Food Chemistry: Molecular Sciences 4 (2022) 100070

Contents lists available at ScienceDirect

Food Chemistry: Molecular Sciences

journal homepage: www.sciencedirect.com/journal/food-chemistry-molecular-sciences

Sorting out the plants responsible for a contamination with pyrrolizidine alkaloids in spice seeds by means of LC-MS/MS and DNA barcoding: Proof of principle with cumin and anise spice seeds

Marie Willocx [a], Iris Van der Beeten [b], Pieter Asselman [b], Lynn Delgat [b], Wim Baert [b], Steven B. Janssens [b,c], Frederik Leliaert [b], Jean-François Picron [a,1], Celine Vanhee [a,*,1]

[a] Department of Chemical and Physical Health Risks, Sciensano, 14 rue Juliette Wytsman, 1050 Brussels, Belgium
[b] Meise Botanic Garden, Nieuwelaan 38, 1860 Meise, Belgium
[c] Department of Biology, KU Leuven, Kasteelpark Arenberg 31, 3001 Leuven, Belgium

ARTICLE INFO

Keywords:
Pyrrolizidine alkaloids
Seed spices
UHPLC-MS/MS
DNA barcoding
Phytotoxin producing plant species

ABSTRACT

High value commodities such as spices suffer from occasional contaminations of both chemical and biological origin. Consequently, quality control and safety monitoring has become a pressing issue for the spice industry. Two recent independent studies showed that at least one third of the analyzed cumin and green anise spice seeds samples surpassed the by the European Union recently established threshold value for toxic pyrrolizidine alkaloids (PAs) and their corresponding N-oxides (PANOs). These heterocyclic secondary plant metabolites are produced by a large number of different plant families. In those spice seeds, it was found by means of DNA metabarcoding, that predominant contamination was due to the presence of herbal material from the *Heliotropium* genus (Boraginaceae). Unfortunately, the use of this specific type of DNA-based identification remains controversial for the majority of the official instances and preference is still given to the use of more tangible classical approaches, including microscopy and chemical analysis. However, these methodologies often suffer from inherent drawbacks. Here we demonstrate that at least for spice seeds, a combinatory approach of microscopy, chemical analysis and classical DNA barcoding of the isolated contaminants using the *matK* and *trnH-psbA* loci, provides qualitative and quantitative information on the amount of plant material responsible for the contaminations and the extent of the contamination. The generated data also demonstrates that the presence of a very limited number of *Heliotropium* sp. seeds in a standard commercially available canister is sufficient to surpass the allowed threshold value, illustrating once more the importance of weed control.

1. Introduction

Spices have been used for millennia both for flavoring and medicinal purposes. These high value commodities often have a long and complex supply chain, making them susceptible to contaminations of various origins (Krishnaraj, Gunaseelan, Arunmozhi, & Sumandiran, 2020, Modupalli, Naik, Sunil, & Natarajan, 2021). Indeed, chemical contaminations with pesticides, heavy metals (Bua, Annuario, Albergamo, Cicero, & Dugo, 2016, Reinholds, Pugajeva, Bavrins, Kuckovska, & Bartkevics, 2017, Kowalska, 2021), persistent organic pollutants (POPs), including polychlorinated dibenzo-*p*-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs) and plasticizers, have been documented (Di Bella et al., 2018, Rivera-Pérez, Romero-González, & Garrido Frenich, 2021). Moreover, there have also been several reports illustrating the contamination of spices with microorganisms, their spores or their toxins (Waśkiewicz, Beszterda, Bocianowski, & Goliński, 2013, Hariram & Labbé, 2015, Man, Mare, Toma, Curticăpean, & Santacroce, 2016, Pickova, Ostry, Malir, Toman, & Malir, 2020, Pantano et al., 2021, Mathot, Postollec, & Leguerinel, 2021). Consequently, quality control and safety monitoring has become a pressing issue for the spice industry. Therefore, spices are increasingly subjected to systematic quality control programs while analytical methodologies for previously unassessed possible contaminants are being generated (Modupalli et al., 2021).

* Corresponding author at: rue Juliette Wytsman, 14 - 1050 Brussels, Belgium.
E-mail address: celine.vanhee@sciensano.be (C. Vanhee).
[1] Equally contributing project leaders.

https://doi.org/10.1016/j.fochms.2021.100070
Received 3 November 2021; Received in revised form 19 December 2021; Accepted 28 December 2021
Available online 1 January 2022

Very recently, the occurrence of pyrrolizidine alkaloids (PAs) and their corresponding N-oxides (PANOs), originating from a contamination with this type of phytotoxin producing plant species, was assessed in different herbal matrices, including spices (Chung & Lam, 2017, Kaltner, Rychlik, Gareis, & Gottschalk, 2020, Picron et al., 2021).

PA/PANOs are heterocyclic secondary plant metabolites with a typical pyrrolizidine motif. They display a wide structural diversity and occur in a vast number of species with novel structures and occurrences continuously being discovered (Schramm, Köhler, & Rozhon, 2019, Mädge, Gehling, Schöne, Winterhalter, & These, 2020). These alkaloids can exhibit a strong hepatotoxic, genotoxic, cytotoxic, tumorigenic, and neurotoxic activity, and can thereby pose a serious threat to the consumer (reviewed in Schrenk et al., 2020). Indeed, the subset of 1,2-unsaturated PA/PANOs with at least one ester bond at a specific location, as illustrated in Fig. 1, have been shown to be responsible for large incidents of acute and subacute food poisoning, with high morbidity and mortality (Kakar et al., 2010, Li, Gao, Liu, Pan, & Liu, 2021, Zhu et al., 2021). Additionally, long-term intake of small amounts of PA/PANOs has been associated with chronic diseases such as liver cirrhosis, cancer or pulmonary arterial hypertension (Edgar, Colegate, Boppré, & Molyneux, 2011, Dusemund et al., 2018). Therefore, the European Commission has just amended the European Regulation (EC) 1881/2006 regarding maximum thresholds of 21 PA/PANOs in certain foods, including seed spice and cumin seeds (European commission regulation, 2020). The new maximum levels will apply from July 2022.

Recently, two independent studies have shown that at least one third of the analyzed cumin and green anise spice seeds samples surpassed the set threshold value of 400 μg/kg (Kaltner et al., 2020; Picron et al., 2021). Moreover, the predominant type of PA/PANOs present in these samples belonged to the heliotrine-type PA/PANO, suggesting a contamination with a plant species belonging to the *Heliotropium* genus of the Boraginaceae plant family (El-Shazly & Wink, 2014, Mädge et al., 2020, Picron et al., 2021). This hypothesis was confirmed by DNA metabarcoding which demonstrated the presence of DNA of a plant species belonging to the *Heliotropium* genus in these seed samples (Picron et al., 2021).

