Data Article

Data on screening and identification of genetically modified papaya in food supplements

Theo W. Prins a,*, Ingrid M.J. Scholtens a, Arno W. Bak b, Jeroen P. van Dijk a, Marleen M. Voorhuijzen a, Emile J. Laurensse c, Esther J. Kok a

a RIKILT Wageningen UR, Akkermaalsbos 2, 6708 WB Wageningen, Netherlands
b Netherlands Food and Consumer Product Safety Authority (NVWA), Akkermaalsbos 4, 6708 WB Wageningen, Netherlands
c Netherlands Food and Consumer Product Safety Authority (NVWA), Catharijnesingel 59, 3511GG Utrecht, Netherlands

ABSTRACT

This article contains data related to the research article entitled “A case study to determine the geographical origin of unknown GM papaya in routine food sample analysis, followed by identification of papaya events 16-0-1 and 18-2-4” (Prins et al., 2016) [1]. Quantitative real-time PCR (qPCR) with targets that are putatively present in genetically modified (GM) papaya was used as a first screening to narrow down the vast array of candidates. The combination of elements P-nos and nptII was further confirmed by amplification and subsequent sequencing of the P-nos/nptII construct. Next, presence of the candidate GM papayas 16-0-1 and 18-2-4 were investigated by amplification and sequencing of event-spanning regions on the left and right border. This data article reports the Cq values for GM elements, the nucleotide sequence of the P-nos/nptII construct and the presence of GM papaya events 18-2-4 and/or 16-0-1 in five samples that were randomly sampled to be analysed in the framework of the official Dutch GMO monitoring program for food.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
**Specifications Table**

| Subject area                  | Biology                                           |
|------------------------------|---------------------------------------------------|
| More specific subject area   | Biotechnology, qPCR                                |
| Type of data                 | Table, Figure                                     |
| How data was acquired        | Data are Cq qPCR values calculated with Bio-Rad CFX software, and the nucleotide sequence of P-nos/nptII amplicon generated by Sanger sequencing |
| Data format                  | Analysed                                          |
| Experimental factors         | DNA isolated from 5 papaya food supplements        |
| Experimental features        | qPCR and Sanger sequencing                        |
| Data source location         | Wageningen, Netherlands                           |
| Data accessibility           | Data is within this article and available at GenBank via the accession numbers GenBank: KU376437, GenBank: KU376438, GenBank: KU376439, GenBank: KU376440, GenBank: KU376441, GenBank: KU376442 |

**Value of the data**

- Data presented here provide a screening strategy for the identification of unknown genetically modified papaya.
- We show Cq (quantification cycle) values of common elements putatively present in genetically modified papaya to show that a combination of qPCRs can narrow down the suspect known GM papaya events.
- Using the forward primer of a promoter and a reverse primer of a coding sequence in PCR and subsequent sequencing of the amplicon reveals part of the construct P-nos/nptII.
- We show that PCR with event-specific primers for the left and right border of GM papaya 16-0-1 and 18-2-4 confirms the presence or absence of the GM papaya(s) in food samples.

1. **Data**

The data presented in this article show the Cq values of the available qPCRs for known GM elements in GM papaya (Table 1). The nucleotide sequence (partial) of the P-nos/nptII construct (Fig. 1), and the event-specific borders of GM papaya are presented (Table 2). Data were obtained by extracting DNA from five papaya samples.

2. **Experimental design, materials and methods**

2.1. **Materials**

The analytes consisted of five different papaya food supplements that were sold as capsules with dried and powdered papaya, or papaya powder. See [1] for detailed information.

2.2. **DNA extraction**

According to [1], in brief, the Promega Wizard™ Magnetic DNA Purification system for food (Promega, Madison WI, USA) was used according to the manufacturer’s protocol with minor
modifications. A Nanodrop spectrophotometer (ND-1000, Montchanin, DE, USA) was used to quantify and assess the purity of the DNA.

2.3. Primers and probes

Primer and probes used for the detection of targets described in Table 1, the amplification and sequencing of the partial P-nos/nptII cassette (Fig. 1), and the amplification and sequencing of the left border and right border of events 16-0-1 and 18-2-4 (Table 2), are given in Prins et al. [1].

2.4. PCR reactions

qPCR was performed on a CFX (Bio-Rad). qPCR trace files were analysed with the Bio-Rad CFX manager software version 3.0. For full PCR conditions, consult [1].

2.5. Amplicon analysis

Samples were bidirectionally sequenced by Macrogen Europe (Amsterdam, the Netherlands) using the forward and reverse primers. Sequence data were analysed using Staden Package software (version 1.6.0: for trimming, aligning, and consensus building), BLAST [2] and Clustal Omega [3].
Acknowledgments

This research was funded by Dutch Ministry of Economic Affairs WOT02 Food Safety (WOT-02-004-005).

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.08.028.

References

[1] T.W. Prins, et al., A case study to determine the geographical origin of unknown GM papaya in routine food sample analysis, followed by identification of papaya events 16-0-1 and 18-2-4, Food Chem. 213 (2016) 536–544.
[2] S.F. Altschul, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (17) (1997) 3389–3402.
[3] F. Sievers, et al., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. (2011) 7.

Table 2

| Sample | Method | 16-0-1 LB (105 bp) | 16-0-1 RB (249 bp) | 18-2-4 LB (140 bp) | 18-2-4 RB (309 bp) |
|--------|--------|-------------------|-------------------|-------------------|-------------------|
| A      | +      | +                 | –                 | –                 | –                 |
| B      | +      | +                 | –                 | –                 | –                 |
| C      | +      | +                 | –                 | –                 | –                 |
| D      | +      | +                 | +                 | +                 | +                 |
| E      | +      | +                 | –                 | –                 | –                 |

–: not detected; +: detected.