JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Development of a Locked Nucleic Acid Real-Time Polymerase Chain Reaction Assay for the Detection of *Pinus armandii* in Mixed Species Pine Nut Samples Associated with Dysgeusia

Sara M. Handy,* Ruth E. Timme, Salena M. Jacob, and Jonathan R. Deeds

Center for Food Safety and Applied Nutrition, Office of Regulatory Science, United States Food and Drug Administration (U.S. FDA), College Park, Maryland 20740, United States

ABSTRACT: Recent work has shown that the presence of the species *Pinus armandii*, even when occurring as species mixtures of pine nuts, is correlated with taste disturbance (dysgeusia), also referred to as "pine mouth". Because of this known possibility of pine nut mixtures, a need was identified for a rapid streamlined assay to detect the presence of this species in the presence of other types of pine nuts. A locked nucleic acid probe was employed in a real-time polymerase chain reaction (RT-PCR) format to detect a single nucleotide polymorphism (SNP) unique to this species. This assay was able to detect *P. armandii* in homogenates down to ~1% concentration (the lowest level tested) in the presence of several commonly co-occurring and closely related species of pine and should prove to be a useful tool for the detection of this species in food products.

KEYWORDS: *Locked nucleic acid, RT-PCR, pine nut, pine mouth, Pinus armandii, dysgeusia*

INTRODUCTION

From July 2008 to June 2012, the United States Food and Drug Administration (U.S. FDA) received 501 complaints from consumers who reported dysgeusia consistent with a clinical entity referred to as "pine mouth syndrome".[1] This syndrome consists of a delayed (hours to days), persistent (up to several weeks) bitter or metallic aftertaste following consumption of pine nuts. Recent work correlated the presence of the pine species *Pinus armandii* with these taste disturbances based on fatty acid analysis,[2] nuclear magnetic resonance (NMR) spectroscopy,[3] deoxyribonucleic acid (DNA) sequencing,[4] and flow cytometry.[5] This species has been prohibited for use in foods by the European Union. As of April 13, 2011, *P. armandii* can no longer be exported from China to Europe (http://www.nutfruit.org/en/chinese-pine-nuts-bitter-after-taste-update_36923, accessed Aug 23, 2012). Additionally, *P. armandii* has not been included on the Food and Agricultural Organization (FAO) of the United Nations list of edible pine species and has been further deemed by the Codex Committee on Pesticide Residues as being unfit for food (http://www.codexalimentarius.net/web/archives.jsp?lang=en; paragraph 88 of Alinorm REP11/PR, Report of the 43rd session, accessed Aug 23, 2012).

To confirm the association between *P. armandii* and U.S. FDA consumer complaints of taste disturbances, a genetic method was developed to identify various pine nut species.[4] However, this method requires a laborious screening of each individual nut (DNA extraction, amplification, and DNA sequencing), which can be tedious when *P. armandii* is present as mixtures with other pine nut species. A total of 11 of 15 of the consumer complaint samples analyzed previously by the U.S. FDA contained mixtures of *P. armandii* and other pine nut species, with the remaining 4 samples being 100% *P. armandii* of the nuts tested.[4,6] The goal of the research described in this paper was to develop a more streamlined method that could be used with pine nut species mixtures and would give a simple presence/absence answer for *P. armandii* in a rapid assay.

Because of the closely related genetic nature of the edible pine species, a real-time polymerase chain reaction (RT-PCR) method using a locked nucleic acid (LNA) probe was proposed. LNA probes for RT- or quantitative (Q)-PCR have select bases replaced by LNA monomers (2′-*O*,4′-*C* methylene bridge), which enhances the performance of the probe.[7] LNA probes have recently been used to differentiate between the same species (a wild-type and a mutant) with a single base mutation[8] and have also been employed in food-safety-based assays.[7] RT-PCR has been widely used to recognize unwanted materials or contaminants in foods, including pathogens,[9,10] allergens,[11−13] and other unexpected products.[14−16]

Designing RT-PCR assays for the identification of closely related species of plants is challenging because there is often little or no differentiation in the gene regions typically used to identify plants.[17] The gene previously used for the sequencing-based identification of *P. armandii*[4] (*ycf*1) could not be used for a RT-PCR approach. Although *ycf*1 is useful in differentiating *Pinus* species throughout its length, there is no single position unique to *P. armandii* and no other closely related species (data not shown). This paper describes our efforts to identify an effective region using the full chloroplast genome from a large alignment of *Pinus* species,[18] allowing for the detection of *P. armandii* in a rapid RT-PCR assay that could be used to analyze pure or mixed samples of pine nuts.