Although DNA metabarcoding has been proven to be very useful to trace the possible origin of the contamination, it is unable to provide reliable data on the extent of the contamination. Therefore, the use of this technology remains controversial and official instances still prefer the use of more classical biological approaches such as macro- and microscopic analysis. The latter approaches can only be performed on intact plant samples (leaves, roots, seeds, …). In 2017, Chung and Lam demonstrated the utility of microscopy for the analysis of pyrrolizidine alkaloids in cumin seeds. It was found that the cumin seeds were contaminated with "foreign matter" and upon removal of these items a sharp decrease in PA/PANO content was observed. However, no identification of the herbal nature of this matter was performed. Indeed, the identification of contaminants when exclusively based on morphological similarities, can often be misleading or ambiguous. Therefore, microscopic analyses are often combined with other chemical analyses such as mass spectroscopy, spectroscopy or other (reviewed in Galvin-King & Haughey, Elliot, 2018). Although these techniques were proven to be very useful when dealing with adulteration, they may not be able to pinpoint the plant species that might be responsible for a PA/PANO contamination, as the presence of only a small amount of PA/PANO producing seeds might be responsible for exceeding the established threshold (Chung & Lam, 2017). DNA barcoding on the other hand is often used to efficiently identify a species but is, in contrast to DNA metabarcoding, often not applicable to highly processed mixtures (Newmaster, Grguric, Shanmughanandhan, Ramalingam, & Ragupathy, 2013, Parvathy et al., 2014, Parvathy, Swetha, Sheeja, & Sasikumar, 2015, Huang, Li, Liu, & Long, 2015, Raclariu, Heinrich, Ichim, & de Boer, 2018, Howard, Lockie-Williams, & Slater, 2020). However, a combinatory approach of an in-depth morphological analysis and sorting of foreign material and intact plant material such as seeds and subsequent DNA barcoding of all the different items could pinpoint the nature of a PA/PANO contamination. In the present this study, we utilize this multidisciplinary approach to provide quantitative information on the chemical nature of the contamination, as well as to identify the

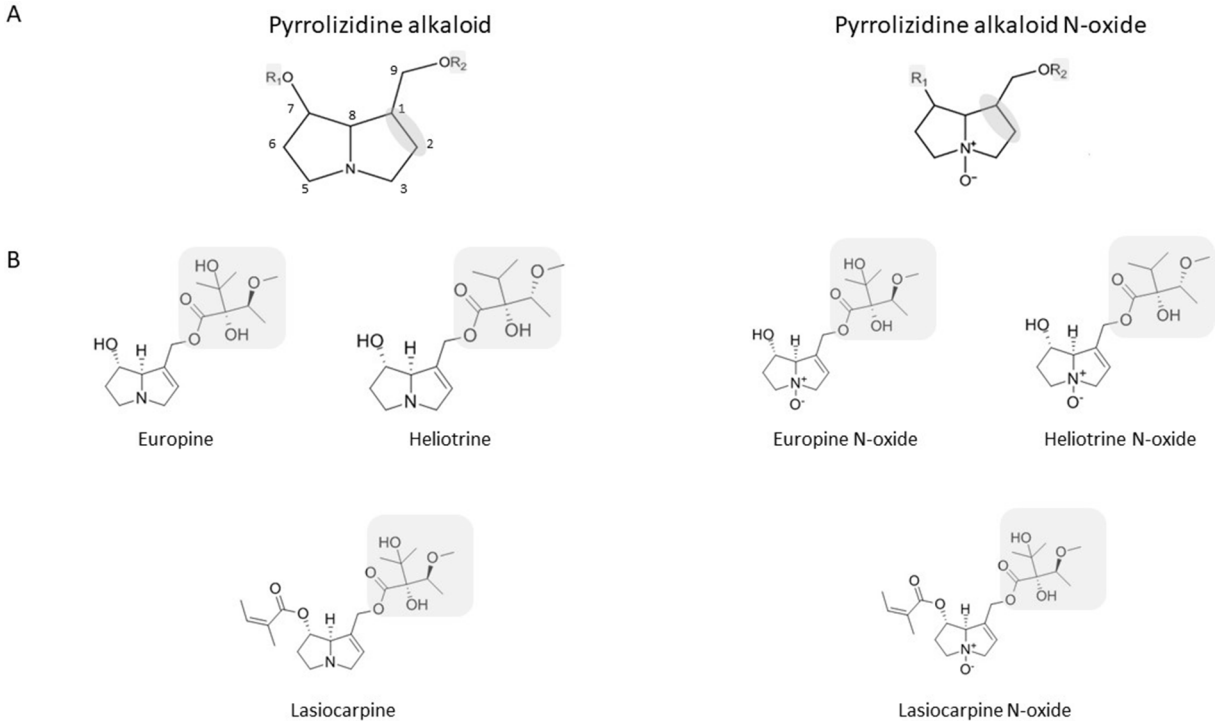

Fig. 1. General chemical structure of a PA and a PANO (A) consisting of a necine base and necic acids at the R1 or R2 position. PA/PANOs are assumed toxic when the necine base is unsaturated at the 1,2-position and esterified with a branched necic acid on the C7 and/or C9 position (grey boxes). This is the case for the heliotrine-type compounds (B) targeted in this study and found to be present in seed spice.

potential plant source and quantification of the amount of plant material that caused the contamination with PA/PANOs. This information is pivotal when raising awareness amongst the producers and the local farmers.

2. Material and methods

2.1. Solvents, reagents and standard solutions

Methanol and acetonitrile were HPLC grade and purchased from Biosolve BV (Valkenswaard, the Netherlands). Water was purified using a Millipore Milli-Q system (Millipore Corp., Bedford, MA, USA). Formic acid (98–100%) and sulphuric acid (95–97%) were ordered from Merck (Kenilworth, New Jersey, USA). Supelclean™ ENVI™-Carb SPE tubes 500 mg/6 mL (Supelco, Bellefonte, Pennsylvania, USA) were obtained from Sigma-Aldrich and Strata™-XC 33 μm Polymeric Strong cation exchange SPE cartridges, 500 mg/6 mL were bought from Phenomenex®.

Standards of pyrrolizidine alkaloids (PAs) and related N-oxides (PANOs) were ordered from PhytoLab GmbH and Co. KG (Germany): 16 PAs i.e. echimidine, erucifoline, europine, heliotrine, indicine, intermedine, jacobine, lasiocarpine, lycopsamine, monocrotaline, retrorsine, senecionine, seneciphylline, senecivernine, senkirkine, trichodesmine and 14 PANOs i.e. echimidine N-oxide, erucifoline N-oxide, europine N-oxide, heliotrine N-oxide, indicine N-oxide, intermedine N-oxide, jacobine N-oxide, lasiocarpine N-oxide, lycopsamine N-oxide, monocrotaline N-oxide, retrorsine N-oxide, senecionine N-oxide, seneciphylline N-oxide, senecivernine N-oxide. Purity was between 89% and 100%. Individual solutions of each compound at the concentration of 1 mg/mL were prepared in methanol and have proven to be stable for 1 year at −20 °C.

2.2. Sampling

Reference seeds of *Heliotropium europaeum* were purchased online while reference *Heliotropium europaeum* plant material was provided by Meise Botanic Garden.

Samples were selected based on the targeted LC-MS/MS analysis for PA/PANOs of Picron et al. (2021). One highly contaminated cumin sample (sample S06) and one highly contaminated green anise sample (sample S57) were selected, as well as one cumin (S23) and one green anise (S53) sample with low PA/PANO content. All samples were bought in supermarkets. Prior to the analysis two jars of each sample, with same lot number, were pooled and homogenized. Half of the total amount was finely ground for the LC-MS/MS analysis and half of the amount was kept in the dark and stored at room temperature for the morphological analysis and subsequent molecular analysis.

