherichia coli* cells

(Recombinant DNA; genomic DNA; homologous primer; amplification; cloning; sequence analysis; intron; antisense gene)

Hiroaki Shimada and Yuichi Tada \*

Mitsui Plant Biotechnology Research Institute, 2-1-6 Sengen, Tsukuba 305 (Japan)

Received by R. Wu: 20 August 1990

Revised: 24 September 1990

Accepted: 5 October 1990

---

### SUMMARY

Polymerase chain reaction (PCR) is an efficient method for obtaining a desired nt sequence if both required primers can hybridize to the target DNA molecule specifically. A rapid and simple PCR-based method for analyzing plasmids using intact cells was established. An attempt to target a rice *waxy* sequence by PCR using homologous primers was also carried out. In three cases, specific fragments were amplified and their nucleotide sequences were determined. However, the cloned rice *waxy* gene contained base substitution mutations. The cumulative frequency of mutation after 30 polymerization cycles was estimated to be one in 500 bp.

---

### INTRODUCTION

PCR is an in vitro method for the enzymatic synthesis of specific DNA sequences, using two oligo primers that can hybridize to opposite strands and flank the region of interest in the target DNA (Saiki et al., 1988). A repetitive series of cycles are carried out that involve template denaturation, primer annealing, and the synthesis of a specific DNA fragment whose termini are defined by the 5' end of the primers. Using this method, a particular fragment of DNA can be specifically amplified and then cloned.

The amylose content of endosperm starch strongly

affects the taste of rice. Many mutants involved in starch synthesis have been isolated (Yano et al., 1988), suggesting that a set of genes in combination controls starch synthesis. Among them, the *waxy* gene which encode a major starch synthesizing enzyme, UDP-glucose starch glycosyl transferase (EC 2.4.1.11), has been well investigated. Most *waxy* mutants exhibit varying levels of amylose and amylopectin in their endosperms, depending upon the allele (Sano et al., 1986). This observation suggests that the affected enzyme may control the amylose content of endosperm starch. Therefore, manipulation of *waxy* gene expression level may be a valuable tool for breeders to improve the quality of rice grains. In the case of the *waxy* gene, it may be possible to use antisense RNA to down-regulate gene expression and thereby reduce amylose content in the endosperm starch.

Maize and barley *waxy* genes were previously cloned and their nt sequences were determined (Klößgen et al., 1986; Rohde et al., 1988). It was suggested that the rice sequence might also be similar to the maize sequence (Okagaki and Wessler, 1988). The fact that the maize and barley sequences have highly conserved regions suggests that the same sequences should be conserved in rice. Using primers

---

Correspondence to: Dr. H. Shimada, Mitsui Plant Biotechnology Research Institute, 2-1-6 Sengen, Tsukuba 305 (Japan) Tel. (81)298-58-6115; Fax (81)298-58-6191.

\* Present address: Plant Biotechnology Laboratory, Mitsui Toatsu Chemicals, Inc., 1144 Togo, Mobara 297 (Japan) Tel. (81)475-25-6729.

Abbreviations: bp, base pair(s); kb, kilobase(s); nt, nucleotide(s); oligo, oligodeoxyribonucleotide; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; *waxy*, gene encoding UDP-glucose starch glycosyl transferase.

based upon the barley sequences within these highly conserved regions, it was expected that rice genomic sequences could be amplified by PCR.

In this paper, we report the rapid isolation of the rice *waxy* sequence from genomic DNA. In addition, we also describe a rapid, efficient PCR-based method for analyzing plasmids, that bypasses the necessity to grow up bacteria in liquid culture, to perform explicit lysis steps, or to carry out DNA extraction or precipitation steps.

## EXPERIMENTAL AND DISCUSSION

### (a) A rapid and simple method for the selection of recombinant plasmids

Various procedures for identifying specific DNA within intact cells all contain some steps for cell lysis and/or DNA extraction (e.g., Kim and Smithies, 1988). Therefore, a protocol omitting the DNA extraction step would shorten the procedure.