MATERIALS AND METHODS

Gene Selection. Potential probe sites were investigated using a subset of *Pinus* plastid genomes (Table 1) for species that produce

Received: October 4, 2012
Revised: January 8, 2013
Accepted: January 11, 2013
Published: January 11, 2013

dx.doi.org/10.1021/jf304223a | *J. Agric. Food Chem.* 2013, 61, 1060−1066

Table 1. References for the 33 Plastid Genomes Used in the kSNP Analysis

species	source	GenBank ID
Pinus albicaulis Engelm.	Parks et al.[23]	FJ899566
Pinus armandii Franch.	Parks et al.[23]	FJ899568
Pinus ayacahuite Ehrenb. ex Schltdl.	Parks et al.[23]	FJ899570
Pinus cembra L.	Parks et al.[23]	FJ899574
Pinus cembroides Zucc.	Parks et al.[18]	JN854220
Pinus coulteri D. Don	Parks et al.[18]	JN854215
Pinus culminicola Andresen and Beaman	Parks et al.[18]	JN854213
Pinus discolor D. K. Bailey and F. G. Hawksworth	Parks et al.[18]	JN854207
Pinus edulis Engelm.	Parks et al.[18]	JN854203
Pinus flexilis E. James	Parks et al.[23]	FJ899576
Pinus gerardiana Wall. ex D. Don	Cronn et al.[24]	EU998741
Pinus johannis M.-F. Robert	Parks et al.[18]	JN854192
Pinus koraiensis Siebold and Zucc.	unpublished	NC_004677.2
Pinus lambertiana Dougl.	Parks et al.[23]	FJ899577
Pinus massoniana Lamb.	Parks et al.[18]	JN854185
Pinus maximartinezii Rzedowski	Parks et al.[18]	JN854184
Pinus monophylla Torr. and Frém.	Cronn et al.[24]	EU998745
Pinus monticola Dougl. ex D. Don	Parks et al.[23]	FJ899580
Pinus nelsonii Shaw	Cronn et al.[24]	EU998746
Pinus pinceana Gord.	Parks et al.[18]	JN854174
Pinus pinea L.	Parks et al.[18]	JN854173
Pinus ponderosa Doug. ex C. Lawson var. *benthamiana* (Hartweg) Silba	Parks et al.[18]	JN854172
Pinus pumila (Pall.) Regel	Parks et al.[18]	JN854168
Pinus quadrifolia Parl. ex G. B. Sudworth	Parks et al.[18]	JN854166
Pinus remota (Little) D. K. Bailey and F. G. Hawksworth	Parks et al.[18]	JN854164
Pinus roxburghii Sargent	Parks et al.[18]	JN854162
Pinus sabiniana Dougl. ex D. Don	Parks et al.[18]	JN854161
Pinus sibirica Du Tour	Parks et al.[23]	FJ899558
Pinus strobiformis Engelm.	Parks et al.[23]	JN854159
Pinus tabuliformis Carr.	Parks	unpublished
Pinus torreyana ssp. *torreyana* Parry ex Carriére	Parks et al.[23]	FJ899564
Pinus wallichiana A. B. Jackson	Parks et al.[18]	JN854154
Pinus yunnanensis Franch	Parks et al.[18]	JN854151

edible nuts. This was a subset of a larger data set alignment provided by M. Parks, R. Cronn, and A. Liston (pre-publication).[18] Locations of candidate single nucleotide polymorphisms (SNPs) unique to *P. armandii* were identified using the program kSNP v1.0,[19] with a *k*-mer size of 25. Three of the SNPs discovered *in silico* to be specific to *P. armandii* were targeted, and primers were designed (Geneious Pro[20]) to target each SNP with ∼75 base pairs of the flanking sequence on either side (total of ∼150 base pairs for the expected amplicon length).

These three primer sets (8.6kF/R, 27kF/R, and 129kF/R), purchased from Integrated DNA Technologies (Coralville, IA) were investigated further for their utility in identifying *P. armandii* from the other types of pine species. A subset of pine specimens used by Handy et al.,[4] including *P. armandii* ARMA 09, *P. armandii* A140, *P. armandii* (LN), *Pinus koraiensis* (MA71577), and *Pinus monophylla* (PP), were amplified with all three primer sets. Full sample preparation and DNA extraction methods for these samples can be found in the study by Handy et al.[4] Briefly, sterile scalpel blades were used to cut into each pine nut, and a small segment of megametophyte (tissue inside the seed coat not including the embryo) from the inside of the seed (∼10 mg) was added to a sterile 1.5 mL microcentrifuge tube. DNA was extracted from tissue using a DNeasy Blood and Tissue kit (with reagent volumes reduced as in the study by Handy et al.[21]). A negative control was included with each set of extractions. An additional step was added for incubation of the washed filters and elution buffer at 37 °C for 30 min to increase successful elution of DNA. Extracted DNA was used directly in the PCR with no dilution.

The PCR cocktail consisted of 6.25 μL of 10% trehalose, 2 μL of molecular-grade water, 1.25 μL of 10× PCR reaction buffer (Invitrogen), 0.625 μL of 50 mM $MgCl_2$, 0.125 μL of each primer (10 μM), 0.062 μL of 10 mM deoxyribonucleotide triphosphates (dNTPs; New England Biolabs), and 0.06 μL of Platinum Taq (Invitrogen), with 1 μL of DNA template, for a total of 11.5 μL. Thermocycling conditions for all primer sets were 94 °C for 2 min, 30 cycles of 94 °C for 15 s, 55 °C for 30 s, and 72 °C for 1 min, with a final extension at 72 °C for 1 min.

