

Development of a multiplex DNA-based traceability tool for crop plant materials

Marleen M. Voorhuijzen · Jeroen P. van Dijk ·
Theo W. Prins · A. M. Angeline Van Hoef ·
Ralf Seyfarth · Esther J. Kok

Received: 30 June 2011 / Revised: 22 September 2011 / Accepted: 23 October 2011 / Published online: 10 November 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract The authenticity of food is of increasing importance for producers, retailers and consumers. All groups benefit from the correct labelling of the contents of food products. Producers and retailers want to guarantee the origin of their products and check for adulteration with cheaper or inferior ingredients. Consumers are also more demanding about the origin of their food for various socioeconomic reasons. In contrast to this increasing demand, correct labelling has become much more complex because of global transportation networks of raw materials and processed food products. Within the European integrated research project ‘Tracing the origin of food’ (TRACE), a DNA-based multiplex detection tool was developed—the padlock probe ligation and microarray detection (PPLMD) tool. In this paper, this method is extended to a 15-plex traceability tool with a focus on products of commercial importance such as the emmer wheat Farro della Garfagnana (FdG) and Basmati rice. The specificity of 14 plant-related padlock probes was determined and initially validated in mixtures comprising seven or nine plant species/varieties. One nucleotide difference in target sequence was sufficient for the distinction between the presence or absence of a specific target. At least 5% FdG or Basmati rice was detected in mixtures with cheaper bread wheat or non-fragrant rice, respectively. The results suggested that even lower levels of (un-)intentional

adulteration could be detected. PPLMD has been shown to be a useful tool for the detection of fraudulent/intentional admixtures in premium foods and is ready for the monitoring of correct labelling of premium foods worldwide.

Keywords Authenticity · DNA · Ligation · Detection · Traceability · Multiplex

Introduction

The authenticity of food is of increasing importance for producers, retailers and consumers. Producers and retailers want to be able to guarantee, for instance, claims of origin on their products and do not want to be confronted with abuse of such claims by third parties. Consumers show increasing interest in where their food comes from, how it is produced, and whether it is safe and healthy for consumption. For both consumers and retailers, it is important that specific raw materials or derived food products have not been adulterated with ingredients of inferior quality. In the past, obtaining details on products with respect to the origin and type of processing was relatively easy since most food and food products were produced locally. Nowadays, this has become much more complex as a consequence of the global transportation networks of raw materials and processed food products.

Regulation EC/178/2002, the EU General Food Law, defines traceability as the ability to trace and follow food, feed and ingredients through all stages of production, processing and distribution [1]. In addition to the general requirements, sector-specific legislation applies to certain categories of food products such as fruit and vegetables, enabling consumers to identify their origin and authenticity [1]. Furthermore, authenticity and the related labelling are

M. M. Voorhuijzen (✉) · J. P. van Dijk · T. W. Prins ·
A. M. A. Van Hoef · E. J. Kok
RIKILT—Institute of Food Safety, Wageningen UR,
P.O. Box 230, 6700 AE Wageningen, the Netherlands
e-mail: marleen.voorhuijzen@wur.nl

R. Seyfarth
Biolytix AG,
Benkenstrasse 254,
4108 Witterswil, Switzerland

particularly important for special products such as agricultural products or derived foodstuffs with a specific geographical origin. These products can either be listed as products of designated origin (PDOs) or products of protected geographical indication (PGIs) [2]. Within the European integrated research project 'Tracing the origin of food' (TRACE), analytical methods have been developed that can distinguish specialty products from potentially present undesired admixtures. Within the TRACE project, different approaches were tested in order to develop reliable methods for the traceability of specialty products in the different parts of the production and processing chain. One of the goals of TRACE was to develop a DNA-based multiplex detection tool which is rapid, robust, accurate and cost-effective.

The TRACE project focused on a limited number of specialty products. One of these products is the Italian wheat known as Farro della Garfagnana, a PGI registered in 1996. This is an emmer wheat produced from the local population of *Triticum dicoccum*, traditionally used for making pastas or baked products such as bread. Farro della Garfagnana is different from emmer wheat produced in other areas mainly by the fact that it is adapted to the local environment [3].

Farro della Garfagnana contains the AABB genome, unlike the widely used 'modern' common bread wheat (*Triticum aestivum*) or the rarely cultivated spelt (*Triticum spelta*), which have the AABBDD genome [4]. This means that admixtures of Farro della Garfagnana-labelled products with these more modern varieties can be traced on the basis of the presence of the D genome.

The only other AABB-containing cereal of commercial importance is the widely cultivated durum wheat (*Triticum durum*). Durum wheat also lacks the D genome, so it cannot be identified as an admixture in Farro della Garfagnana in this way. Instead, the presence of durum wheat can be determined by genotyping the *Q* locus, which is located on the A genome of both durum and Farro della Garfagnana. *Q* determines the phenotype for the important agronomical trait 'threshability', which distinguishes hulled from free-threshing seeds [5]. Thus, durum wheat, a free-threshing (*Q* allele) wheat species, can be distinguished from Farro della Garfagnana, which is a hulled wheat (*q* allele) [6].