The denaturation step of PCR is normally carried out at a high temperature (90–94°C). Under such conditions, any

*E. coli* cells present in the solution should be lysed or made permeable, and would allow DNA to leak out, especially small multicopy plasmids. Thus, we tried to simplify the procedure for analyzing the DNA fragments inserted in the pUC plasmid vector. Commercially available forward and reverse sequencing primers (5'–GTTTTCCAGTCAC–GAC and 5'–CAGGAAACAGCTATGTGAC from Takara Shuzo Co., Japan) were used for the PCR amplification. Fresh transformant colonies grown on selective medium plates were picked using round toothpicks. Each toothpick was stabbed into a second agar plate prior to transfer into a tube containing 50 µl of reaction mixture. The small number of transformed cells was found not to inhibit subsequent steps of the PCR. Prior to DNA synthesis cycles, a 10-min incubation at 90°C was added to ensure sufficient DNA leakage from the cells. After PCR, 5 µl from each reaction was examined by agarose gel electrophoresis.

All reactions gave distinct bands on agarose gels, consistent with the presence of various inserts (Fig. 1). Based upon the ethidium bromide staining intensity, all reactions yielded approximately the same mass of DNA product. A 3.4-kb control insert was also precisely amplified, indicating that such large inserts present no insurmountable obstacle to amplification. Colonies more than three days old, how-

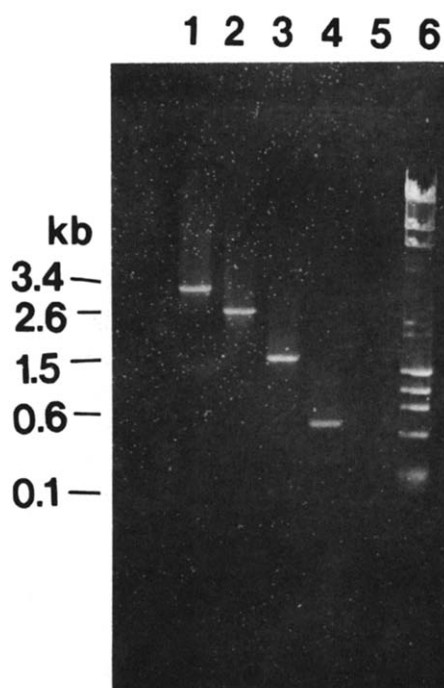


Fig. 1. Agarose gel analysis of PCR products starting from intact cells. Cells containing pSS3C (Sugita et al., 1987), pRP4-3, pRP4-8, PRP1-4 (Shimada et al., 1989) and pUC119 (Vieira and Messing, 1987) were used in lanes 1–5, respectively. Lane 6 contains *Hind*III/λ DNA and *Hae*III/φX174 DNA (size marker). Numbers to the left indicate sizes of insert DNA corresponding to each plasmid. The standard solution for PCR was prepared according to Innis and Gelfand (1990). The reactions were performed in a Perkin-Elmer-Cetus (Norwalk, CT) DNA thermal Cycler. After an initial 10 min incubation at 90°C, 30 cycles of DNA synthesis were carried out, each consisting of denaturation (94°C, 1 min), annealing (60°C, 2 min) and extension by *Taq* DNA polymerase (72°C, 3 min).

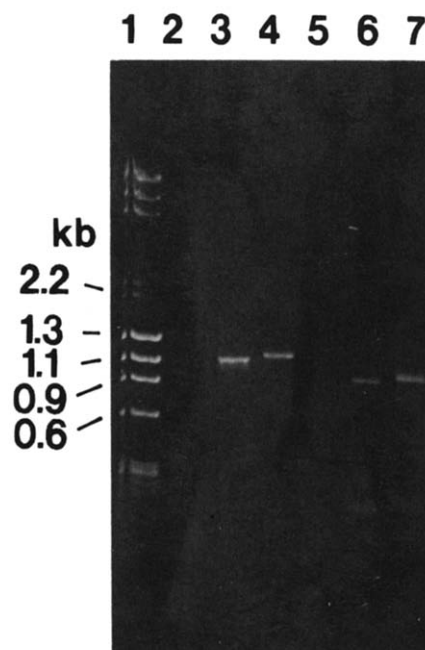


Fig. 2. Amplification of the rice *waxy* sequence. Lane 1 contains the same size marker as in Fig. 1. Lanes 2–7 show the results of the reactions using pairs of primers (see Fig. 3) HVW-2 and HVW-4R, HVW-2 and HVW-6R, HVW-2 and HVW-7R, HVW-3 and HVW-4R, HVW-3 and HVW-6R, and HVW-3 and HVW-7R, respectively. PCR was performed using 1 µg of rice genomic DNA in each reaction. Genomic DNA was extracted from leaves of 2 weeks old rice (*Oryza sativa* L. (Japonica)) plants according to the method of Rogers and Bendich (1985).

ever, were found unsuitable for this method. These results were consistently obtained over the course of many experiments, suggesting that heating for 10 min at 90°C prior to PCR releases a sufficient amount of plasmid DNA from the bacteria. DNA isolation and purification steps can be replaced by a simple heating step carried out in the complete PCR reaction mixture. This protocol should facilitate rapid screening of recombinant plasmids.