To verify amplification, a selection of products were analyzed using precast 4% E-gel agarose gels according to the protocols of the manufacturer with the E-Base Integrated power supply. Gels were run for 15 min at 60−70 V (constant voltage) and then visualized using a Gel Doc 2000 gel documentation system. Amplified products were purified by adding 2 μL of Exosap-IT to 5 μL of PCR product and incubating at 37 °C for 15 min, followed by 15 min at 80 °C, and then sequencing.

Each sequencing reaction contained 0.25 μL of BigDye Terminator v3.1, 1.875 μL of 5× sequencing buffer, 5 μL of 10% trehalose, 1 μL of 10 μM primer, and 0.875 μL of molecular-grade water, for a total of 9 μL, to which 1 μL of purified PCR product was added. Sequencing products were purified as described by Handy et al.[21] using an EdgeBio short well plate and sequenced bidirectionally on an ABI 3730 instrument.

Sequenced products were examined in Geneious Pro. The sequences were assembled and trimmed for quality (error probability limit of 0.05 from both ends), assembled into contigs (default settings), converted to consensus sequences, and aligned using the "muscle alignment" tab,[22] with default settings.

Examination of Authenticated Pine Specimens. Genomic DNA from all 27 pine specimens from Handy et al.[4] (Table 2) was amplified with primer set 129k (129kF, 5′-TTC CAC CAT GTC AAG GTG AC-3′; 129kR, 5′-CTG TCT TCA AGT TGT TCG AGA-3′). These products were then sequenced as described above.

Development of a LNA Probe. A LNA RT-PCR probe (ParmaLNA) was developed targeting a *P. armandii* SNP sequence amplified by primers 129kF/R in Geneious Pro. The probe was 24 bases long with carboxyfluorescein (a 6-FAM) label on the 5′ end and a black hole quencher (3BHQ_1) on the 3′ end. It also contained 6 LNA bases, indicated in the sequence by "+". The full sequence of ParmaLNA is 5-6-FAM-TA+T+G+A+A+T+AAGT-CAGTTCTCCTTTC-3BHQ_1, with the LNA bases indicated by the "+" signs. This probe was compared to the larger 113 genome alignment and the public database Genbank (using BLAST; http://blast.ncbi.nlm.nih.gov) to determine *in silico* specificity. To reduce or eliminate non-specific binding of the ParmaLNA probe to closely related non-*P. armandii* targets, a blocker probe was designed to match the most closely related species, including *Pinus gerardiana*, *Pinus lambertiana*, *Pinus griffithi* (*wallichiana*), *Pinus koraiensis*, *Pinus pumila*, and *Pinus wallichiana*. The NonParmaLNA probe was also 24 bases and contained 6 LNA bases in addition to a hexachlorofluorescein (HEX) label on the 5′ end and a black hole quencher on the 3′ end. The full sequence of NonParmaLNA was 5-HEX-TA+T+T+A+A+T+AAGTCAGTTCTCCTTTC-3BHQ_1. These reporter dyes (HEX and FAM) can be spectrally resolved from each other because they emit at different wavelengths. Both probes were purchased from Integrated DNA Technologies, using their LNA PrimeTime technology.

Optimization of the LNA Assay. Primers and probe (ParmaLNA) concentrations were optimized for the Mx3005P QPCR system (Agilent Technologies, Santa Clara, CA) RT-PCR instrument using the PCR product (cleaned with Exosap-IT according to the instructions of the manufacturer) from *P. armandii* A140 diluted to 0.1 ng/μL. Optimized conditions for the assay were 12.5 μL reactions with 6.25 μL of Brilliant II Master Mix (Agilent Technologies, Santa Clara, CA), 0.186 μL of a 1:500 dilution of provided reference dye, ROX (1 mM), 0.313 μL of water, 2 μL of diluted template, and the following concentrations of primers and

Table 2. Pine Specimens, with More Information Available in Table 1 of the Study by Handy et al.[4]

species and authority	from	ParmaLNA probe
Pinus armandii Franch	Oregon State University	+
Pinus armandii Franch	International Nut and Dried Fruit Foundation 288	+
Pinus armandii Franch	F. W. Schumacher, U.S.A.	+
Pinus armandii Franch	Lawyer Nursery, Inc., U.S.A.	+
Pinus cembra var. *sibirica* Du Tour	Lawyer Nursery, Inc., U.S.A.	−
Pinus cembra L.	Lawyer Nursery, Inc., U.S.A.	−
Pinus edulis Engelm.	Lawyer Nursery, Inc., U.S.A.	−
Pinus gerardiana Wall. ex D. Don	Lawyer Nursery, Inc., U.S.A.	−
Pinus griffithi syn. *wallichiana* A. B. Jacks	F. W. Schumacher, U.S.A.	−
Pinus kochiana Klotzsch ex K. Koch (potentially a synonym of *Pinus sylvestris* var. *hamata* Steven)	U.S. Department of Agriculture, Agriculture Research Service	−
Pinus koraiensis Siebold et Zucc.	U.S. Department of Agriculture, Agriculture Research Service	−
Pinus koraiensis Siebold et Zucc.	International Nut and Dried Fruit Foundation 288	−
Pinus koraiensis Siebold et Zucc.	F. W. Schumacher, U.S.A.	−
Pinus koraiensis Siebold et Zucc.	Lawyer Nursery, Inc., U.S.A.	−
Pinus lambertiana Douglas	Lawyer Nursery, Inc., U.S.A.	−
Pinus lambertiana Douglas	Pinyon Penny's, U.S.A.	−
Pinus lambertiana Douglas	U.S. Department of Agriculture, Agriculture Research Service	−
Pinus massoniana Lamb.	Oregon State University	−
Pinus monophylla Torr. et Frém.	Pinyon Penny's, U.S.A.	−
Pinus pinea L.	Sardinia/Italy	−
Pinus pumila (Pall.) Regel	F. W. Schumacher, U.S.A.	−
Pinus pumila (Pall.) Regel	Lawyer Nursery, Inc., U.S.A.	−
Pinus sibirica Du Tour	International Nut and Dried Fruit Foundation 288	−
Pinus tabulaeformis Carriére	F. W. Schumacher, U.S.A.	−
Pinus taeda L.	U.S. Department of Agriculture, Agriculture Research Service	−
Pinus wallichiana A. B. Jacks	Lawyer Nursery, Inc., U.S.A.	−
Pinus yunnanensis Franch	F. W. Schumacher, U.S.A.	−