2.3. PA/PANOs extraction and LC-MS/MS analysis

The PA/PANOs extraction protocol for the ground seed material was performed as described by Picron et al., 2021. Briefly, 0.5 g (±0.005 g) of homogenized sample was weighed into a 50 mL polypropylene tube to which 20 mL of 0.05 M H_2SO_4 were added. The tube was capped, vortex-mixed for 15 min, placed in an ultrasonic bath for 15 min (35 kHz) and shaken for 15 min. The sample solution was centrifuged at 3900 rpm for 15 min at 5 °C prior to the removal of the upper aqueous solution into a 50 mL polypropylene tube. The remaining solid/solution was washed a second time by adding 10 mL of 0.05 M H_2SO_4. The tube was vortex-mixed for 10 min, shaken for 10 min and centrifuged at 3900 rpm for 15 min at 5 °C. The new aqueous phase was combined with the previous one and the crude extract was homogenized. The SPE cartridges (Strata-XC, 33 μm Polymeric Strong Cation, 500 mg/6 mL, Phenomex®) were pre-conditioned with 10 mL of MeOH and activated with 10 mL of 0.05 M H_2SO_4 prior to the application of a 12 mL aliquot of the acidic crude extract. After washing the cartridge with 10 mL of aqueous formic acid solution (0.1%) and 10 mL of MeOH, the analytes were eluted with 14 mL ammoniated methanol (3%) into a 15 mL polypropylene tube. The purified extract was evaporated to dryness under a nitrogen gas flow in a water bath at 45 °C. Next, the residue was reconstituted in 1 mL of H_2O/MeOH mixture (80:20, v/v) and transferred into an injection vial ready for analysis.

The measurements were performed on a UHPLC device for chromatographic separation followed by detection using tandem mass spectrometry (MS/MS) in multi reaction mode. Detailed LC-MS/MS conditions can be found in the Supplementary Data (see supplemental data).

2.4. Morphological sorting, DNA extraction and amplification

Prior to the DNA extraction, the aliquot of stored intact spice seed samples was subjected to phenotypical analysis by means of macroscopic sorting analysis and photographed with a Keyence VHX 5000 digital microscope, equipped with a VH-Z20R/W/T lens (Keyence, Osaka, Japan).

Different types of foreign material were encountered and grouped according to morphological similarities which were all subjected to the DNA extraction protocol. Total genomic DNA was extracted following a modified CTAB protocol (Tel-zur, Abbo, Myslabodski, 1999). After mechanical sample disruption, 800 μl CTAB lysis buffer (incl.1% PVP-40) was added prior to the incubation of the samples for 1.5 h at a constant temperature of 60 °C. Subsequently, the DNA extract, including all other co-lysed components, was washed twice with chloroform-isoamylalcohol (24/1 v/v). A final cleaning step through isopropanol precipitation was carried out and left overnight at − 20 °C. During the next step, the samples were centrifuged after which the pellet was washed with 70% ethanol. Finally, the DNA pellet obtained was air-dried, and dissolved in 50 μl TE buffer (10 mM TrisHCl pH 8, 1 mM EDTA pH 8). The amount and purity of DNA was assessed by spectrophotometry (OD 260:280 ratio) using a Nanodrop instrument (Thermo Scientific, Waltham, Massachusetts, USA). PCR reactions of either *rbcL*, *matK*, *trnH-psbA* and ITS were carried out using standard PCR conditions containing 4 μl DNA, 2 × 0.2 μl oligonucleotide primer (20 μM), 1.6 μl of 10 mM dNTPs, 1.6 μl BSA, 1.6 μl $MgCl_2$ (25 μM), 2 μl Taq Buffer, 0.1 μl DreamTaq DNA polymerase and 8.7 μl MilliQ water. Amplification reactions were performed on a T100 Thermal Cycler (Bio-Rad Laboratories, Hercules, California, USA) and carried out as described in the Supplementary Data. The obtained PCR products were then further analyzed on a 5200 Fragment Analyzer System with the dsDNA 915 Reagent Kit (35–5000 bp) (Agilent Technologies, Santa Clara, California, USA). After the purification of the PCR products using an ExoSap protocol (Wallis & Morrell, 2011), samples were sent for sequencing to Macrogen, Inc. (Seoul, South Korea).

2.5. DNA barcode identification approaches

Two approaches, as described by Ghorbani, Saeedi, and de Boer (2017) albeit with more stringent criteria for species-level identifications, were used for DNA barcode identification. These methods are solely based on BLAST similarity-based identifications of the sequences present in GenBank (Benson et al., 2013) and consisted either of considering only the top hits of the maximum score or alternatively of the similarity method, where extra weight is put on the identity value of the query-reference comparison. For both methods the obtained DNA sequences were sequentially queried using megablast online at NCBI nucleotide BLAST against the nucleotide database (Altschul, Gish, Miller, Myers, & Lipman, 1990, Morgulis et al., 2008). For the top hit method only those hits within 10 points deviation down of the max score were considered, providing that a 70% query coverage was obtained and a 98% identity was obtained. If these hits included only a single species then a species level identification was assigned, if these hits contained multiple species from the same genus, a genus level identification was

assigned or if these top hits included multiple species in different genera in the same family then a family level identification was assigned. For the optimized method a similarity score was calculated for up to 100 BLAST hits if the query cover was 70% or higher: max score*(% identity/query coverage). Subsequently all hits were ordered by this score, and the deviation for each similarity score value from the highest similarity score was calculated (see Supplementary Data). The identifications were also assigned based on the similarity score and the obtained % identity (Ghorbani et al., 2017). A species level confidence was assigned if no other species were listed within the 1% deviation from the maximum similarity score, provided that the obtained % identity $\geq$ 98%. A genus level confidence was assigned to those samples with a % identity $\geq$ 98% but with multiple species from the same genus listed within the 1% deviation from the maximum similarity score. Additionally, a genus level confidence was also assigned if no other species were listed within the 1% deviation from the maximum similarity score for hits with a medium identity score (90% $\leq$ % identity $\leq$ 98%). A family level identity was assigned for those samples where the identity score $<$ 90%.

3. Results and discussion

3.1. PA/PANOs content in selected spice seed samples

Recently the European Commission has amended Regulation (EC) No 1881/2006 regarding the maximum threshold of pyrrolizidine alkaloids in certain foodstuffs (European commission regulation, 2020). The new regulation, which is mandatory from the first of July 2022, only allows for a maximum level of 400 μg/kg of the sum of 21 pyrrolizidine alkaloids and their corresponding N-oxides for cumin seeds and seed spices. These 21 PA/PANOs belong to the four different PA/PANOs families: heliotrine-type, lycopsamine-type, senecionine-type and monocrotaline-type. For the heliotrine-type, heliotrine, heliotrine N-oxide, europine, europine N-oxide, lasiocarpine and lasiocarpine N-oxide are being monitored (see Table 1). Chemically, the PAs are composed of a saturated or 1,2-unsaturated necine base esterified with one or two necic acids and they can occur in either the free tertiary base form or the corresponding PANO (see Fig. 1A).