Another special product that was included in the present study is Basmati rice, which is not a PDO or PGI but which has similar commercial importance as its specific fragrance is highly valued [7].

The fragrance of rice (*Oryza sativa*) is associated with the presence of the potent aromatic component 2-acetyl-1-pyrroline (2AP) [8]. Non-fragrant rice varieties contain low levels of 2AP, whilst these levels are considerably higher in aromatic varieties such as Basmati rice [9]. The aroma that

is based on the accumulation of 2AP is the result of a deletion in an allele at the fragrance locus. The betaine aldehyde dehydrogenase2 (*BAD2*) gene on chromosome 8 of rice contains the functional *nfgR* allele in non-fragrant rice. In fragrant rice, the *fgr* allele contains a cluster of single nucleotide polymorphisms and an eight-nucleotide deletion, resulting in an early stop codon in the open reading frame of the *BAD2* gene. The presence of this stop codon leads to a truncated protein and results in the accumulation instead of the breakdown of 2AP [10]. However, aromatic varieties containing the native *nfgR* allele are also found. This suggests that the deletion in the *fgr* allele is not the only cause of 2AP accumulation but that there is at least one other genetic factor involved in controlling fragrance [11, 12].

Specialty products Farro della Garfagnana and Basmati rice can be sold with a premium and are therefore sensitive to fraudulent admixture with cheaper varieties. To detect adulteration of specific raw materials or derived food products, traditionally, singleplex methods are used. DNA-based methods such as quantitative PCR (q-PCR) [13–15], PCR restriction fragment length polymorphism [6] and capillary electrophoresis [16] are considered to be rapid, robust and accurate. However, today's complex matrices and related ingredients demand for multiplex detection methods that can determine in a single assay which type of raw materials have been used to obtain the product and, indeed, whether adulteration of specific raw materials may have occurred. One example of multiplex detection is the previously published strategy of padlock probe ligation and microarray detection (PPLMD). In this strategy, ten different linear padlock probes (PLPs) for event-, element- and species-specific sequences derived from genetically modified organisms were circularized, ligated and detected on microarray in a DNA mixture comprising four species [17].

This method contains a number of steps: detection, amplification and identification. It uses PLPs that contain five specific regions. Two target specific regions, located at the extremities, will cause the PLPs to hybridize in a circular fashion to the template as the complementary target sequences are located next to each other. Only when this hybridization is flawless will a ligase covalently bind the two PLP ends, forming a circular molecule. This is the basis of the detection. For amplification, two primer binding sites are included in the PLPs. When in linear form, these are oriented away from each other, forestalling PCR amplification. When circularized, the primer sites are orientated towards each other, enabling PCR amplification. For identification, a complementary ZIP (cZIP) sequence is included that can hybridize to a ZIP sequence on a microarray. When the PCR is performed with a fluorescently labelled primer, common microarray analysis can be

used for the identification of the amplified circularized PLPs. In a previous work, criteria for the design of PLPs were developed and discussed [18]. A so-called SpikeLock probe was developed as a control for a successful ligation reaction [17].

Here, we describe an extension of this strategy to a 15-plex traceability tool and the subsequent initial validation. In this study, also a comparison between DNA and DNA/locked nucleic acid (LNA) PLPs was made to practically assess the advantage of the incorporation of LNA nucleotides to overcome the poor performance of long PLPs. The main focus of the present study was on the detection of small amounts of specific DNA sequences in more complex DNA mixtures.

Materials and methods

Design and quality control of PLPs

The PLPs for wheat discrimination were designed for the detection of the D genome and both genotypes of the *Q* locus. For rice, PLPs were designed around the difference in sequence of the BAD2 gene to discriminate between Basmati (fragrant) and non-fragrant rice. Melting temperatures of the LNA-containing PLPs were calculated using the LNA Oligo Tm Prediction Tool (Exiqon, Denmark; see Table 1 for the sequences and cZIP codes used). Synthetic single-stranded target sequences were ordered (Biolegio, the Netherlands) to test the performance of the newly designed PLPs in q-PCR, as described by Prins et al. [17].

Plant materials

All cereal ground seed materials, except for rice, were provided by Biolytix (Switzerland) within the TRACE framework (project no. FOOD-CT-2005-006942). For rice, no certified reference material was available, so Basmati rice and conventional (non-fragrant) rice were purchased at a local biological food store.

Ground seed material for soy, maize and sugar beet was purchased from IRMM (Belgium), and canola leaf DNA was purchased from AOCS (Urbana, IL, USA). For detailed information, see Table 2.

DNA extractions

All plant materials were isolated using the Qiagen DNeasy Plant Minikit. Soy (100 mg) was isolated according to the supplied protocol. For all other materials, the extraction step was adapted. Fifty milligrams was used per DNA isolation. MQ (150 μ l) and CTAB buffer (350 μ l; 20 g/l CTAB, 1.4 M NaCl, 0.1 M Tris-HCl, 20 mM EDTA) were added together

with 10 μ l 20 mg/ml Proteinase K and incubated overnight at 42 °C. After a short centrifugation (5 min at 14,000 rpm), the supernatant was incubated with 5 μ l RNaseA (Qiagen, 100 mg/ml) and incubated at 65 °C for 15 min. Of buffer AP2 (Qiagen DNeasy Plant Minikit), 260 μ l was added and was incubated on ice for 5 min. Further steps continued without deviation from step 4 of the Qiagen DNeasy Plant Minikit protocol (Qiagen, DNeasy Plant Mini and DNeasy Plant Maxi Handbook 01/2004). DNA concentrations were measured with the NanoDrop spectrophotometer (NanoDrop ND-1000, software version 3.3.0). DNA mixtures were prepared by combining equal volumes of 50 ng/ μ l stock solutions.