probes: 1.25 μL of 9 μM for 129kF, 1.25 μL of 3 μM for 129kR, and 1.25 μL of 2 μM each of the ParmaLNA probe and NonParmaLNA blocker. The cycling parameters for all reactions were 10 min at 95 °C for activation of the polymerase, followed by 40 cycles of 15 s at 95 °C, 30 s at 54 °C, and 2 min at 60 °C. RT-PCR data were analyzed using MxPro software (Agilent Technologies, Santa Clara, CA). Cycle threshold (Ct) values were determined from amplification curves, by setting the Ct line at a normalized relative fluorescence (ΔRn) at 1.0 for FAM and 0.1 for HEX. In all cases, a cut-off value of Ct < 30 (for the ParmaLNA probe) was used to evaluate whether a sample was positive for *P. armandii*.

LNA Probe Assay of Pine Specimens and Consumer Complaint Samples. A total of 27 authenticated pine specimens from Handy et al.[4] (Table 2) were tested using the optimized LNA assay. Each sample was diluted to 25 ng/μL or less (in cases where the DNA of the original extraction was not high enough) and run using the conditions listed above. With all runs, cleaned PCR product (as above) from sample *P. armandii* A 140 and *P. koraiensis* NA 71577 were used as controls (0.1 ng/μL) to confirm that both LNA probes were functioning properly (i.e., not giving false positives or false negatives). Additionally, a negative control was run with no DNA template added.

DNA from 105 nuts (Table 3), used previously to diagnose 15 consumer complaint samples from the original examination,[4] was used to further test the LNA assay. As noted in the previous study, 5 or more nuts from each consumer complaint sample were extracted depending upon visual sample homogeneity.[4] Each sample was diluted to 25 ng/μL if possible (some starting concentrations of DNA were <25 ng/μL).

LNA Probe Assay of Spiked Pine Nuts. A total of 5 g (30−40 nuts, depending upon the species) of *P. armandii* A140, *P. koraiensis* K116, and *Pinus sibirica* C120 were ground separately into a paste using a sterile closed ultra tissue grinder system (Fisher Healthcare), to which 10 mL of DNA/RNA free water was added. The slurries were vigorously vortexed to homogenize. A total of 100 μL of ground pine nut homogenate was extracted for each of the following samples: 100% of either *P. koraiensis*, *P. sibirica*, or *P. armandii* and both *P. koraiensis* and *P. sibirica* spiked with either 20, 10, 5, or 1% wet slurry weight of *P. armandii*, for a total of 11 samples (100 μL in total, diluted accordingly, i.e., 80 μL of *P. sibirica* with 20 μL of *P. armandii* for a 20% *P. armandii* spike, etc.). Extractions were performed using the Qiagen Dneasy plant extraction kit (according to the instructions of the manufacturer). DNA was quantified using a Nanodrop ND 1000 spectrophotometer (ThermoScientific, Wilmington, DE) and diluted to 25 ng/μL. The 11 extracted samples of either pure pine nuts or pine nut mixtures were analyzed using the LNA RT-PCR pine nut assay with the same reagent recipe and thermocycling conditions as mentioned above.

Table 3. Summary of LNA RT-PCR Pine Nut Assay Results for the U.S. FDA Consumer Complaint Samples

nuts amplified (105)	identification based on sequencing (Handy et al.[4])	number of nuts positive for *P. armandii* DNA based on LNA probe	notes
63	*Pinus armandii*	62	1Ct > 30[a]
15	*Pinus cembra/sibirica*	0	
1	*Pinus pumila*	0	
26	*Pinus koraiensis*	0	2Ct > 30[b]

[a]DNA was at a low concentration. [b]Potentially indicates low levels of contamination from *P. armandii*, which was known to be present in these samples.