In addition, we examined whether this procedure could be applied to a mixture of clones. Reaction mixture composed of five independent clones whose inserts varied in size between 0.2 kb and 3.4 kb, showed distinct bands of each fragment on agarose gels, but mixtures with ten clones did not (data not shown). In the case of mixtures composed of plasmids containing both small inserts (less than 0.6 kb) and longer ones (more than 1.0 kb), longer inserts occasionally resulted in only faint bands on the gels. Nevertheless, our results indicate the usefulness of this procedure for rapid screening of plasmids carrying inserts of at least up to 3 kb in length.

#### (b) Amplification of the rice *waxy* sequence

In PCR, selection of primers is important. If both primers can hybridize to the expected sites, even if only at low stringency, a DNA fragment will be amplified. Thus, a low level of mismatch may be tolerated for the amplification of specific sequences. As described in the Introduction, *waxy* genes contain sequences highly conserved at least among some grass species. An alignment of *waxy* genes indicated that, except for exon-1, exon sequences were highly conserved (data not shown). These results suggested that the sequence of PCR primers could be based upon the sequences in the exons.

For targeting rice *waxy* sequences five different primers, designated as HVW-2, HVW-3, HVW-4R, HVW-6R and HVW-7R, were synthesized (Fig. 3). The 5' ends of the oligos were phosphorylated using polynucleotide kinase. HVW-2, HVW-4R and HVW-6R are completely identical to both the barley and maize sequences, while HVW-3 and HVW-7R were taken from barley sequence. PCR reactions pairing either HVW-2 or HVW-3 with either HVW-6R or HVW-7R yielded single bands near 1 kb in size as judged by gel electrophoresis (Fig. 2). Reactions primed with HVW-4R failed to yield detectable product. Next, amplified fragments were analyzed for size on agarose gels (Fig. 2). The major product obtained in each reaction was consistent in size (between 1.0 and 1.2 kb) with that predicted from the barley and maize sequences.

#### (c) Cloning and sequence analysis of PCR products

The fragment amplified between HVW-2 and HVW-6R was inserted into the *Sma*I site of pUC13 after making it blunt-ended using T4 DNA polymerase. Transformed bac-

terial colonies were analyzed by the PCR method described in section a. From 40 transformants, three independent clones, designated pWX12, pWX15 and pWX27, were selected.

Sequences of pWX12, pWX15 and pWX27 were shown to vary from each other, presumably due to nt misincorporation during the PCR (Fig. 3). Over the 1024-bp sequence, there were six positions at which one sequence deviated from the other two. Assuming that the altered nt resulted from misincorporation, this gives a 0.2% error frequency in the final PCR amplification product. This value is comparable with the results of Saiki et al. (1988) who found an overall error frequency of 0.25%.

Comparison of sequences shows a high level of sequence identity between the rice sequence obtained by PCR and the maize and barley sequences, especially in the regions of the putative exons. Therefore, it seems that our cloned PCR sequences are derived from the rice *waxy* gene. Though it has been reported that two introns were lost from the barley sequence, the rice sequence contains both putative introns at the same positions as in maize.

The priming sites of HVW-3 and HVW-4R were contained within the amplified region. Comparison of these artificial primer sequences with the actual genome sequences revealed three differences in the case of HVW-3 and two differences in the case of HVW-4R (Fig. 3). Since amplification with HVW-3 was successful while amplification with HVW-4R was not, it seems that the number of mismatches may not have been the decisive factor in their ability to function. The most serious obstacle to effective priming by HVW-4R may have been self-hybridization of the primer due to a local palindromic structure.

#### (d) Conclusions

We have developed an efficient PCR-based method for screening recombinant plasmids. The primary constraints limiting the number of plasmid clones which can be screened by an investigator in a single day are the availability of wells in a thermocycler and the number of gel slots available.

Using this procedure, we successfully screened the clones in a rice genomic *waxy* PCR amplification library. Though PCR occasionally resulted in base-substitution mutations in the derived clones, PCR enabled us to rapidly isolate the target sequences. Such amplified fragments may be used as probes in Southern-blot analysis, as homologous RFLP markers or in the construction of antisense genes. Using one of the isolated *waxy* clones as probe, a distinct *waxy* band was detected on a Southern transfer of rice genomic DNA (data not shown).

