Food Chemistry 205 (2016) 272–279

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Whole genome sequence analysis of unidentified genetically modified papaya for development of a specific detection method

Kosuke Nakamura [a,*], Kazunari Kondo [a,*], Hiroshi Akiyama [a], Takumi Ishigaki [a], Akio Noguchi [a], Hiroshi Katsumata [b], Kazuto Takasaki [b], Satoshi Futo [b], Kozue Sakata [a], Nozomi Fukuda [a], Junichi Mano [c], Kazumi Kitta [c], Hidenori Tanaka [d], Ryo Akashi [d], Tomoko Nishimaki-Mogami [a]

[a] National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[b] FASMAC CO., LTD, 5-1-3 Midorigaoka, Atsugi, Kanagawa 243-0041, Japan
[c] Analytical Science Division, National Food Research Institute, National Agriculture and Food Research Organization, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan
[d] Graduate School of Agriculture, University of Miyazaki, 1-1 Gakuen Kibanadai Nishi, Miyazaki 889-2192, Japan

ARTICLE INFO

Article history:
Received 2 November 2015
Received in revised form 22 February 2016
Accepted 27 February 2016
Available online 3 March 2016

Keywords:
Genetically modified
Transgenic sequence
Next-generation sequencer
Genome sequence
Real-time PCR
Carica papaya L.

ABSTRACT

Identification of transgenic sequences in an unknown genetically modified (GM) papaya (*Carica papaya* L.) by whole genome sequence analysis was demonstrated. Whole genome sequence data were generated for a GM-positive fresh papaya fruit commodity detected in monitoring using real-time polymerase chain reaction (PCR). The sequences obtained were mapped against an open database for papaya genome sequence. Transgenic construct- and event-specific sequences were identified as a GM papaya developed to resist infection from a *Papaya ringspot virus*. Based on the transgenic sequences, a specific real-time PCR detection method for GM papaya applicable to various food commodities was developed. Whole genome sequence analysis enabled identifying unknown transgenic construct- and event-specific sequences in GM papaya and development of a reliable method for detecting them in papaya food commodities.

1. Introduction

Genetically modified (GM) crops for food (GM food) are being increasingly developed (James, 2014). Potential toxicity or allergenicity derived from introduction of transgenic sequences in the genome is one of the concerns over acceptance of GM food by consumers. Therefore, use of GM food is regulated by many countries. Safety assessments of GM food are mandatory and independently performed based on safety assessment criteria, such as in the EU (Regulation (EC) No 1829/2003, 2003) and Japan (Standards for the Safety Assessment of GM Foods (Seed Plants), 2004). In these countries, any unauthorized GM foods are prohibited from the market since authorization of GM food use is stipulated under law.

In some countries, unintended use of unauthorized GM food has been reported in various food commodities, such as rice (*Oryza sativa* L.) (Grohmann et al., 2015; Nakamura et al., 2013a), flax (*Linum usitatissimum* L.) (Young et al., 2015), wheat (*Triticum aestivum* L.) (Kim et al., 2015), potato (*Solanum tuberosum* L.) (Akiyama et al., 2002a, 2002b), and papaya (*Carica papaya* L.). In the case of papaya, an unauthorized GM papaya, which resists infection from *Papaya ringspot virus* (*PRSV*), was found in some processed papaya commodities (papaya-leaf tea, papaya jam, papaya pickles) in Japan in 2011 (Nakamura et al., 2011). Another unauthorized GM papaya line developed to effectively resist infection from the *PRSV* strains in Thailand was found in a processed papaya commodity (dried papaya) in Japan in 2013 (Nakamura et al., 2014). In the EU, GM papaya has been found in some imported foods (RASFF Portal. Unauthorised genetically modified papaya from unknown origin, 2015). As compared to non-GM papaya, no *de novo* toxicity or allergenicity by GM papaya ingestion has been reported so far (Fermín et al., 2011; Guo et al., 2009; Lin et al., 2013; Lin et al., 2015). Some GM papaya lines have been safely consumed as food by human more than a decade. Monitoring foods