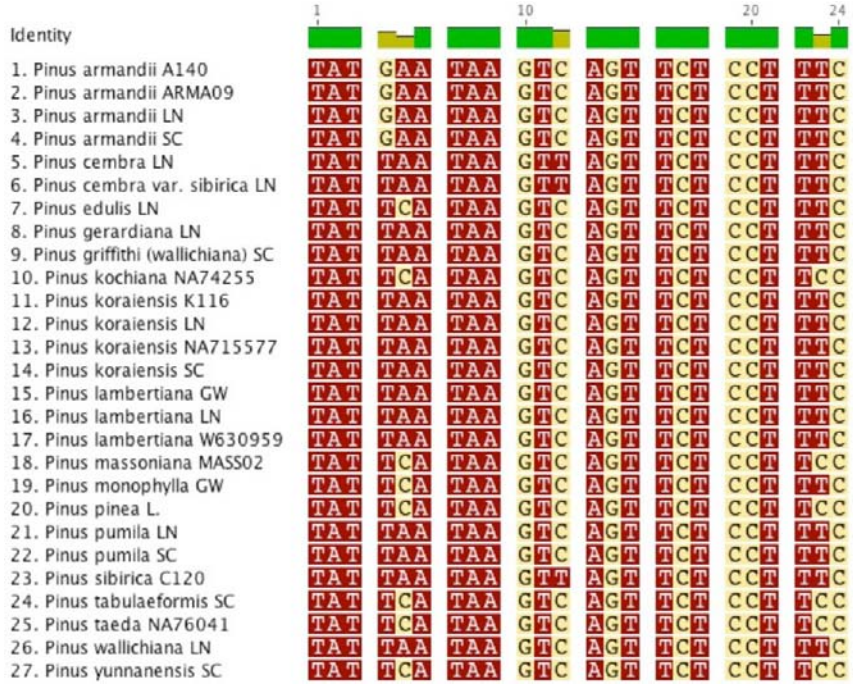

Figure 1. Geneious Pro alignment of sequences of the LNA probe region of the pine specimens (as in the study by Handy et al.[4]). Many of the specimens have several variable regions compared to the four *P. armandii* representatives, and dark regions indicate adenine (A) or thymine (T) residues, while light regions indicate cytosine (C) or guanine (G) residues. *P. armandii* is unique from all of the pine samples in base position 4. ParmaLNA uses this region.

RESULTS AND DISCUSSION

Gene Selection. All three primer sets produced amplicons from the five samples [*P. armandii* ARMA 09, *P. armandii* A140, *P. armandii* (LN), *P. koraiensis* (MA71577), and *P. monophylla* (PP)], but only one primer set (129k) amplified a canonical SNP unique to *P. armandii* (Figure 1). The SNPs in the other two primer sets, 8.6k and 27k, were not consistent across our three *P. armandii* samples (data not shown). Primer set 129k, which targets the chloroplast ATP synthase CF1 gene (α subunit), was selected for the rest of the experiments. Additionally, it was determined that no SNP pairs were located in proximity to each other to facilitate a less expensive PCR method, and it is for this reason that we chose the RT-PCR assay approach.

Examination of Authenticated Pine Specimens. All 27 pine specimens were amplified successfully by primer set 129kF/R (data not shown). When all 27 pines were sequenced and examined, only the representatives of *P. armandii* had a guanine residue at the fourth base position (Figure 1), while the rest of the species had a thymine residue.

Development of a LNA Probe. The ParmaLNA probe was designed *in silico* to be specific only to *P. armandii*.[18] In addition, a comparison by BLAST to the public database Genbank only returned *P. armandii* species. This discrimination power along with the increased affinity of LNA probes for their targets provided confidence that the ParmaLNA probe would specifically target only *P. armandii*. LNA bases were added around the SNP site to increase the melting temperature and decrease non-specific binding.

LNA Probe Assay of Pine Specimens and Consumer Complaint Samples. Of the 27 pine specimens examined, only the 4 *P. armandii* samples produced a positive result [indicated by a Ct value of <30 for the ParmaLNA probe when the threshold was set at a ΔRn of 1 (Table 2)]. Our analysis supported these findings because the probe was able to differentiate *P. armandii* from closely related pine species, including *P. koraiensis*, which was previously found to be mixed with *P. armandii*.[4] None of the other pine specimens gave Ct values (below the maximum of 40 cycles tested) using the ParmaLNA probe, indicating a lack of non-specific binding. The NonParmaLNA probe merely functions as a blocker, and it targets many species; it will not be discussed further besides noting that a control was always included for both *P. armandii* and *P. koraiensis* to confirm that the probes were working properly. Figure 2 illustrates the use of the control samples *P. armandii* and *P. koraiensis*, showing that the two probes are specific for their targets. In all cases, the positive controls had Ct < 30 for their respective probes and the negative controls were always negative (no Ct value below 40 cycles tested). Again, the two probes can always be differentiated from each other even in mixtures because they emit at different wavelengths and are compared on different scales (because of differences in brightness).