Previous studies (Kaltner et al., 2020, Picron et al., 2021) on seed spices demonstrated a contamination by heliotrine-type compounds. In fact, these heliotrine-type compounds were the dominant source of PA/PANOs present in previously analyzed seed spices (and aromatic mixtures mainly consisting of seed spices) with up to 98.9% of the PA/PANOs encountered belonging to this family (Picron et al., 2021). The presence of these types of compounds is generally attributed to a contamination with a species from the Boraginaceae plant family, and likely but not exclusively to a contamination with a member of the *Heliotropium* genus (El-Shazly & Wink, 2014 and Mädge et al., 2020). Indeed, DNA metabarcoding suggested a contamination with a *Heliotropium* species (Picron et al., 2021). In order to verify these results obtained with DNA metabarcoding, one highly contaminated cumin sample and one highly contaminated green anise sample was subjected to both targeted LC-MS/MS analysis for PA/PANOs and morphological analysis to separate the desired plant seeds from foreign material (Picron et al., 2021). Additionally, samples (one cumin and one green anise) with low PA/PANO content were subjected to the same analysis. As can be seen from Table 1, the highly contaminated cumin sample, named sample S-06 contained 8515.0 μg/kg of total PA/PANOs from which 99.9% of the found PA/PANOs belonged to the heliotrine-type of compounds, including europine, europine N-oxide, heliotrine, heliotrine N-oxide, lasiocarpine and lasiocarpine N-oxide (see Fig. 1B). This tendency was also seen for the heavily contaminated green anise seeds, named sample S-57, with a total amount of PA/PANOs of 1653.1 μg/kg, where only heliotrine-type PA/PANOs were detected.

Table 1
The amount of PA/PANOs determined for the different samples. The PA/PANOs that are included into the list of 21 PA/PANOs to be monitored in regard to the by the European union set maximum level for PA/PANOs in food stuff, are marked in bold.

		Cumin		Green Anise	
Type PA/PANO		**S-06**	**S-23**	**S-53**	**S-57**
Heliotrine type	**Europine**	463,9	0,7	1,1	170,7
	Europine-N-Oxide	5029,0	0,3	2,9	1459,4
	Heliotrine	282,0	2,9	1,5	7,2
	Heliotrine-N-Oxide	1829,0	0,5	3,6	8,1
	Lasiocarpine	115,7	0,3	0,6	3,2
	Lasiocarpine-N-Oxide	790,4	0,0	2,3	4,5
Lycopsamine type	**Echimidine**	0,0	0,0	0,0	0,0
	Echimidine-N-Oxide	0,0	0,0	0,0	0,0
	Indicine	0,0	0,0	0,0	0,0
	Indicine N-oxide	0,0	0,0	0,0	0,0
	Intermedine	0,0	0,0	0,0	0,0
	Intermedine-N-Oxide	0,9	0,0	0,0	0,0
	Lycopsamine	0,0	0,0	0,0	0,0
	Lycopsamine-N-Oxide	0,0	0,0	0,0	0,0
Senecionine type	**Retrorsine**	3,2	0,0	0,0	0,0
	Retrorsine-N-Oxide	0,0	0,0	0,0	0,0
	Senecionine	0,0	0,0	0,0	0,0
	Senecionine-N-Oxide	0,0	0,0	0,0	0,0
	Seneciphylline	0,0	0,0	0,0	0,0
	Seneciphylline-N-Oxide	0,0	0,0	0,0	0,0
	Senecivernine	0,0	0,0	0,0	0,0
	Senecivernine-N-Oxide	0,0	0,0	0,0	0,0
	Senkirkin	0,0	0,0	0,0	0,0
	Jacobine	0,0	0,0	0,0	0,0
	Jacobine-N-oxide	0,0	0,0	0,0	0,0
	Erucifoline	0,0	0,0	0,0	0,0
	Erucifoline-N-oxide	0,0	0,0	0,0	0,0
Mono-crotaline type	Monocrotaline	0,0	0,0	0,0	0,0
	Monocrotaline-N-Oxide	0,0	0,0	0,0	0,0
	Trichodesmine	0,9	0,0	0,0	0,0
Total		**8515**	**4,7**	**12**	**1653,1**
Heliotrine type		8510	4,7	12	1653,1
Lycopsamine type		0,9	0,0	0,0	0,0
Senecionine type		3,2	0,0	0,0	0,0
Monocrotaline type		0,9	0,0	0,0	0,0

3.2. Morphological analysis

All samples were subjected to phenotypical analysis. Different types of foreign material were encountered and grouped according to macroscopic and microscopic similarities and possible type of material (seed/twigs/wood leaf/unknown). These in total 55 different morphological items, depicted in Tables 2 and 3, were sorted from the main constituents being either cumin seeds or green anise seeds. As can be seen from Table 2, the cumin seeds sample S-23 that contained very little PA/PANOs contamination also contained very little foreign material compared to the heavily contaminated sample S-06. However, both the green anise samples (see Table 3), S-53 and S-57, contained several different groups of contaminants. Interestingly, both samples (the cumin sample S-06 and the green anise sample S-57) contained similar foreign material which could be described as a “sculptured nutlet”, which was not encountered in the samples with low PA/PANOs levels.

3.3. DNA barcoding

DNA barcoding makes use of specific regions of DNA to identify species. In 2009, the Consortium for the Barcode of Life (CBOL) Plant Working Group proposed the use of the chloroplast genes *rbcL* and *matK*

Table 2
Overview of the different morphological items encountered in the different cumin samples and the identifications obtained of the DNA barcoding. Different barcode identification approaches were utilized, the top hit method for both loci (*matK* or *trnH-psbA*) separately, the combined top hit method, the similarity method for both loci separately and the combined similarity approach. n.d.: not determined. a: query coverage < 70%; b: no amplification orsequencing failed; c: presence of multiple PCR bands; d: no DNA present.

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
Cumin seeds	Dried cumin seeds			*mat K:Cuminum cyminumtrnH-psbA:Cuminum cyminum*	*Cuminum cyminum*	*mat K:Cuminum cyminumtrnH-psbA:Cuminum cyminum*	*Cuminum cyminum*
	Cumin seed embryo and seed coats			*mat K:Cuminum cyminumtrnH-psbA:* n.d.[a]	*Cuminum cyminum*	*mat K:Cuminum cyminumtrnH-psbA:* n.d.[a]	*Cuminum cyminum*
Contaminants sample C-1	Contaminant 1: sculptured nutlet	14		*mat K:Heliotropium ellipticumtrnH-psbA:* n.d.[a]	*Heliotropium ellipticum*	*mat K:Heliotropium* sp.*trnH-psbA:* n.d.[a]	*Heliotropium* sp.
	Contaminant 2: seed with spines	11		*mat K:Torilis scabratrnH-psbA:Torilis* sp.	*Torilis scabra*	*mat K:Torilis* sp.*trnH-psbA: Torilis japonica*	*Torilis japonica*