Ligation

A total of 200 ng DNA was used in a ligation assay: 1 \times *Pfu* ligation buffer (Stratagene); 12% (w/v) PEG6000 (Fluka, Germany); 0.1 U/ μ l *Pfu* ligase (Stratagene), 1.0 pM SpikeTarget and 25 pM of each PLP in a final volume of 10 μ l) to allow circularisation (94 °C for 5 min; 95 °C for 30 s, 65 °C for 5 min for 30 cycles) in the BioRad iCycler 3.021 [18].

LATE-PCR and microarray detection

Amplification by linear after the exponential (LATE)-PCR [19] of the circularised PLPs was performed as described by Prins et al. [17]. Denatured labelled PCR product (2 μ l) was applied to 63 μ l hybridisation mixture (5 \times SSC, 0.1% SDS, 0.1 mg herring sperm DNA, 12.5 nM Cy3-labelled ZB3 oligo (Table 1) as the positioning marker after hybridisation.

Slides with eight microarrays were custom-made by Eppendorf (Eppendorf Array Technologies Namur, Belgium) and contained 100 ZIP codes (20-mer oligonucleotide sequences from Affymetrix) with a 10-mer-A-tail and C6 to a linker (Biolegio) in quadruplicate per microarray. The ZIP codes were demarcated by ZB3 spots and ZIP-P spots (Table 1) for positioning purposes.

The microarray slides were pre-hybridized O/N at 42 °C in denatured pre-hybridisation mix (5 \times SSC, 0.1% SDS, 0.1 mg herring sperm DNA). After a repeated wash with 0.1 \times SSC for 5 min at room temperature (RT) on a rotary tablet, the slides were rinsed with MQ and dried by centrifugation (2 min at 1,000 rpm).

For hybridisation, an eight-microarray gasket in combination with a hybridisation chamber was used and hybridisation was carried out for 2 h at 65 °C in a hybridisation oven (all Agilent).

After hybridisation, the slides were washed twice with 1 \times SSC, 0.1% SDS for 5 min; twice with 0.1 \times SSC, 0.1% SDS for 5 min; twice with 0.1 \times SSC for 1 min; and once

Table 1 ssDNA sequences (5'-3' orientation) of PLPs

Name	5' target	cZIP sequence	3' target	Target gene (Genbank)	Reference
Spikelock	CGTCGGACAGGTTACTTTCGAAAGAGCCGGAATACTC	GCTGAGGTCGATGCTGAGGTCGCA	CGAAGGTCATATCTCG		[17]
Canola	GTGACGCATACGTTCTATAACATCAGCCGTGTCC	CGTCGGTTAGACAGCTCAT	CCGATCTTTCTTGTATTTC	DQ173668	[17]
Soy	GCATCATAGGTAATGAGAACTTGGCTACTTTTATTTGTTGGCC	ACTCCAGTGCCAAAGTACGAT	AGAGGCTGGTGGAG	K00821	[17]
Maize <i>zein</i>	CTGTGGCATCATCACTGGCATCGT	GTACTACATTCGTGCGATGG	TTAGGCCGTCATCAT	AF371266	[17]
Oat	CATCTAATAAAGTGGATTAATCGGACGAGGACAAAGAGAGATCCC	TGTGATAATTTCCGACGAGGC	TGAAAGGAAGGGCTTTATA	DQ131557	[18]
Barley	CGGTTTGGCTGCTCGTATTGGCTCACGACCTG	GTTGACCGTTAGTTATGCGA	CCTTGGATGGCACCT	CK125594	[18]
Wheat_D-genome	ATGGCTGGCTTCTAATTCATGATCCATCGTCTCTTCTG	ATTGACCAAACTGCGGTGCG	TTTTCAAAATATCCAGCTAG	A92706*	[18]
Wheat_Q-allele	AATAGAAAAATTTGCTGTTGGCCCAAGTTGTGTCGGTGG	CGTACCAATGGATGCGGTCT	GTGGTTAGATTTGCTCA	AY702956	This paper
Wheat_q-allele	AATAGAAAAATTTGCTGTTGGCCCAAGTTGTGTCGGTGG	GTCACGTATGGTTGCGTCT	GTGGTTAGATTTGCTCG	AY702957	This paper
Rye	CTCTGTCCGTGTCGTAGGAGGACTCGGTCC	TCCTTAGTTGTCGAGCAGG	TTGCCGATTCGATATGC	AF000227	This paper
Maize <i>HMG</i>	CACACAAAACGACCGGTAAACAAATTAATCAGCACGAG	CTGCGGTGTCAGTGATCTCT	GCCTTGTCTACAATC	AJ131373	This paper
Sugar beet	GAAAGTTAAAAGTATGTGCCGCTCTCAAGACTGAACAATGGC	GAGGTCAGTTACGAAAGCTC	GATCAAITTAAGTGTCTAC	BQ5864848	This paper
Generic rice	CCATTGCTGTCTCTGCAAGCTCACGGC	ATGCAGCGTAGGTATCGACT	CGGCAGCAACTCTCA	U12171	This paper
Non-fragrant rice	TTCAGCTGCTCCTAIGGTTAAGGTTTGTTCCAAATTTCTGIGG	ATTCTAAATCAATCGTCGCGG	CTGGTAA AAA GAAITATGGC	AP004463	This paper
Basmati rice_short	TTTCAGCTGCTCCTAIGGTTAAGGTTTGTTCCAA	CTGGAATGTGACCGTGCTCT	ATGAAACTGGTATATA	Bradbury et al. [10]	This paper
Basmati rice_short with LNA	TTTCAGCTGCTCCTAIGGTTAAGGTTTGTTCCAA	CTGGAATGTGACCGTGCTCT	ATGAAACTGGTATATA	Bradbury et al. [10]	This paper
Basmati rice_long	TTTCAGCTGCTCCTAIGGTTAAGGTTTGTTCCAA	CTGGAATGTGACCGTGCTCT	AGTTATGAAACTGGTATATA	Bradbury et al. [10]	This paper