Of the original DNA extractions of 116 nuts from consumer complaint samples from Handy et al.,[4] 11 samples had no DNA detectable, as determined by the Nanodrop, presumably as the result of either degradation or DNA binding to the storage tube walls. This left 105 nut DNA extracts, consisting of 63 identified as *P. armandii* and 42 identified as non-*P. armandii*. Of these 105 nut DNA extracts, 102 yielded expected answers: Ct values of <30 for all *P. armandii* samples and no measurable Ct value for non-*P. armandii* (Table 3). In one sample, identified as *P. armandii*, the Ct value was higher than 30. In this case, the starting DNA concentration was lower than the ideal of 25 ng/μL (11.2 ng/μL). Two samples that were identified genetically as *P. koraiensis* in the study by Handy et al.[4] had ParmaLNA Ct values of 33 and 39. We speculate that this occurred in both cases because they were isolated from

dx.doi.org/10.1021/jf304223a | J. Agric. Food Chem. 2013, 61, 1060−1066

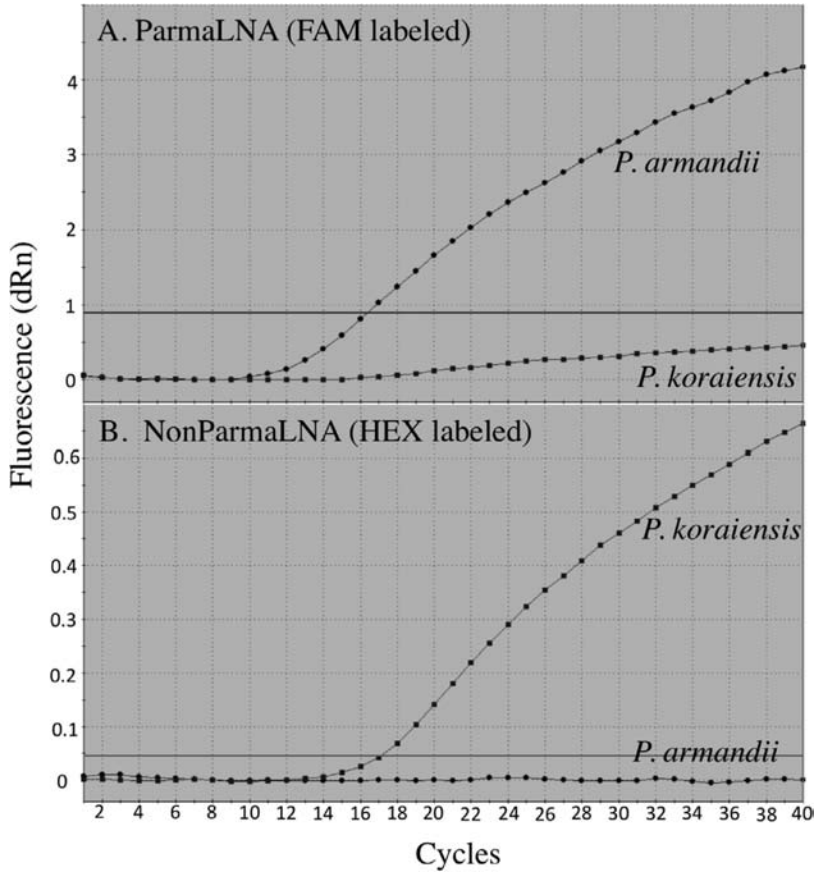

Figure 2. Demonstration of the *P. armandii*-specific probe ParmaLNA mixed with the non-specific blocker probe, NonParmaLNA. In each panel, two samples are being shown, one with *P. armandii* and another with *P. koraiensis* (cleaned PCR product at 0.1 ng/μL; thus, 0.2 ng in the final reaction used as the template). The panels show FAM (ParmaLNA) channel in view A and one with the HEX (NonParmaLNA) channel in view B. Panel A (FAM) shows *P. armandii* amplification, while *P. koraiensis* is not amplified. In panel B (HEX), *P. koraiensis* is amplified and *P. armandii* is not. Panel B indicates that *P. koraiensis* is amplifiable and *P. armandii* is not targeted by the NonParmaLNA probe. Note that the two reporter dyes (HEX and FAM) have different emission wavelengths and different intensities (FAM is brighter), which is why they are shown on separate panels.

