Application of Single-Nucleotide Polymorphism Analysis of the trnK Gene to the Identification of Curcuma Plants

Yohei SASAKI, a Hirotoshi FUSHIMI, a and Katsuko KOMATSU a,b

Received August 27, 2003; accepted October 27, 2003

We previously found that Curcuma plants and drugs derived from Curcuma longa, C. phaeocaulis, C. zedoaria, and C. aromatica could be identified by the nucleotide differences at two sites and the existence of a 4-base indel on trnK gene. In this paper, based on species-specific nucleotide sequences, the application of a new method, single-nucleotide polymorphism (SNP) analysis was investigated to identify Curcuma plants more conveniently. First, three types of reverse primer were synthesized in different lengths, 34 mer, 26 mer, and 30 mer, to anneal the template DNAs from each species at sites immediately upstream from substitution positions 177 and 645, and at the site including the 4-base insertion from 728 to 731, respectively. After single-base extension reaction of these primers using fluorescent-labeled ddNTPs and PCR products of the trnK gene region as template, the resulting products were detected using an ABI PRISM 310 Genetic Analyzer. The electrophoretogram showed three or two peaks at different positions depending on the 27 mer, 31 mer, and 35 mer product lengths. Each peak was derived from the incorporated fluorescent-labeled ddNMPs complementary to template nucleotides at positions 645, 724, and 177, respectively. C. phaeocaulis showed three peaks of ddCMP, ddAMP, and ddAMP. The other three species showed two peaks derived from 27 mer and 35 mer product peaks: peaks of ddCMP and ddAMP in C. longa, those of ddCMP and ddTMP in C. zedoaria, and those of ddTMP and ddAMP in C. aromatica. Thus SNP analysis to identify four Curcuma plants was newly developed.

Key words Curcuma; trnK gene; single-nucleotide polymorphism (SNP); amplification-refractory mutation system (ARMS); identification; Zingiberaeace

In the field of clinical laboratory medicine, genetic polymorphisms on a human gene coding drug-metabolizing enzymes, cytochrome P450, has attracted attention because patients from different populations or individuals show different drug responses. To analyze the phenomenon of genetic polymorphisms, several single-nucleotide polymorphism (SNP) analyses using PCR-restriction fragment-length polymorphism (RFLP), amplification-refractory mutation system (ARMS), TaqMan assay, and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry have been developed and provide beneficial information. In the present study, one SNP analysis, a multiplexing primer-extension analysis method recently established, is applied to the authentication of Curcuma plants. We previously clarified nucleotide differences on the nuclear 18S rRNA gene and plastid trnK gene in six Curcuma species distributed in China and Japan. Curcuma longa, C. phaeocaulis, C. zedoaria, and C. aromatica were distinguished from each other by nucleotide substitutions at positions 177 and 645 (from upstream), and by the existence of a 4-base indel from 728 to 731 in the trnK gene (Fig. 1). That is, both C. longa and C. phaeocaulis had thymine and guanine, C. zedoaria had adenine and guanine, and C. aromatica had thymine and adenine at these substitution sites, respectively. Moreover, C. phaeocaulis only had the insertion. These species-specific substitutions and insertion made it possible to identify Curcuma plants and drugs using sequence analysis. In the previous study, we developed an ARMS method to detect these substitutions and indel, and identified Curcuma drugs by both sequence analysis and the ARMS method. However, the ARMS method requires searching optimal reaction conditions for every template DNA and primer set. Because of the strict reaction conditions, the results probably differ based on the performance of the thermal controller. Moreover, there is no convenient method to detect the existence of indels among available SNP analyses. Therefore we attempted to develop a multiplexing primer-extension analysis method including an annealing reaction to the insertion site to identify Curcuma plants.

As materials, four specimens of different species were used: C. longa, C. zedoaria, and C. aromatica from Japan, and C. phaeocaulis from China, with voucher numbers Komatsu, O, 06, and 02, and Cao 9956, respectively.