Bold font represents an LNA nucleotide. Underlined font represents the deletion and single nuclear polymorphisms resulting in the difference in fragrance between non-fragrant and Basmati rice

Table 2 Plant materials

Latin name	Common name	Source
<i>Triticum dicoccum</i>	Emmer (Farro della Garfagnana)	Biolytix, Switzerland
<i>Triticum dicoccum</i>	Emmer	Biolytix, Switzerland
<i>Triticum monococcum</i>	Einkorn	Biolytix, Switzerland
<i>Triticum aestivum</i>	Common bread wheat	Biolytix, Switzerland
<i>Tritium spelta</i>	Spelt (Rotkorn)	Biolytix, Switzerland
<i>Triticum durum</i>	Durum	Biolytix, Switzerland
<i>Hordeum vulgare</i>	Barley	Biolytix, Switzerland
<i>Secale cereale</i>	Rye	Biolytix, Switzerland
<i>Avena sativa</i>	Oat	Biolytix, Switzerland
<i>Oryza sativa indica</i>	Non-fragrant rice	Indica rice (Xieqingzao)
<i>Oryza sativa</i>	Basmati rice	'Biological Basmati rice from India'
<i>Glycine max</i>	Soy	BF410D (3) 1% RRS
<i>Zea mays</i>	Maize	BF411A (2) 0% Bt176
<i>Brassica napus</i>	Canola	AOCS 0306G 100% RF3
<i>Beta vulgaris</i>	Sugar beet	BF419A 0% H7-1

with $0.01 \times$ SSC for 30 s, all at RT. Slides were dried by centrifugation (2 min at 1,000 rpm).

Slides were scanned using the ScanArray Express HT microarray scanner (Perkin Elmer) at 543 nm (Cy3). A PhotoMultiplier Tube gain of 60 was used at a laser power of 90%. The individual signals were quantified using the optical system software programme ArrayVision, version 8.0 (Imaging Research Inc.) The subtracted density (sDens) was used for further data analysis in Excel (Microsoft). sDens is defined as the density minus the background on a scale of 0–65,279 in arbitrary units (AU). Density is the average intensity of all pixels within the spots, and background is the average intensity of four corners between the spots. Each sample was ligated, amplified and hybridized in duplicate or triplicate. Each ZIP code was printed on the microarray quadruplicate.

Data analysis

Outliers caused by stains after washing were removed by hand. The average sDens values from all signals (8 or 12 per ZIP maximum) were calculated and used to calculate the standard errors of the mean.

Results

Design and quality control of PLPs

PLPs were designed for multiplex detection of the Q and q alleles in wheat, the *nfg*r allele in rice and four plant species: generic rice (all rice varieties), rye, sugar beet and maize (*HMG* gene). All newly designed PLPs were found to show a difference of at least seven threshold cycles (C_t s)

in q-PCR comparing its synthetic target to water (data not shown). Previously designed PLPs for the species oat, barley, maize (*zein* gene), soy and canola, and the wheat D genome, had already been shown to meet these criteria [17, 18].

Due to the high AT content of the Basmati rice discriminating sequence (*fgr* allele), another approach for PLP design had to be used. Target sites of the Basmati PLP were to be adjusted to meet T_m requirements by the insertion of LNA molecules. To test the effect of LNA molecules on the performance of a PLP, three different PLPs for Basmati rice were designed: one meeting the set T_m criteria and therefore with a long target site; one with a shorter target site and thus low T_m ; and this same PLP with LNA molecules increasing the T_m to meet the set criteria (Table 1). All three PLPs were found to show a difference in C_t value when its synthetic target was compared with water, although the delta C_t was 3 at best (data not shown). For comparison of the different target characteristics (length, T_m), probes were tested on microarray with a mixture of equal amounts on DNA mass basis of Basmati and non-fragrant rice in a background of genomic DNA (Table 2). As shown in Fig. 1, the PLP comprising the long target site showed also a signal ligation with non-fragrant rice. The two short-target PLPs did not show this signal. Lower signals on microarray were observed with the PLP comprising the LNA nucleotides than in the same PLP with the normal DNA nucleotides.