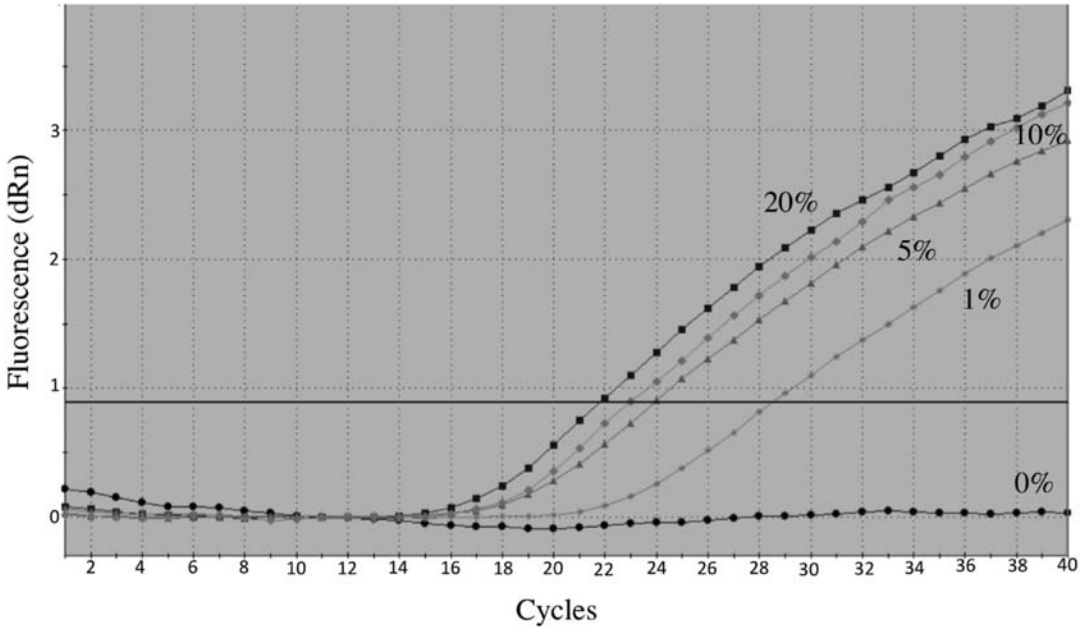

Figure 3. ParmaLNA probe fluorescence (FAM) of assays generated from DNA extracted from ground *P. koraiensis* pine nut samples spiked with noted percentages of ground *P. armandii* pine nut samples. Note that *P. armandii* can be detected down to 1%. The 100% *P. koraiensis* is indicated by 0% *P. armandii*.

mixed samples to start with and some trace amount of *P. armandii* DNA could potentially still be found in the sample.

We used these 132 pine nut samples (105 from the consumer complaint samples and 27 authenticated pine specimen) to verify that the LNA RT-PCR assay would yield the same answers as the sequencing-based method used in the study by Handy et al.[4] The assay supports the data presented in that study, but the current goal was ultimately to be able to identify this unwanted species in a mixture of nuts (using a greater number of nuts at a time).

LNA Probe Assay of Spiked Pine Nuts. To confirm that the assay was applicable to mixed samples, we conducted spiking experiments with varying levels of contaminating *P. armandii* in the presence of other pine nut species. A slurry of ground *P. armandii* nuts and water was added to a background of another, closely related species, and several different percentages of spiked samples were extracted. Specifically, *P. armandii* (A140) was spiked into the two species that it was most often found to be mixed with in previous U.S. FDA consumer complaint samples, *P. koraiensis* (K116) and *P. sibirica* (C120).[4] The 100% *P. armandii* sample gave a positive (<30) Ct value, while both 100% *P. koraiensis* (curve labeled 0% in Figure 3) and *P. sibirica* (data not shown) were negative (no CT values of <40 cycles) with the ParmaLNA probe. Spikes of *P. armandii* from 20 to 1% (ground wet volume) were all measurable with Ct values of <30 (spikes of 1−20% in *P. koraiensis* in Figure 3).

The LNA RT-PCR assay for the detection of *P. armandii* described here complements the DNA sequencing method using the chloroplast gene *ycf*1 previously described in the study by Handy et al.[4] While the previously described DNA sequencing method does allow for a more complete confirmatory analysis of the species of pine nuts present in a given sample, the advantages of the RT-PCR assay are a reduction in time (from days to hours) and costs (when considering a 96-well plate of samples, sequencing is 37 times more costly per sample). Most importantly, RT-PCR works on the mixed samples (without sorting nuts based on morphology) known to be a large proportion (73%) of the consumer complaint samples analyzed by the U.S. FDA.[4] This assay allows users to specifically detect the species of pine associated with complaints of "pine mouth" and demonstrates how RT-PCR assays can be designed for closely related plant species of interest for food safety studies and beyond.

AUTHOR INFORMATION

Corresponding Author

*Telephone: 240-402-3063. E-mail sara.handy@fda.hhs.gov.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are sincerely grateful to Matthew Parks of Griffith University, Richard Cronn of the United States Forest Service, and Aaron Liston of Oregon State University for sharing their plastid alignment, pre-publication, as well as to M.P. and A.L. for reviewing this manuscript before submission. In addition, we thank Aftan Vander Zwaag (IDT) for help with probe design and John Callahan and Christine Hebling for help with experimental design and manuscript edits. We appreciate Ali-Reza Fardin-Kia and Jeanne Rader for sharing pine specimens, Anne Eischeid and the Bioanalytical Methods Branch at the U.S. FDA for use of their Mx3005P QPCR system and help with troubleshooting, and Greg Diachenko for initiating our part of this project. Finally, we are thankful for the comments of three anonymous reviewers who improved the manuscript.

ABBREVIATIONS USED

RT-PCR, real-time polymerase chain reaction; DNA, deoxyribonucleic acid; FAO, Food and Agriculture Organization; LNA, locked nucleic acid

REFERENCES

(1) Kwegyir-Afful, E.; DeJager, L.; Handy, S. M.; Wong, J.; Begley, T.; Luccioli, S.An investigational report into the incidence and causes of pine mouth syndrome events in U.S. consumers. Unpublished work.

(2) Destaillats, F.; Cruz-Hernandez, C.; Giuffrida, F.; Dionisi, F.; Mostin, M.; Verstegen, G. Identification of the botanical origin of commercial pine nuts responsible for dysgeusia by gas−liquid chromatography analysis of fatty acid profile. *J. Toxicol.* **2011**, 1−7.

(3) Köbler, H.; Monakhova, Y. B.; Kuballa, T.; Tschiersch, C.; Vancutsem, J.; Thielert, G.; Mohring, A.; Lachenmeier, D. W. Nuclear magnetic resonance spectroscopy and chemometrics to identify pine nuts that cause taste disturbance. *J. Agric. Food Chem.* **2011**, *59* (13), 6877−6881.