First, we designed three types of reverse primer of different lengths, 26 mer, 30 mer, and 34 mer, to anneal the template DNA from each species on the basis of species-specific nucleotide sequences at positions 177, 645, and 728—731. The primers synthesized were: CT26MER, 5’-AAA ATG GTT TAT TCT CGT TAT TAA TA-3’, designed for annealing to template DNAs at a site immediately upstream from the nucleotide position 645; CT30MER, 5’-ATA TAC AAT AAT AAA AAA CCT GTA TG-3’, designed for annealing to the region including the 4-base insertion from positions 728 to 731 in C. phaeocaulis; and CT34MER, 5’-ATT TTT TCA ATG AAA ATA AAA TAA TAA TAA TCA ATG AAA ATA AAA TAA TAA TAC TAT T-3’, designed for annealing to template DNAs immediately upstream from position 177. After single-base extension reactions of these three primers were performed using PCR products of the trnK gene region from each species as a template, together with fluorescent-labeled ddNTPs, the resulting products were detected using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, U.S.A.). The reaction mixture of 10 µl consisted of the three primers (1.0 µM of CT26MER, 0.6 µM of CT30MER, and 0.2 µM of CT34MER), purified PCR products of the trnK gene region (about 0.2 pmol) from each species, and SNaPshot Multiplex Ready Reaction Mix (Applied Biosystems) containing the fluorescent-labeled 2’,3’-dideoxynucleoside triphosphates (F)ddNTPs composed of ddGTP, ddATP, ddCTP and ddTTP, labeled with different fluorescence). DNA polymerase, and reaction buffer. The cycling conditions were as follows: hot start at 96 °C for 1 min, followed by 25 cycles at 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 30 s. In this reaction, single [F]ddNTP complementary to a template nucleotide was incorporated onto the 3’ end of each primer. After treatment with 1.0 unit of shrimp alkaline phosphatase (SAP; Sigma, U.S.A.) at 37°C for 1 h, deactivation at 75°C for 15 min, and dilution with Hi-Di formamide (Applied Biosystems), the reaction mixtures were
analyzed using the ABI PRISM 310 Genetic Analyzer with POP-4 polymer (Applied Biosystems) and the software Gene Scan 3.1.2. In the electrophoretograms, the fluorescent labeled 2',3'-dideoxynucleoside monophosphates ([F]ddN-MPs) incorporated onto the 3' end of the primer were detected as colored peaks. The incorporated [F]ddNMPs are complementary to the nucleotides at the positions 177, 645, and 724 (if the 4-base insertion exists from 728 to 731).

The resulting products, the single-base extended primers, were oligonucleotides of 27 mer, 31 mer, and 35 mer in length with specific [F]ddNMP attached at 3’ end. In the electrophoretograms, three or two colored peaks were detected at different positions depending on the length of products. Each peak was derived from the incorporated [F]ddNMPs at nucleotide positions 645, 724, and 177, respectively. *C. phaeocaulis* showed three peaks of ddCMP (black in color in Fig. 2), ddAMP (green), and ddAMP (green), respectively. The other three species showed two peaks derived from 27 mer and 35 mer products: peaks of ddCMP (black) and ddAMP (green) in *C. longa*, those of ddCMP (black) and

Fig. 1. Sites of Nucleotide Substitutions and Indels in Four Curcuma Species and Locations of Primers Designed for SNP Analysis

The numerals in italics indicate the aligned nucleotide position from upstream. The broad half-arrows indicate primers used in this analysis. Incorporated ddAMP, ddCMP, and ddTMP onto the 3' end of each primer in the single-base extension reaction are indicated by uppercase A (green), C (black), and T (red), respectively.

![Fig. 1](image)

Fig. 2. Electrophoretograms of the Reaction Products in Four Curcuma Species

The resulting products using 3 primers, CT26MER, CT30MER, and CT34MER, in the single-base extension reaction were 27 mer, 31 mer, and 35 mer in length, respectively. Letters over the electrophoretograms indicate the incorporated [F]ddNMPs. ddAMP, ddCMP, and ddTMP are indicated by uppercase A (green), C (black), and T (red), respectively.

![Fig. 2](image)
ddTMP (red) in *C. zedoaria*, and those of ddTMP (red) and ddAMP (green) in *C. aromatica*, respectively (Fig. 2). Thus the SNP analysis to detect nucleotide differences and the existence of an insertion site of four *Curcuma* species was newly developed.

This method has the advantages of not requiring the fluorescent-labeled primer used in the Taq Man assay and the optimal thermal conditions of single-nucleotide extension can be established easily, *i.e.*, the effort to determine the optimal conditions as in the ARMS assay is unnecessary. Moreover, multiple-nucleotide polymorphisms in the same gene region or several different gene regions can be targeted simultaneously as only one template, although only the *trnK* gene region was examined in this study.

Since it is difficult to identify the botanical origins of *Curcuma* drugs such as Ezhu morphologically and phytochemically, the SNP analysis method developed in the present study promises to become a useful identification method.

**Acknowledgments** We are grateful to Professor Toshio Miyawaki and Dr. Hirokazu Kanegane, Faculty of Medicine, Toyama Medical and Pharmaceutical University, for permitting us to use the ABI PRISM 310 Genetic Analyzer. This work was supported by a Grant-in-Aid for Scientific Research (B), No. 11695086 in 1999—2001 and No. 14406030 in 2002—2003 from the Japan Society for the Promotion of Science, and for the 21st Century COE Program from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

**REFERENCES**