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
	Contaminant 3: ovoid seeds	5		*mat K:Pimpinella* sp.*trnH-psbA:* n.d.[b]	*Pimpinella* sp.	*mat K:Pimpinella* sp.*trnH-psbA:* n.d.[b]	*Pimpinella* sp.
	Contaminant 4:shiny ovoid seeds	5		*mat K:Polygonum avicularetrnH-psbA:* n.d.[c]	*Polygonum aviculare*	*mat K:Polygonum* sp.*trnH-psbA:* n.d.[c]	*Polygonum* sp.
	Contaminant 5	4		*mat K:Phalaris* sp.*trnH-psbA:* n.d.[a]	*Phalaris* sp.	*mat K:Phalaris* sp.*trnH-psbA:* n. d.[a]	*Phalaris* sp.
	Contaminant 6: round yellow seeds	3		*mat K:Vicia monanthatrnH-psbA:* n.d.[c]	*Vicia monantha*	*mat K:Vicia monanthatrnH-psbA:* n.d.[c]	*Vicia monantha*

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
	Contamint 7: black sculptured seed	2		*mat K:* n.d.[b]*trnH-psbA: Gypsophila* sp.	*Gypsophila* sp.	*mat K:* n.d.[b]*trnH-psbA: Caryophyllaceae*	Caryophyllaceae
	Contaminant 8: black seed	1		*mat K:Ixilirion tataricumtrnH-psbA: n.d.*	*Ixilirion tataricum*	*mat K:Ixilirion tataricumtrnH-psbA: Asparagaceae*	*Ixilirion tataricum*
	Contaminant 9: camel colored seed	1		mat K:*Lathyrus inconspicuus*trnH-psbA: *Lathyrus inconspicuus*	*Lathyrus inconspicuus*	mat K:*Lathyrus inconspicuus*trnH-psbA:*Lathyrus inconspicuus*	*Lathyrus inconspicuus*
	Contaminant 10: round structured seed	1		*mat K:* n.d.[b]*trnH-psbA:Galium aparine*	*Galium aparine*	*mat K:* n.d.[b]*trnH-psbA:Galium aparine*	*Galium aparine*

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
	Contaminant 11: light brown seed	1		*mat K:Euphorbia* sp.*trnH-psbA:Euphorbia* sp.	*Euphorbia* sp.	*mat K:Euphorbia nicaeensis* subsp. *nicaeensistrnH-psbA: Euphorbiaceae*	*Euphorbia nicaeensis* subsp. *nicaeensis*
	Contaminant 12: orange-colored seed	1		*mat K:Euphorbia* sp.*trnH-psbA:* n.d.	*Euphorbia* sp.	mat K:*Euphorbia nicaeensis* subsp. *nicaeensistrnH-psbA:* n.d.	*Euphorbia nicaeensis* subsp. *nicaeensis*
	Contaminant 13: coriander like seed	1		*mat K:* n.d.[b]*trnH-psbA:* n.d.[b]	unidentified	*mat K*: n.d.[b]*trnH-psbA:* n.d.[b]	unidentified
	Contaminant 14: light brown seed	1		*matK:*Poaceae*trnH-psbA:* n. d.[b]	Poaceae	*matK:*Poaceae*trnH-psbA:* n.d.[b]	Poaceae

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
	Contaminant 15: part of a seed	1		*mat K:Vicia monanthatrnH-psbA:* n.d.[b]	*Vicia monantha*	*mat K:Vicia monanthatrnH-psbA:* n.d.[b]	*Vicia monantha*
	Contaminant 16: twigs of different size	multiple		*mat K:Cuminum cyminumtrnH-psbA:* n.d.[b]	*Cuminum cyminum*	*mat K:Cuminum cyminumtrnH-psbA:* n.d.[b]	*Cuminum cyminum*
	Contaminant 17: green twigs	2		*mat K:Euphorbia* sp.*trnH-psbA: Euphorbia* sp.	*Euphorbia* sp.	*mat K:Euphorbia nicaeensis* subsp. *nicaeensistrnH-psbA:* Euphorbiaceae	*Euphorbia nicaeensis* subsp. *nicaeensis*
	Contaminant 18: leaf-like structure			*mat K:Cuminum cyminumtrnH-psbA:* n.d.[c]	*Cuminum cyminum*	*mat K:Cuminum cyminumtrnH-psbA*: n.d.[c]	*Cuminum cyminum*

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
	Contaminant 19: white flake like structures	6		*mat K:Vicia monanthatrnH-psbA:* n.d.[b]	*Vicia monantha*	*mat K:Vicia monanthatrnH-psbA:* n.d.[b]	*Vicia monantha*
	Contaminant 20: sail ship structures	2			n.d.[d]		
	Contaminant 21: miscellaneous	8		n.d.[d]			
Contaminants sample C-2	Contaminant 22: elongated yellow–red seed	3		*mat K:Plantago ovatatrnH-psbA:Plantago ovata*	*Plantago ovata*	*mat K:Plantago ovatatrnH-psbA: Plantago ovata*	*Plantago ovata*

(*continued on next page*)

Table 2 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to combined top hit	ID according to the similarity method	ID according to combined similarity method
	Contaminant 23: little brown seed	1		*mat K:* n.d.[b]*trnH-psbA: Brassica napus*	*Brassica napus*	*mat K:* n.d.[b]*trnH-psbA:Brassica napus*	*Brassica napus*
	Contaminant 24: green–brown seeds	1		*mat K:Melilotus* sp.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*	*mat K:Melilotus* sp.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*
	Contaminant 25: elongated dark brown seed	1		*mat K:Plantago* sp.*trnH-psbA*: n.d.[b]	*Plantago* sp.	*mat K:Plantago* sp.*trnH-psbA:* n.d.[b]	*Plantago* sp.
	Contaminant 26: white and brown striped seed	1		*mat K:Anethum graveolenstrnH-psbA:Anethum graveolens*	*Anethum graveolens*	*mat K:Anethum* sp.*trnH-psbA: Anethum graveolens*	*Anethum graveolens*

Table 3
Overview of the different morphological items encountered in the different green anise samples and the identifications obtained of the DNA barcoding. Different barcode identification approaches were utilized, the top hit method for both loci (*matK* or *trnH-psbA*) separately, the combined top hit method, the similarity method for both loci separately and the combined similarity approach. n.d.: not determined. a: query coverage < 70%; b: no amplification or sequencing failed; c: presence of multiple PCR bands; d: no DNA present.