Specificity

In order to demonstrate the specificity of all plant-related PLPs, a mixture of 14 probes plus the SpikeLock (Table 1) was prepared and tested on a genomic DNA mixture

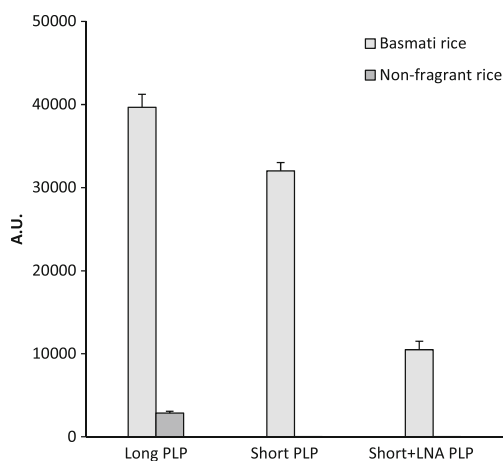


Fig. 1 Difference in performance of three distinct Basmati PLPs. *Long PLP* indicates the PLP meeting the set T_m criteria. *Short PLP* indicates the PLP with a shorter target sequence and thus a low T_m . *Short+LNA-PLP* indicates this same PLP with LNA nucleotides increasing the T_m . On the y-axis, signals are plotted in arbitrary units. Error bars indicate standard errors of the mean

comprising nine plant species/varieties: soy, maize, rye, barley, oat, canola, sugar beet, Basmati rice and non-fragrant rice at levels of 11.1% each. Specificity of the probes was determined in a series of experiments in which each species or variety was excluded once, resulting in levels of the remaining species of 12.5% each. All PLPs showed positive signals when their corresponding target was present in the mixture and from 2 to over 300 times lower when the target was left out (Fig. 2). The PLPs for the rice fragrance locus showed clear discrimination

between Basmati and non-fragrant rice, whereas the rice generic PLP detected both (data not shown).

For demonstration of the specificity of the wheat PLPs, the same PLP mixture was used. Six wheats (Table 2) were tested separately in a genomic DNA background of seven different species (soy, maize, rye, barley, oat, canola, sugar beet), giving a total of eight different DNA sources per sample, all at levels of 12.5% per ingredient. The specific ligation pattern was observed for all six wheats (Fig. 3). None of the wheat-specific PLPs showed a significant signal with any other target than its own (data not shown).

Sensitivity

The sensitivity of the wheat and rice PLPs was determined by comparing pure samples with samples with 5% admixture of a cheaper variety. In the case of Farro della Garfagnana, 5% common bread wheat was mixed, on DNA mass basis. In the case of Basmati rice, 5% non-fragrant rice was mixed, also on DNA mass basis. The same 15-plex PLP mix was used as in the previous paragraph. Furthermore, the experiments were performed with an extra replication. The results are depicted in Fig. 4.

Discussion

Currently, labelling of food and food products is a daily routine in compliance with the EU Food Law regulations

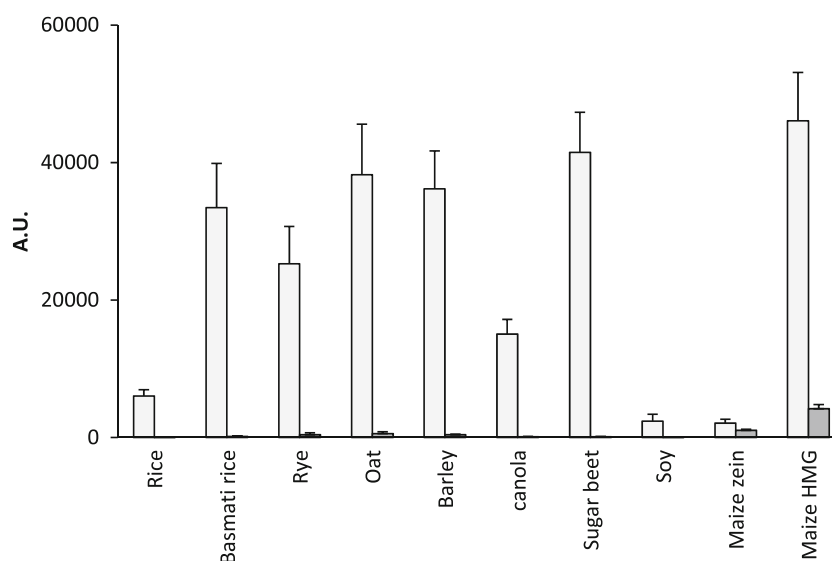


Fig. 2 Specificity of ten PLPs in complex DNA mixtures. Specificity was tested for 10 out of 15 PLPs in a series of ten duplicate experiments for nine different targets (two PLPs for maize were tested). In one experiment, all targets were present (*light grey bars*). In the other nine experiments, each target was left out once. The *dark*

grey bars represent the signals of each PLP when its corresponding target was not present. They are grouped with the signals for the same PLP in the ‘all present’ experiment for comparison. All other PLP signals in these experiments are not shown