We anticipate that the low level of base substitutions will not seriously impair the use of our clones in antisense strategies to reduce the level of the rice *waxy* gene product.

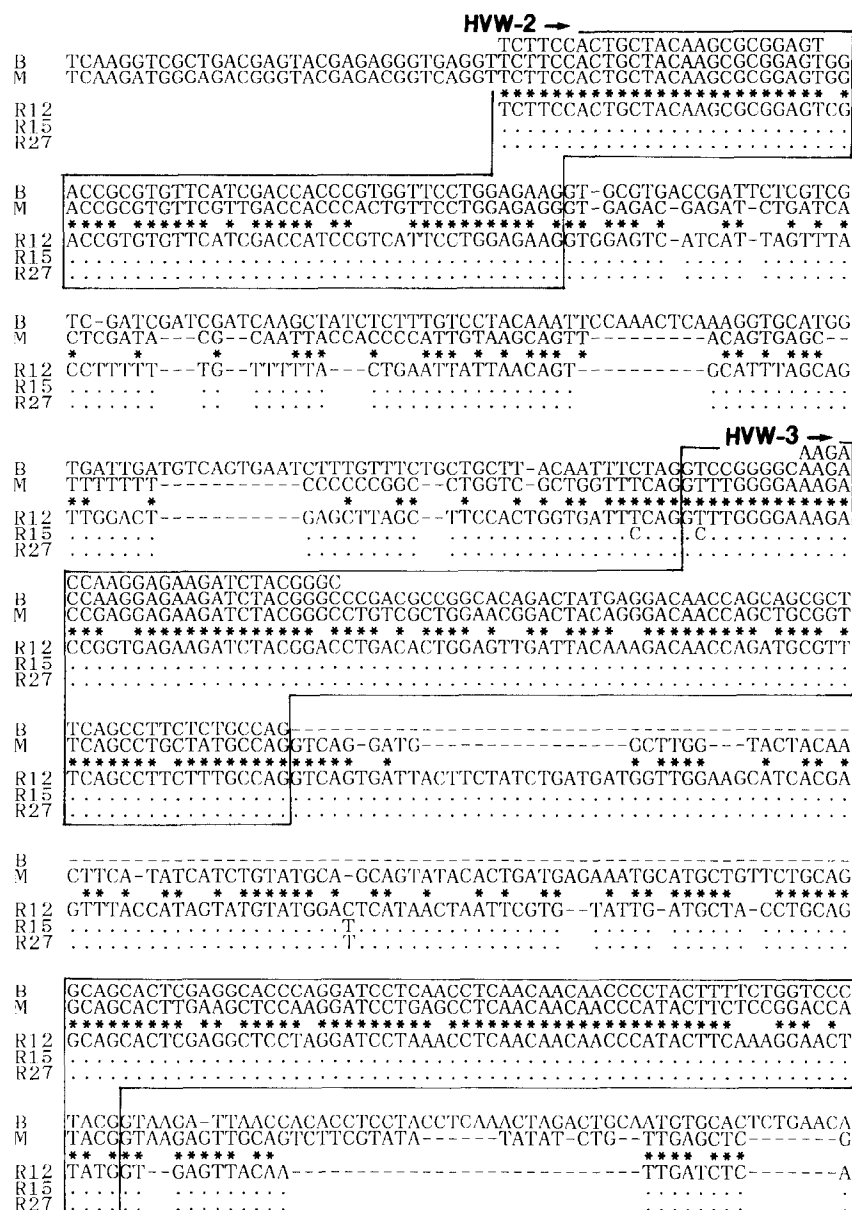


Fig. 3. Alignments of the rice, maize and barley *waxy* sequences. Lines marked by B, M, R12, R15 and R27 indicate the sequences of barley, maize, and rice clones pWX12, pWX15 and pWX27, respectively. In rice clones pWX15 and pWX27, nt identical to that in the pWX12 clones are shown as small dots. Primer sequences used in this experiment are indicated above or below the lines, labeled HVW-2, HVW-3, HVW-4R, HVW-6R or HVW-7R, accordingly. Hyphens indicating gaps are introduced to maximize homology. The nt identical between maize and pWX12 are denoted with asterisks. Computer-aided analysis of the nt sequences was carried out using the GENETYX program (Software Development Co., Japan) on an NEC PC9801 computer.

```

B   AGTTGAAA---TTGAGAGCAA-CGGATGATCACCATTTTCTTGATCTTGGTGCCGCGG
M   AGA-----ATCTTCACAGGAAGCGGCCCATCAGACGGACTGTCATTTTACACTGACTACT
    **      *      *      *      *      *      *      *      *      *
R12 AG-----ATCTTT-----ATAACTTTCTTC-GAAGGAAT-CCATGATGATCAGACTAAT
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   CCCCCTGTCGTTCTTC----GCCGTGCA-GGGGAAGACGTGGTGTTCGTGTGCAACGAC
M   GCTGCTGC---TCTTCGTCCATCCATACAAGGGGAGGACGTCGTGTTTCGTGTGCAACGAC
    **      *      *      *      *      *      *      *      *      *
R12 TC-CTTCC---GGTTTGTACTGACGAC-AGGTGAGGATGTTGTGTTTCGTGTGCAACGAC
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   TGGCACACGGGCTTCTGGCCTGTACCTCAAGAGCAACTACCAGTCCAATGGCATCTAC
M   TGGCACACCGGCCCTCTCTCGTGTACCTCAAGAGCAACTACCAGTCCAATGGCATCTAC
    **      *      *      *      *      *      *      *      *      *
R12 TGGCACACTGGGCCACTGGCGAGCTACCTGAAGAACACCTACCAGACCAATGGCATCTAC
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   AGGACGGCCAAAGCTTTTGAATTTGCA-TCTTTTGAATCTACATATTCTCTCTGCATATA
M   AGGGACGCAAAAGCTTGCCTTCTCTGAACGTAA--CAACGCCGTTTTCGTTCTCCATGCTC
    **      *      *      *      *      *      *      *      *      *
R12 AGGAATGCAAAAG-----TCT--ATG-----GTTGTTCTTGCCATACCAACTC
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   TATATATATATATTGTTGTCCATCTGAACCTACCCCTGGAA-----GAAACTGACTG
M   GTATATACCTCGT--CTGGTAGTGGTGTGCTTCTCTGAGAACTAACTGAAACTGACTG
    **      *      *      *      *      *      *      *      *      *
R12 AAATCTGCAT-----GC-ACACTGCATTCTGTTTCAGAACTGA---
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   -----CCATCTGCTT-----TTGCTAGGTGGCCTTCTGCATCCACAACATCTCG
M   CATGCTGTCTGACCATCTTACGTACTACCAGACCGCTTCTGCATCCACAACATCTCC