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
Green Anise seeds	Dried green anise seeds			*mat K:Pimpinella* sp.*trnH-psbA:Pimpinella* sp.	*Pimpinella* sp.	*mat K:Pimpinella* sp.*trnH-psbA: Pimpinella* sp.	*Pimpinella* sp.
	Green Anise seed embryo and seed coats			*mat K:Pimpinella anisumtrnH-psbA:* n.d.[b]	*Pimpinella anisum*	*mat K:Pimpinella* sp.*trnH-psbA:* n.d.[b]	*Pimpinella* sp.
	Immature green anise seed			*mat K:Pimpinella* sp.*trnH-psbA:Pimpinella* sp.	*Pimpinella* sp.	*mat K:Pimpinella* sp.*trnH-psbA: Pimpinella* sp.	*Pimpinella* sp.
Contaminants sample S-57	Contaminant 1: coriander like seeds	multiple		*mat K:Coriandrum sativumtrnH-psbA: Coriandrum sativum*	*Coriandrum sativum*	*mat K:Coriandrum sativumtrnH-psbA: Coriandrum sativum*	*Coriandrum sativum*

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 2: shiny ovoid seeds	18		*matK:Polygonum aviculare*trn*H-psbA:* n.d.[c]	*Polygonum aviculare*	*matK:Polygonum* sp.*trnH-psbA:* n.d.[c]	*Polygonum* sp.
	Contaminant 3: green–brown seeds	17		*matK:Melilotus* sp.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*	*matK:Melilotus* sp.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*
	Contaminant 4:	6		*matK:Amaranthaceae*trn*H-psbA:* n.d.[c]	*Amaranthaceae*	*matK:Amaranthaceae*trn*H-psbA:* n.d.[c]	*Amaranthaceae*
	Contaminant 5: elongated seeds	4		*matK:*Poaceae*trnH-psbA:* Poaceae	Poaceae	*matK:*Poaceae*trnH-psbA:* Poaceae	Poaceae

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 6: round nutlets	4		n.d.[d]			
	Contaminant 7: dark brown balls	2		*matK:Sinapis arvensistrnH-psbA:Sinapis arvensis*	*Sinapis arvensis*	*matK:Sinapis arvensistrnH-psbA:Sinapis arvensis*	*Sinapis arvensis*
	Contaminant 8: structured nutlet	1		*matK:Melilotus* sp.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*	*matK:Melilotus* sp.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*
	Contaminant 9: cumin like seed	2		*mat K:Cuminum cyminumtrnH-psbA:Cuminum cyminum*	*Cuminum cyminum*	*mat K:Cuminum cyminumtrnH-psbA:Cuminum cyminum*	*Cuminum cyminum*

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 10: structured nutlet	2		*mat K:Heliotropium ellipticumtrnH-psbA: Heliotropium* sp.	*Heliotropium ellipticum*	*mat K:Heliotropium* sp.*trnH-psbA: Heliotropium* sp.	*Heliotropium* sp.
	Contaminant 11: Shiny disc shaped seed	1		*mat K:Amaranthus blitoidestrnH-psbA: Amaranthus* sp.	*Amaranthus blitoides*	*mat K:Amaranthus* sp.*trnH-psbA: Amaranthus* sp.	*Amaranthus* sp.
	Contaminant 12: pale triangular seed	1		*mat K:Nigella* sp.*trnH-psbA: Nigella* sp.	*Nigella* sp.	*mat K:Nigella* sp.*trnH-psbA: Nigella* sp.	*Nigella* sp.
	Contaminant 13: twigs of different size	multiple		*mat K:Pimpinella* sp.*trnH-psbA:* n.d.[b]	*Pimpinella* sp.	*mat K:Pimpinella* sp.*trnH-psbA:* n.d.[b]	*Pimpinella* sp.

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 14: thicker twigs than contaminant 13	Multiple		*mat K:Pimpinella diversifoliatrnH-psbA: Pimpinella* sp.	*Pimpinella diversifolia*	*mat K:Pimpinella anisumtrnH-psbA:Pimpinella sp*	*Pimpinella anisum*
	Contaminant 15: curly twigs	6		*mat K:* n.d.[b]*trnH-psbA: Cuscuta* sp.	*Cuscuta* sp.	*mat K:* n.d.[b]*trnH-psbA: Cuscuta* sp.	*Cuscuta* sp.
	Contaminant 16: pieces of leaves			*mat K:Convolvulus arvensistrnH-psbA:* n.d.[b]	*Convolvulus arvensis*	*mat K:Convolvulus arvensistrnH-psbA:* n.d.[b]	*Convolvulus arvensis*
	Contaminant 17	2		*mat K:Pimpinella diversifoliatrnH-psbA: Pimpinella* sp.	*Pimpinella diversifolia*	*mat K:Pimpinella* sp.*trnH-psbA:Pimpinella sp*	*Pimpinella* sp.

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 18: rock	1		n.d.[d]			
Contaminants sample S-53	Contaminant 19: coriander like seeds	multiple		*mat K:Coriandrum sativumtrnH-psbA: Coriandrum sativum*	*Coriandrum sativum*	*mat K:Coriandrum sativumtrnH-psbA: Coriandrum sativum*	*Coriandrum sativum*
	Contaminant 20: crown shaped seeds	27		*matK:Chenopodium vulvariatrnH-psbA:* n.d.[c]	*Chenopodium vulvaria*	*matK:Chenopodium vulvariatrnH-psbA:* n.d.[c]	*Chenopodium vulvaria*
	Contaminant 21: green brown seeds	10		*matK:* n.d.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*	*matK:* n.d.*trnH-psbA: Melilotus indicus*	*Melilotus indicus*

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 22: shiny ovoid seeds	7		*matK:Polygonum aviculare*trnH-psbA: n.d.[c]	*Polygonum aviculare*	*matK:Polygonum* sp.*trnH-psbA:* n.d.[c]	*Polygonum* sp.
	Contaminant 23: brown-grey seed			matK:Sinapis albatrnH-psbA: Sinapis alba	*Sinapis alba*	matK:Sinapis albatrnH-psbA: Sinapis alba	*Sinapis alba*
	Contaminant 24: camouflage like seed			*matK:Euphorbia* sp.*trnH-psbA:Euphorbia* sp.	*Euphorbia* sp.	*matK:Euphorbia myrsinites*trnH-psbA: *Euphorbiaceae*	*Euphorbia myrsinites*
	Contaminant 25: little brown ball			*matK:* n.d.[b]*trnH-psbA:Cuscuta campestris*	*Cuscuta campestris*	*matK:* n.d.[b]*trnH-psbA:Cuscuta campestris*	*Cuscuta campestris*

(*continued on next page*)

Table 3 (*continued*)

Sample ID	Morphological description	Number encountered	Microscopic analysis	ID according to top hit method	ID according to the combined top hit method	ID according to the similarity method	ID according to the combined similarity method
	Contaminant 26: rock like pieces			*mat K:Convolvulus arvensistrnH-psbA: Convolvulaceae*	*Convolvulus arvensis*	*mat K:Convolvulus arvensistrnH-psbA: Convolvulus* sp.	*Convolvulus arvensis*
	Contaminant 27: irregular structure			*matK:Tanacetum* sp.*trnH-psbA:Tanacetum coccineum*	*Tanacetum coccineum*	*matK:AsteraceaetrnH-psbA: Tanacetum coccineum*	*Tanacetum coccineum*
	Contaminant 28: white structures			*matK:Euphorbia* sp.*trnH-psbA:Euphorbia* sp.	*Euphorbia* sp.	*matK:Euphorbia myrsinitestrnH-psbA: Euphorbiaceae*	*Euphorbia myrsinites*

Table 4
Overview of the identification rates for the different barcode identification approaches.