(4) Handy, S. M.; Parks, M. B.; Deeds, J. R.; Liston, A.; de Jager, L. S.; Luccioli, S.; Kwegyir-Afful, E.; Fardin-Kia, A. R.; Begley, T. H.; Rader, J. I.; Diachenko, G. W. Use of the chloroplast gene ycf1 for the genetic differentiation of pine nuts obtained from consumers experiencing dysgeusia. *J. Agric. Food Chem.* **2011**, *59* (20), 10995−11002.

(5) Zonneveld, B. J. M. Pine nut syndrome: A simple test for genome size of 12 pine nut-producing trees links the bitter aftertaste to nuts of *P. armandii* Zucc. ex Endl. *Plant Syst. Evol.* **2011**, *297*, 201−206.

(6) Fardin-Kia, A. R.; Handy, S. M.; Rader, J. I. Characterization of pine nuts in the U.S. market, including those associated with "pine mouth", by GC−FID. *J. Agric. Food Chem.* **2012**, *60* (10), 2701−2711.

(7) Josefsen, M. H.; Löfström, C.; Sommer, H. M.; Hoorfar, J. Diagnostic PCR: Comparative sensitivity of four probe chemistries. *Mol. Cell. Probes* **2009**, *23*, 201−203.

(8) Ugozzoli, L. A.; Latorra, D.; Pucket, R.; Arar, K.; Hamby, K. Real-time genotyping with oligonucleotide probes containing locked nucleic acids. *Anal. Biochem.* **2004**, *324*, 143−152.

(9) Malorny, B.; Paccassoni, E.; Fach, P.; Bunge, C.; Martin, A.; Helmuth, R. Diagnostic real-time PCR for detection of *Salmonella* in food. *Appl. Environ. Microbiol.* **2004**, 7046−7052.

(10) Fukushima, H.; Katsube, K.; Tsunomori, Y.; Kishi, R.; Atsuta, J.; Akiba, Y. Comprehensive and rapid real-time PCR analysis of 21 foodborne outbreaks. *Int. J. Microbiol.* **2009**, 1−13.

(11) Mustorp, S.; Engdahl-Axelsson, C.; Svensson, U.; Holck, A. Detection of celery (*Apium graveolens*), mustard (*Sinapis alba*, *Brassica juncea*, *Brassica nigra*) and sesame (*Sesamum indicum*) in food by real-time PCR. *Eur. Food Res. Technol.* **2008**, *226*, 771−778.

(12) Brezná, B.; Dudásová, H.; Kuchta, T. A novel real-time polymerase chain reaction method for the detection of Brazil nuts in food. *J. AOAC Int.* **2010**, *93* (1), 197−201.

(13) Herrero, B.; Vieites, J. M.; Espiñeira, M. Fast real-time PCR for the detection of crustacean allergen in foods. *J. Agric. Food Chem.* **2012**, *60* (8), 1893−1897.

(14) Sawyer, J.; Wood, C.; Shanahan, D.; Gout, S.; McDowell, D. Real-time PCR for quantitative meat species testing. *Food Control* **2003**, *14*, 579−583.

(15) Brzezinski, J. L. Detection of sesame seed DNA in foods using real-time PCR. *J. Food Prot.* **2007**, *70* (4), 1033−1036.

(16) Rasmussen-Hellberg, R.; Naaum, A.; Handy, S. M.; Hanner, R.; Yancy, H.; Deeds, J.; Morrissey, M. Interlaboratory evaluation of a real-time multiplex polymerase chain reaction method for identification of salmon and trout species in commercial products. *J. Agric. Food Chem.* **2011**, *59*, 876−884.

(17) Pennisi, E. Wanted: A barcode for plants. *Science* **2007**, *318*, 190−191.
(18) Parks, M.; Cronn, R.; Liston, A. Separating the wheat from the chaff: Mitigating the effects of noise in a plastome phylogenomic data set from *Pinus* L. (Pinaceae). *BMC Evol. Biol.* **2012**, *12*, 100.
(19) Gardner, S. N.; Slezak, T. Scalable SNP analyses of 100+ bacterial or viral genomes. *J. Forensic Res.* **2010**, 01.
(20) Drummond, A. J.; Ashton, B.; Buxton, S.; Cheung, M.; Cooper, A.; Heled, J.; Kearse, M.; Moir, R.; Stones-Havas, S.; Sturrock, S.; Thierer, T.; Wilson, A. *Geneious v5.1*, 2010; http://www.geneious.com.
(21) Handy, S. M.; Deeds, J.; Ivanova, N.; Hebert, P.; Hanner, R.; Ormos, A.; Weigt, L.; Moore, M.; Yancy, H. A single laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. *J. AOAC Int.* **2011**, *94*, 1.
(22) Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792−1997.
(23) Parks, M.; Cronn, R.; Liston, A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* **2009**, *7*, 84.
(24) Cronn, R.; Liston, A.; Parks, M.; Gernandt, D. S.; Shen, R.; Mockler, T. Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Res.* **2008**, *36*, No. e122.