Abbreviations: bp, base pair; *Chy*, *chymopapain*; Ct, threshold cycle; ΔRn, change in normalized reporter signal; DNA, deoxyribonucleic acid; FAM, 6-carboxyfluorescein; GM, genetically modified; LOD, limit of detection; NOSt, *nopaline synthase* terminator; *NPTII*, *neomycin phosphotransferase II*; P35S, *Cauliflower mosaic virus* 35S RNA promoter; *PRSV*, *Papaya ringspot virus*; *CP*, *coat protein*; PCR, polymerase chain reaction; QV, quality value; RSD, relative standard deviation; TAMRA, 6-carboxytetramethyl-rhodamine; UV, ultraviolet.

* Corresponding authors.
E-mail addresses: kosnakamura@nihs.go.jp (K. Nakamura), kondo@nihs.go.jp (K. Kondo).

http://dx.doi.org/10.1016/j.foodchem.2016.02.157
0308-8146/

containing any novel unauthorized GM papaya lines is required to follow the regulation and to ensure food safety.

Some GM papaya lines have already been successful in production and commercialization for foods. *PRSV*-resistant GM papaya line 55-1 (line 55-1) from Hawaii, for example, was the first commercialized GM papaya for food use after regulatory approval in countries such as the United States, Canada and Japan (Fermín et al., 2011; Nakamura et al., 2013b). Many GM papaya lines having different transgenic constructs have been developed worldwide (Mendoza, Laurena, & Botella, 2008). In Japan, only line 55-1 for food use has been authorized. All other GM papaya is unauthorized and prohibited from being distributed to the market in Japan before gaining Japanese regulatory approval. Therefore, laboratories testing GM food have been monitoring food commodities to avoid distribution of unauthorized GM papaya in Japan.

The most frequently used approach to detect the presence of GM food is detection of common transgenic construct sequences derived from viral and bacterial genomes, such as *Cauliflower mosaic virus 35S RNA* promoter (P35S), the 3′-untranslated region of the *Rhizobium radiobacter nopaline synthase* terminator (NOSt) or *neomycin phosphotransferase II* (*NPTII*) using conventional polymerase chain reaction (PCR) and real-time PCR methods (Mano et al., 2009; Reiting, Broll, Waiblinger, & Grohmann, 2007). However, false-positive results are frequently observed in test monitoring various food commodities due to viral and bacterial contamination (Kitagawa, Nakamura, Kondo, Ubukata, & Akiyama, 2014). Also, detection and identification of transgenic construct- or event-specific sequences in an unknown GM food using methods such as chromosome walking and inverse PCR require highly trained experimental skills and are time consuming.

In this study, whole genome sequence data were generated for a GM-positive fresh papaya fruit commodity detected in a monitoring test using real-time PCR. Common transgenic construct sequences, P35S and border regions in *Agrobacterium* T-DNA transformation, used in GM papaya development were exploited to search for sequence reads having sequences specific to GM papaya. Transgenic construct- and event-specific DNA sequences of the GM papaya were identified from the sequencing data. Based on the identified sequences, a specific, repeatable and reproducible real-time PCR detection method applicable to testing various types of papaya food commodities was developed.

2. Materials and methods

2.1. Papaya samples

Papaya food commodities (pickled papaya, papaya-leaf tea, jam) were purchased through the Internet. A fresh papaya fruit was obtained as GM papaya line PRSV-YK using transgenic construct-specific real-time PCR method during a monitoring test for commodities covered by regulations on GM food use in Japan. Fresh non-GM papaya (cultivar Sunset) fruit was purchased from the Hawaii Papaya Industry Association through the Consumer Affairs Agency of Japan.