	% species level identifications	% genus level identifications	% family level identifications	% unidentified
Top hit *matK*	60.4(29/48)	33.3(16/48)	6.3(3/48)	
Similarity method m*atK*	52.1(25/48)	39.6(19/48)	8.3(4/48)	
Top hit *trnH-psbA*	47.2(17/36)	38.9(14/36)	5.6(2/36)	8.3(3/36)
Similarity method *trnH-psbA*	50.0(18/36)	25.0(9/36)	19.4(7/36)	5.5(2/36)
Combined top hit	67.3(37/55)	25.5(14/55)	5.5(3/55)	1.8(1/55)
Combined similarity method	61.8(34/55)	29.1(16/55)	7.3(4/55)	1.8(1/55)

as the main barcodes to be used for plant species identification whether or not supplemented where needed with the use of the chloroplast intergenic sequence *trnH-psbA* and nuclear ribosomal internal transcribed spacer (ITS) region (CBOL Plant Working Group, 2009). Although the *rbcL* gene has been widely used for phylogenetic analysis, variation at the species level is limited, often resulting in a poor ability to discriminate plant species (Newmaster, Fazekas, & Ragupathy, 2006, Gonzalez et al., 2009). Therefore, the *matK* locus was chosen since it has demonstrated an evolution rate which is 2–3 times higher than *rbcL*, making it more suitable for species-level identifications (Lahaye et al., 2008). Nevertheless, the gene is ~ 1,570 bp in length and consequently, might result in a relatively low amplification success rate (Newmaster et al., 2006, Kress & Erickson, 2007, Fazekas et al., 2008, Kool et al., 2012). Additionally, also the *trnH-psbA* sequence was chosen as another independent locus with a shorter amplicon length and a rapid evolution rate. However, this locus is subject to insertion/deletion events leading to variation in fragment length, and causing difficulties in comparing species from different genera (Pang et al., 2012). The ITS region also suffers from this variation in length and often also results in the amplification of fungal DNA due to contamination by this microorganism (Ghorbani et al., 2017, Timpano, Scheible, & Meiklejohn, 2020), which might add an additional difficulty to sequence analysis when doing DNA barcoding.

3.3.1. Optimization of the DNA extraction protocol and choice of barcode regions

Based on the morphological sorting, the samples were contaminated with different types of material (seeds, twigs, leaves and other unknown items). Therefore, prior to the analysis of the different morphological items, the DNA extraction and subsequent PCR amplification of the different barcode loci were assessed on both a fraction of the sorted cumin seeds and *Heliotropium Europaeum* seeds and *Heliotropium europaeum* leaves, as a representative for the Heliotropium genus. The final extraction protocol (see material and methods) was able to extract enough DNA from only 2 seeds to perform multiple PCR reactions and to assess the discriminative power for all tested barcodes. Different primers, amplifying either *rbcL*, *matK*, *trnH-psbA* or ITS locus, were tested. The amplification was successful for all loci for the cumin seed DNA. However, for the *H. europaeum* seeds only the *matK*, *trnH-psbA* and ITS region were successfully amplified while for the DNA obtained from *H. europaeum* leaves DNA there was no amplification of the ITS region. Moreover, further assessment of the sequences obtained for the ITS region of the cumin seed DNA demonstrated that the ITS sequences were contaminated with fungal DNA. Based on this data, we decided to utilize both the *matK* and *trnH-psbA* loci as barcodes.

3.3.2. DNA extraction, amplification success, sequencing and blast searching of the samples

Prior to any sequencing, the concentration of the obtained DNA was determined. From the 59 items 4 items did not result in the occurrence of detectable amounts of DNA and were omitted from further sequencing analysis. The amplification success for *matK* was 81.4 % (48/55 items) while the success rate of *trnH-psbA* was 78.2% (43/55 items). However, for the *trnH-psbA* spacer 7 items resulted in the presence of multiple sequencing fragments, which indicates the non-specificity of the chosen primer pair and amplification settings for these specific items or could alternatively be due to a mixture of different species that were morphologically indistinguishable from each other. Consequently, only 36 *trnH-psbA* fragments were sequenced and subjected to subsequent data analysis.

The simple and optimized BLAST results, called top hit method and the similarity method, both based on sequence similarity searching, as

Table 5
Overview of the different potential phytotoxin producing species/genera encountered in the samples.

ID according to combined similarity method	Sample ID	Morphological description	Number encoun-tered	Toxic plants forbidden by the Belgian government to be present in any plant or plant derived food item and the toxin (s) they produce
Heliotropium sp.	**S-06**	contaminant 1: structured nutlet	10	Species of this genus are known to produce pyrrolizidine alkaloids which are generally present in high concentrations in the seeds (El-Shazly & Wink, 2014)
	S-57	contaminant 10: structured nutlet	2	
Euphorbia sp.	**S-06**	contaminant 11: light brown seed	1	Species from the Euphorbia genus are known to produce several toxic or irritative diterpenoid derivatives (Shi, Su, & Kiyota, 2008)
	S-06	contaminant 17: green twigs	2	
Euphorbia nicaeensis subsp. *nicaeensis*	**S-06**	contaminant 12: orange colored seed	1	
Euphorbia myrsinites	**S-53**	contaminant 24: camouflage-like seed	1	
	S-53	contaminant 28: white structures	2	
Nigella sp.	**S-57**	contaminant 12: pale-like structure	1	*Nigella damascena*, belonging to the Nigella genus is known to produce the potentially toxic alkaloid damascenine (Bouguezza, Khettal, Tir, & Boudrioua, 2015). Based on the seed phenotype, the seeds present in the samples do not correspond to *N. damascena* nor to *N. sativa* (Heiss et al., 2011)
Convolvulvus arvensis	**S-57**	contaminant 16: pieces of leaves	n.a.	*Convolvulvus arvensis* is known to produce tropane alkaloids and pyrrolidine alkaloids (Todd, Stermitz, Schultheis, Knight, & Traub-Dargatz, 1995)
	S-53	contaminant 26: rock like pieces		

well as the putative species identification for the cumin and the green anise samples are present in the Supplementary Material (see supplemental data). The obtained identification success rates for both the top hit method (each locus separately and combined), the similarity method (each locus separately and combined) are mentioned in Table 4. With the top hit method with *matK*, the BLAST search results identified 60.4% (29/48) to species level, 33.3% (16/48) to genus level, and 6.3% (3/48) to family level. The top hit BLAST search results identified 47.2% (17/36) to species level, 38.9% (14/36) to genus level, 5.6% (2/36) to family level and 8.3% (3/36) remained unidentified when the *trnH-psbA* locus was used. The results obtained for *matK* for the similarity method identified 52.1% (25/48) to species level, 39.6% (19/48) to genus level, 8.3% (4/48) to family level. The results obtained for *trnH-psbA* for the similarity method were able to identify 50% (18/36) of the samples to species level, 25% (9/36) to genus level, 19.4% (7/36) to family level and 5.5% (2/36) of the samples remained unidentified due to a low query coverage. Based on the above data it can be stated that for these samples *matK* resulted in a higher identification rate, 81.4% (48/55) compared to the use of the *trnH-psbA* locus which resulted only in a 60% (33/55) positive identification rate. The combined data from both markers using the top hit BLAST sequence similarity search method identified 67.3% (37/55) to species level, 25.5% (14/55) to genus level, 5.5% (3/55) to family level and 1.8% (1/55) could not be identified. Combined data from the similarity method identified 61.8% (34/55) to species level, 29.1% (16/55) to genus level, and 7.3% (3/55) to family level while 1.8% (1/55) could not be identified. As 54 items from the 55 items could be identified, it can be stated that the combined approaches yield a more successful identification rate compared to each locus separately. Moreover, the combined top hit approach resulted in the maximum number of species-level identifications and might therefore be considered the preferential approach. For several of the genus-level identifications obtained by the similarity approach, species-level identifications were generated by the top hit approach. However, in order to minimize the amount of false positive species-level identifications, which might be due to the use of public sequence repositories (Meiklejohn, Damaso, & Robertson, 2019), we decided to use the similarity approach, resulting in more genus-level identifications, to verify for the contamination with a PA/PANOs or other potential phytotoxin producing plant(s).