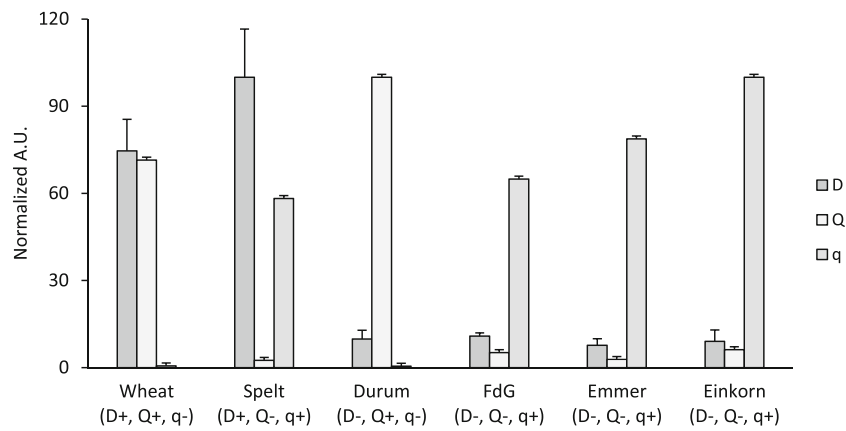


Fig. 3 Identification of six wheats through a combination of three PLPs. *D* D genome specific PLP, *Q* Q locus-specific PLP, *q* q locus-specific PLP. The three discriminating PLPs were present in a mixture of 14 PLPs. *FdG* Farro della Garfagnana. Signals were in different ranges for the different PLPs and have been rescaled in this graph for

visual reasons. For each PLP, separately, signals for all samples were divided by the highest observed value and multiplied by 100. The highest observed values were 9,670, 7,851 and 34,812 AU for the *D*, *Q* and *q* PLP. On the *y*-axis, the normalized signals are plotted in arbitrary units. *Error bars* indicate standard errors of the mean

concerning the traceability of the products and providing information to the consumer. However, maintenance of these regulations in today’s complex food matrices and global food market demand rapid, robust, accurate and cost-effective multiplex detection tools. Here, a tool is presented for multiplex detection and identification of raw materials as well as potentially unintended or fraudulent admixtures that may occur in complex food matrices. The PPLMD tool was tested and initially validated for the detection of different kinds of wheat and rice in mixtures comprising seven or nine other ingredients.

Specificity has been shown for 14 plant-related PLPs, of which nine were designed in the present study. Signal

differences ranged from 2 for the maize *zein* PLP to over 300 for the sugar beet PLP when 0% or 11.1% specific target was present in a background of genomic DNA. Discrimination of fragrant (Basmati) and non-fragrant rice was based on an eight-nucleotide deletion in the target site of fragrant rice compared with non-fragrant rice. Signals from these rice fragrance PLPs were compared in a situation where either fragrant or non-fragrant rice was present in the DNA mixture. The Basmati rice PLP showed a 200 times higher signal when Basmati rice was present compared with the signal when non-fragrant rice was present. In the case of the non-fragrant rice PLP, the difference in signal was almost 150 times. In both cases, the generic rice PLP showed an almost 14 times higher signal when rice DNA was present in the DNA mixture compared with no rice at all being present.

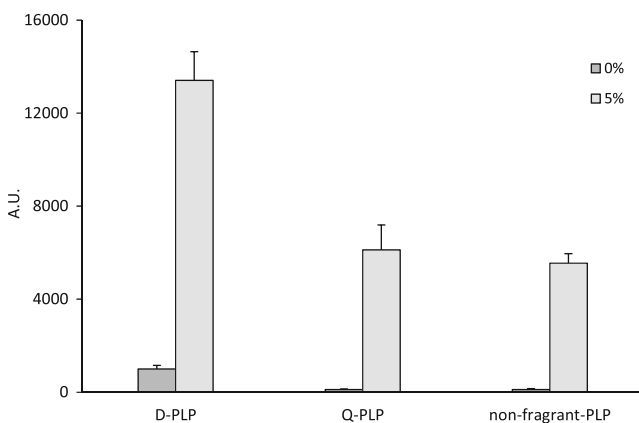


Fig. 4 Sensitivity of two wheats (*D* and *Q* PLP) and one non-fragrant rice discriminating PLP. *Light grey bars* indicate 5% common bread wheat or non-fragrant rice in a background of Farro della Garfagnana or Basmati rice, respectively. *Dark grey bars* indicate no admixture. On the *y*-axis, signals are plotted in arbitrary units. *Error bars* indicate standard errors of the mean