    **      *      *      *      *      *      *      *      *      *
R12 -----CTGTCTGA--ATCTT-TTCACT-GCAGGTTGCTTCTGCATCCACAACATCTCC
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   TACCAGGGCGGCTTCTCCTTCGACGACTTGGCGAGCTCAACCTGCCCGACAGGTTCAAG
M   TACCAGGGCGGCTTCCGCTTCTCCGACTACCCGGAGCTGAACCTCCCGAGAGATTCAAG
    **      *      *      *      *      *      *      *      *      *
R12 TACCAGGGCGGTTTCGCTTTCGAGGATTACCCTGAGCTGAACCTCTCCGAGAGGTTCAAG
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   TCGTCCCTTCGACTTCATTGACGG-----GTTTTCCTGCG-----TGCATGTGAA--
M   TCGTCCCTTCGATTTCATCGACGGGCTCT-----GTTTTCCTGCG-----TGCATGTGAA--
    **      *      *      *      *      *      *      *      *      *
R12 TCATCCCTTCGATTTCATCGACGGGTATGAGTAAGATTCTAAGAGTAACCTACTGTCAATT
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   AGCAGGAAGCTGAAGTAACT
M   AGCAGGAAGCTGAAGTAACT
    **      *      *      *      *      *      *      *      *      *
R12 AGCAGGAAGCTGAAGTAACT
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   ---CATTATGAATGGTAACCCACAACCTGTTCCGG-TCCTGCTGGTTTCAATTA---TCTGA
M   ---CATTATGAATGGTAACCCACAACCTGTTCCGG-TCCTGCTGGTTTCAATTA---TCTGA
    **      *      *      *      *      *      *      *      *      *
R12 CGCCATATATCGACTCAATCCAAGATCCTTTTGAGCTGACAACCCCTGCACCTACTGTCCAT
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   -----CTACGACAAGCCCGTGGAGGGGCGCAAGATCAACTG
M   C---CTGAT--TGCATTATTGCAGCTACGAGAAGCCCGTGGAAAGCCGGAAGATCAACTG
    **      *      *      *      *      *      *      *      *      *
R12 CGTTCAAATCCGGTTAAATTTCAAGCTATGACACGCCGCTGGAGGGCAGGAAGATCAACTG
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   GATGAAGGCCGGGATCCTGCAGCCGACAAGGTGCTGACGGTGAGCCCTACTACGGTGA
M   GATGAAGGCCGGGATCCTGCAGCCGACAAGGTGCTGACGGTGAGCCCTACTACGGTGA
    **      *      *      *      *      *      *      *      *      *
R12 GATGAAGGCCGGAATCCTGGAAGCCGACAGGGTGCTACCCGTGAGCCCTACTACGGTGA
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   GGAGCTCATCTCTGGCGAAGCCAGGGGCTGCGAGCTCGACAACATCATGCGCCTACGGG
M   GGAGCTCATCTCCGGCATCGCCAGGGGCTGCGAGCTCGACAACATCATGCGCCTACGGG
    **      *      *      *      *      *      *      *      *      *
R12 GGAGCTCATCTCCGGCATCGCCAGGGGCTGCGAGCTCGACAACATCATGC
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

B   GATCACCGGGAATCGTCAACGGCATGGACGTGAGTGGGACCCACCAAGGACAAGTT
M   CATCACCGGGATCGTCAACGGCATGGACGTGAGTGGGACCCAGCAGGGACAAGTA
    **      *      *      *      *      *      *      *      *      *
R12 TTAGCAGTTGCCGTACCTGCAGTCA
R15 ..      ..      ..      ..      ..      ..      ..      ..      ..
R27 ..      ..      ..      ..      ..      ..      ..      ..      ..