2.2. Purification of papaya DNA

Sarcocarp was collected from a peeled fresh papaya fruit with the seeds removed, then 10 g aliquots were processed with a food processor and used for DNA purification. Other papaya food commodities, except jam, were ground using a food processor, then 10 g aliquots used for DNA purification. DNA was purified using an ion-exchange resin-type DNA extraction and purification kit (Genomic-tip; QIAGEN, Hilden, Germany) according to methods described previously (Ohmori et al., 2013). Purified DNA was quantified for purity and concentration by measuring UV absorption using an ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA).

2.3. Transgenic construct sequence detection using real-time PCR

DNA sequences coding for *NPTII* and regulatory elements in transgenic constructs, P35S and NOSt, were targeted for real-time PCR detection to monitor papaya food commodities for GM papaya use. Detection of *chymopapain* (*Chy*, GenBank accession No. AY803756) in the papaya genome was used for endogenous gene detection. All nucleotide sequences of primers and probes used are listed in Table 1.

Each reaction contained 12.5 μL TaqMan Gene Expression Master Mix (Thermo Fisher Scientific Inc., Waltham, MA, USA), 0.8 μM each primer, 0.1 μM probe and 50 ng DNA in a final volume of 25 μL. Each sample was tested in duplicate. PCR was performed in MicroAmp Optical 96-well Reaction Plates (Thermo Fisher Scientific Inc.). An ABI PRISM 7900HT instrument was used for the analyses. The PCR conditions were 2 min at 50 °C, 95 °C for 10 min followed by 50 cycles of 15 s at 95 °C and 1 min at 60 °C. After completion of the amplification reaction, data were analyzed using SDS 2.1 sequence detection software (Thermo Fisher Scientific Inc.) for the ABI PRISM 7900HT Sequence Detection System. The ΔRn threshold was set to 0.2 during exponential amplification and threshold cycle (Ct) values were recorded. When both an exponential increase in fluorescence intensity of dye from the target (6-carboxyfluorescein) and a Ct value <48.00 were confirmed, the sample was scored as positive. If no Ct value at the threshold value of 0.2 could be obtained, the reaction was scored as negative. Also, reactions with Ct values <48.00 but without exponential amplification, as judged by visual inspection of respective ΔRn and multi-component plots, were scored as negative.

2.4. Whole genome sequencing and data analyses

Purified DNA samples from fresh papaya fruit were used to prepare DNA libraries. DNA (250 ng) was fragmented using NEBNext dsDNA Fragmentase (New England Biolabs Japan, Inc., Tokyo, Japan). DNA libraries were prepared using NEBNext Ultra DNA Library Preparation Kit (New England Biolabs Japan, Inc.). The quality of the constructed libraries was confirmed using a Qubit DNA HS Kit (Thermo Fisher Scientific Inc.) and Bioanalyzer DNA High Sensitivity Kit (Agilent Technologies, Santa Clara, CA, USA). A single whole genome sequencing was run by the MiSeq sequencing system (Illumina, San Diego, CA, USA) using a Reagent V3 (600 cycles) Kit. The obtained reads were trimmed at quality value (QV) 20 using Sickle ver. 1.290 software (https://github.com/najoshi/sickle). Sequencing and quality trimming of low quality reads resulted in a data set containing 18,177,038 paired-end reads (10,906,222,800 bases). Within the data set, a total of 17,375,285 paired-end reads having more than 50 bases at QV 20 were mapped using Bowtie 2 software (Langmead & Salzberg, 2012) against the papaya genome sequence (GenBank assembly accession: GCA_000150535.1) (Ming et al., 2008) as a reference. Mapped paired reads were extracted and assembled together using Velvet (Zerbino & Birney, 2008) at *k*-mer = 61. PRICE contig extension software (Ruby, Bellare, & DeRisi, 2013) was used to iteratively increase the size of existing contigs. Sequence reads were mapped to the reference genome sequences and visualized using Integrative Genomics Viewer version 2.3.40 (Robinson et al., 2011).

2.5. Analysis of transgenic construct sequences in GM papaya

DNA was PCR-amplified using oligonucleotide primers designed using the obtained genome sequences. PCR reactions were