Five wheat species were easily distinguished using three wheat PLPs based on the detection of the *D* genome and on genotyping of the *Q* locus. For genotyping the *Q* locus, a single nucleotide polymorphism (SNP) on the ligation site was used. No false positive or negative results were observed when a SNP on the ligation site was used as the only difference between PLPs. The difference in signal between the presence and absence of the specific target was 9-fold for the *D*-PLP, 20-fold for the *Q*-PLP and almost 135-fold for the *q*-PLP, which showed again the great discriminatory power of a good PLP. The usefulness of PPLMD has also been shown for the detection of fraudulent/intentional admixtures. Since (very) low-level admixtures will not lead to a significant economic advantage and are therefore not expected, admixtures down to 5% have been tested. In both cases, the admixture of common bread wheat in Farro della Garfagnana or of non-fragrant

rice in Basmati rice could easily be detected. Unintentional admixture to specialty products is likely to occur because of postharvest operations such as transport and storage, but is expected to occur at much lower levels. In the case of wheat, samples with 5% common bread wheat in 95% Farro della Garfagnana showed approximately 13 times higher signals on microarray using the D genome PLP than in samples where no bread wheat was added (Fig. 4). For rice, the discriminating non-fragrant PLP showed an increase in signal of almost 50 times when 5% non-fragrant rice was added to Basmati rice. These results suggested that, if needed, the tool is likely to detect even lower levels of (un-)intentional adulteration.

It has furthermore been shown that a number of aspects are important for optimisation of the tool. An important aspect for the successful performance of the tool is the design of a PLP. Characteristics as target length, GC%, secondary structures such as hairpins and self or primer dimers influence the performance of a PLP. PLPs are designed meeting the criteria which were set on the basis of experimental data with regard to specificity and sensitivity [17, 18, 20]. Each PLP shows a specific background signal, which is always observed even when no target is present. The height of this background signal has a broad range. Signals range from very low (30 AU) for the rye PLP to high (2,000 AU) for the generic rice PLP. The PLP designed for maize *HMG* showed a very high background signal of up to 4,000 AU when no maize is present in the mixture, but this can be considered as an exception. It must be taken into account that the used cereal and rice materials were no certified reference materials. For this reason, false positive signals due to a minor contamination of the raw materials cannot be excluded.

Another indication for the influence of PLP design is the observation that different PLPs that were used to detect the same target resulted in very different signals. For the detection of endogenous maize, two PLPs for the sequences of *zein* and *HMG* were used. Both are supposed to be present only once in the maize genome. An approximately 22 times higher signal was observed when the *HMG* PLP was used instead of the *zein* PLP (Fig. 2). This implied that the height of the signal is dependent on the design of the PLP used.

For the AT-rich discriminating sequence of Basmati rice, a technical adaptation of the target sites of the PLP had to be made in order to meet the set criteria. LNA nucleotides were incorporated to keep the target site rather short, but with a T_m meeting our requirements. LNA nucleotides are a class of nucleic acid analogues in which the ribose ring is adjusted, resulting in a higher thermal stability [21, 22] of the molecules in which the LNA is incorporated. For Basmati rice, three different PLPs were designed: one with a long target site meeting the set T_m criteria, one with a

shorter target site and subsequent low T_m , and this same PLP with LNA substitutes increasing the T_m to meet the set criteria. The Basmati PLP with the long target site showed nonspecific ligation with non-fragrant rice. The two PLPs with the shorter target site did not show this nonspecific ligation. Surprisingly, however, microarray signals were lower using the LNA PLP compared with the same PLP without LNA substitutes (Fig. 1). Because of the increased thermal stability of LNA bases, the risk of stable secondary structures as self and primer dimers or hairpins increases as well [23]. These structures can compete with or inhibit the annealing of the PLP to the genomic DNA, resulting in reduced signals. This may explain the low signals of the Basmati rice PLP.

To further assess the sensitivity and specificity of each PLP, a validation study needs to be performed including more known positive and negative samples. For wheat, these should include other *Triticum* species [24] to check for cross-reactions with closely related wheat species. Mainly the domesticated variants as *Triticum turgidum* and *Triticum polonicum* are of importance. For rice, varieties from the *indica* (i.e. Jasmine varieties) and the *tropical japonica* groups [7] could be included as positive samples and some of the varieties found in South and Southeast Asia that do not carry the eight-nucleotide deletion [11, 12] as negative samples. For all the different rice species that are sold as a specialty product, it would be interesting to investigate whether there is a genetic basis for their special status. If the genetic basis for fragrance is the same as was used here, the same PLP can be used; if not, other PLPs should be designed applying the same principle. On the basis of the observed background signals, specific cutoff levels can be set per PLP. In all cases, the PPLMD tool can be used as a screening approach. If required, confirmation experiments as q-PCR can be performed for individual species/varieties that will also provide further quantitative data.

The main advantage of PPLMD is its multiplexing capacity compared with previously mentioned methods for food detection [6, 13–16]. One mixture of PLPs can be used to test simultaneously different ingredients within one food or feed sample. PPLMD has been shown to be a rapid, robust, accurate and cost-effective multiplex detection tool for the detection of unintended admixture in raw materials for food and feed production. The specificity of 14 plant-related PLPs has been shown in complex DNA mixtures. Five of these PLPs have been tested for the detection of low percentages of fraudulent admixture such as cheaper bread wheat or non-fragrant rice in varieties of superior quality. The results suggest that even lower admixtures, <5%, can easily be detected. This tool can thus be used to monitor the correct labelling of premium foods within Europe and abroad.