```

← HVW-4R

← HVW-6R

← HVW-7R

## ACKNOWLEDGEMENTS

We thank T. Fujimura and R.F. Whittier for critical reading of this paper and discussions, M. Sugita and M. Sugiura for giving plasmids and valuable suggestions, M. Ishikawa for technical assistance and A. Hirose for continuous encouragements to carry out this project. The *waxy* sequence reported here has been communicated to the GenBank database (accession No.: M55039).

## REFERENCES

- Innis, M.A. and Gelfand, D.H.: Optimization of PCRs. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (Eds.), *PCR Protocols, a Guide to Methods and Applications*. Academic Press, San Diego, CA, 1990, pp. 3–12.
- Kim, H.-S. and Smithies, O.: Recombinant fragment assay for gene targeting based on the polymerase chain reaction. *Nucleic Acids Res.* 16 (1988) 8887–8903.
- Klößgen, R.B., Gierl, A., Schwarz-Sommer, Z. and Saedler, H.: Molecular analysis of the *waxy* locus of *Zea mays*. *Mol. Gen. Genet.* 203 (1986) 237–244.
- Okagaki, R.J. and Wessler, S.: Comparison of non-mutant and mutant *waxy* genes in rice and maize. *Genetics* 120 (1988) 1137–1143.
- Rogers, S.O. and Bendich, A.J.: Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5 (1985) 69–76.
- Rohde, W., Becker, D. and Salamini, F.: Structural analysis of the *waxy* locus from *Hordeum vulgare*. *Nucleic Acids Res.* 16 (1988) 7185–7186.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A.: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239 (1988) 487–491.
- Sano, Y., Katsumata, M. and Okuno, K.: Genetic studies of speciation in cultivated rice. 5. Inter- and intraspecific differentiation in the *waxy* gene expression in rice. *Euphytica* 35 (1986) 1–9.
- Shimada, H., Whittier, R.F., Hiratsuka, J., Maeda, Y., Hirai, A. and Sugiura, M.: A physical map and clone bank of the rice (*Oryza sativa*) chloroplast genome. *Plant Mol. Biol. Repr.* 7 (1989) 284–291.
- Sugita, M., Manzara, T., Pichersky, E., Cashmore, A. and Gruissem, W.: Genomic organization, sequence analysis and expression of all five genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from tomato. *Mol. Gen. Genet.* 209 (1987) 247–256.
- Vieira, J. and Messing, J.: Production of single-stranded plasmid DNA. *Methods Enzymol.* 153 (1987) 3–11.
- Yano, M., Okuno, K., Satoh, H. and Omura, T.: Chromosomal location of genes conditioning low amylose content of endosperm starches in rice, *Oryza sativa* L. *Theor. Appl. Genet.* 76 (1988) 183–189.