Acknowledgements The authors acknowledge financial support concerning scientific activities within the Sixth Framework Programme of the European Union, project TRACE—‘TRacing food Commodities in Europe’ (project no. FOOD-CT-2005-006942). Also, support from the Dutch Ministry of Agriculture, Nature and Food Quality is gratefully acknowledged. The publication reflects only the authors’ views and the Community is not liable for any use that may be made of the information contained herein.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. European Commission. General food law. Traceability factsheet. ISBN 978-92-79-05725-0
2. European Commission. Agriculture and rural development. Quality policy. <http://ec.europa.eu>
3. European Commission. Agriculture and rural development. DOOR—Database of Origin and Registration. IT_0017_1528_SUM_EN[1].pdf
4. Pestsova E, Ganal MW, Roder MS (2000) Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome* 43(4):689–697
5. Muramatsu M (1963) Dosage effect of the spelta gene q of hexaploid wheat. *Genetics* 48:469–482
6. Asakura N, Mori N, Nakamura C, Ohtsuka I (2009) Genotyping of the Q locus in wheat by a simple PCR-RFLP method. *Genes and Genetic Systems* 84(3):233–237
7. Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch SR (2009) The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc Natl Acad Sci U S A* 106(34):14444–14449
8. Buttery RG, Ling LC, Juliano BO, Turnbaugh JG (1983) Cooked rice aroma and 2-acetyl-1-pyrroline. *J Agric Food Chem* 31(4):823–826
9. Widjaja R, Craske JD, Wootton M (1996) Comparative studies on volatile components of non-fragrant and fragrant rices. *J Sci Food Agric* 70(2):151–161
10. Bradbury LMT, Fitzgerald TL, Henry RJ, Jin Q, Waters DLE (2005) The gene for fragrance in rice. *Plant Biotechnology Journal* 3(3):363–370
11. Fitzgerald MA, Sackville Hamilton NR, Calingacion MN, Verhoeven HA, Butardo VM (2008) Is there a second fragrance gene in rice? *Plant Biotechnology Journal* 6(4):416–423
12. Sakthivel K, Sundaram RM, Shobha Rani N, Balachandran SM, Neeraja CN (2009) Genetic and molecular basis of fragrance in rice. *Biotechnol Adv* 27(4):468–473
13. Lopez SJ (2008) TaqMan based real time PCR method for quantitative detection of Basmati rice adulteration with non-Basmati rice. *Eur Food Res Technol* 227:619–622
14. Pasqualone A, Montemurro C, Grinn-Gofron A, Sonnante G, Blanco A (2007) Detection of soft wheat in semolina and durum wheat bread by analysis of DNA microsatellites. *J Agric Food Chem* 55(9):3312–3318
15. Sonnante G, Montemurro C, Morgese A, Sabetta W, Blanco A, Pasqualone A (2009) DNA microsatellite region for a reliable quantification of soft wheat adulteration in durum wheat-based foodstuffs by real-time PCR. *J Agric Food Chem* 57:10199–10204
16. Vemireddy LR, Archak S, Nagaraju J (2007) Capillary electrophoresis is essential for microsatellite marker based detection and quantification of adulteration of Basmati rice (*Oryza sativa*). *J Agric Food Chem* 55(20):8112–8117
17. Prins TW, Van Dijk JP, Beenen HG, Van Hoef AMA, Voorhuijzen MM, Schoen CD, Aarts HJM, Kok EJ (2008) Optimised padlock probe ligation and microarray detection of multiple (non-authorized) GMOs in a single reaction. *BMC Genomics* 9:854
18. Prins TW, Van Dijk JP, Van Hoef AMA, Voorhuijzen MM, Broeders S, Trapmann S, Seyfarth R, Pardigol A, Schoen CD, Aarts HJM, Kok EJ (2010) Towards a multiplex cereal traceability tool using padlock probe ligation on genomic DNA. *Food Chem* 118(4):966–973
19. Sanchez JA, Pierce KE, Rice JE, Wangh LJ (2003) Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis. *PNAS* 101(7):1933–1938
20. Szemes M, Bonants P, De Weerd M, Baner J, Landegren U, Schoen CD (2005) Diagnostic application of padlock probes—multiplex detection of plant pathogens using universal microarrays. *Nucleic Acids Res* 33(8):e70
21. Koshkin AA, Singh SK, Nielsen P, Rajwanski VK, Kumar R, Meldgaard M, Olsen CE, Wengel J (1998) LNA (locked nucleic acids): synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* 54:3607–3630
22. Wengel J. Locked Nucleic Acid Technology™: a brief overview. <http://www.exiqon.com>
23. Devor E (2005) Locked nucleic acids (LNAs). Integrated DNA technologies. <http://eu.idtdna.com>
24. Magness JR, Markle GM, Compton CC (1971) Food and feed crops of the United States. Interregional Research Project IR-4, IR Bul. 1 (Bul. 828 New Jersey Agr. Expt. Sta)