3.3.3. *Phytotoxin producing contaminants identified in the samples*

Based on the outcome from the similarity approach, the main part of the identified contaminants are plant species or belong to plant genera or families (e.g. Poaceae), not known to produce phytotoxins. However, both PA/PANOs samples that surpassed the by the European union established threshold, contained a nut-like structure which corresponded to a seed of a Heliotropium species (see Table 5). Unfortunately with the current barcodes it was impossible to pinpoint the exact Heliotropium species responsible for the contamination. Nevertheless the entire Heliotropium genus is taken up by the Belgian government published list of plants which are not allowed for use in or as foodstuffs due to their toxicity (Royal decree of 29 August 1997 accessible at https://www.health.belgium.be/en/food/specific-foods/food-supplements-and-enriched-foodstuffs/plants). Moreover, based on the microscopic observations, sample S-06, containing 8510 μg/kg of PA/PANOs, and sample S-57, containing 1653 μg/kg, were found to be contaminated, in one jar, with 14 and 2 *Heliotropium* sp. seeds respectively. In order to verify if such an in absolute mass percentage neglectable contamination, is responsible for these high amounts of PA/PANOs, the commercially acquired *Heliotropium europaeum* seeds were subjected to quantification by LC-MS/MS. The results (see supplemental data) demonstrate that one single seed contained already 20824 μg/kg, which upon extrapolation would require the sample S-06, with a total mass of 40 g, to be contaminated with 16.4 seeds and the samples S-57, with the similar mass of 40 g per jar, with 3.2 seeds. These numbers are quite close to the observed number of Heliotropium seeds present in the samples. However, care has to be taken with such extrapolations since we were only able to identify the genus responsible for the contamination and the level of PA/PANOs production depends on the plant species. Moreover, also the quantity of PA/PANOs produced by one single plant species depends on the environmental circumstances (e.g. soil and water supply) and on the developmental stage (reviewed in Kopp, Abdel-Tawab, & Mizaikoff, 2020).

In addition to a contamination with *Heliotropium* sp. seeds the sample S-06 was also contaminated with twigs and seeds from the Euphorbia genus, also listed on the prohibited plants list (see Table 5). Moreover, a contamination with material from a *Euphorbia* plant was also encountered in the PA/PANOs negative sample S-53. In addition, material originating from the tropane alkaloid and pyrrolidine alkaloid producing *Convolvulus arvensis* was also detected in that sample, as in the sample S-57. Furthermore, this latter sample also contained one single seed from the *Nigella* genus (see Table 5). Based on the phenotypical traits of this seed, it is very unlikely that this represents a contamination with the by the Belgian government forbidden and phytotoxin producing *Nigella damascena* (Heiss, Kropf, Sontag, & Weber, 2011).

4. Conclusion

Taken together, the DNA barcoding data confirmed the presence of plant material originating from a *Heliotropium* species in the samples S-06 and S-57, concurring previously obtained DNA metabarcoding data (Picron et al., 2021). Moreover, the occurrence of one single Heliotropium seed in one jar can result in an amount of PA/PANOs present that surpasses the threshold limits set forward by the European commission (European commission regulation, 2020). The latter is crucial information for the local farmers growing cumin and/or green anise and for the producers upscale the supply chain, to raise awareness of the danger of this type of plants in close proximity to their field.

Additionally, the provided barcoding data also illustrates the danger of relying solely on targeted chemical analysis for phytotoxin detection and clearly demonstrates the added value of molecular methodologies. Indeed, as DNA barcoding is not specifically targeting PA/PANOs producing plant species, also other types of phytotoxin producing plants were identified in these samples, demonstrating the added value of a more open screening approach for an overall assessment for the possible presence of different types of phytotoxin producing species. It would be interesting to assess if molecular techniques such as DNA barcoding (in combination with morphological sorting) and DNA metabarcoding, could be deployed as a more open screening approach, prior to analysis of the sample by a set or subset of different targeted (bio)chemical analyses to identify and quantify the potential phytotoxins present in spice mixtures and other herbal mixtures.

Declaration of Competing Interest

There are no competing interests.

Acknowledgments

The chemical analysis part of this research was funded by the Belgian Federal Public Service of Health, Food Chain Safety and Environment through the contract PASFOODEXTRA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochms.2021.100070.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology, 215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research, 41*(Database issue), D36–D42. https://doi.org/10.1093/nar/gks1195

Bouguezza, Y., Khettal, B., Tir, L., & Boudrioua, S. (2015). Damascenine induced hepatotoxicity and nephrotoxicity in mice and in vitro assessed human erythrocyte toxicity. *Interdisciplinary Toxicology, 8*, 118–124. https://doi.org/10.1515/intox-2015-0018

Bua, D. G., Annuario, G., Albergamo, A., Cicero, N., & Dugo, G. (2016). Heavy metals in aromatic spices by inductively coupled plasma-mass spectrometry. *Food additives & contaminants Part B, Surveillance, 9*, 210–216. https://doi.org/10.1080/19393210.2016.1175516

CBOL Plant Working Group. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 12794–12797. https://doi.org/10.1073/pnas.0905845106

Chung, S., & Lam, A. (2017). Investigation of pyrrolizidine alkaloids including their respective N-oxides in selected food products available in Hong Kong by liquid chromatography electrospray ionisation mass spectrometry. *Food additives & contaminants Part A, Chemistry, analysis, control, exposure & risk assessment, 34*, 1184–1192. https://doi.org/10.1080/19440049.2017.1319579

Di Bella, G., Ben Mansour, H., Ben Tekaya, A., Beltifa, A., Potortì, A. G., Saiya, E., … Lo Turco, V. (2018). Plasticizers and BPA Residues in Tunisian and Italian Culinary Herbs and Spices. *Journal of Food Science, 83*, 1769–1774. https://doi.org/10.1111/1750-3841.14171

Dusemund, B., Nowak, N., Sommerfeld, C., Lindtner, O., Schäfer, B., & Lampen, A. (2018). Risk assessment of pyrrolizidine alkaloids in food of plant and animal origin. *Food and Chemical Toxicology, 115*, 63–72. https://doi.org/10.1016/j.fct.2018.03.005

Edgar, J. A., Colegate, S. M., Boppré, M., & Molyneux, R. J. (2011). Pyrrolizidine alkaloids in food: A spectrum of potential health consequences. *Food Additives & Contaminants Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment, 28*, 308–324. https://doi.org/10.1080/19440049.2010.547520

El-Shazly, A., & Wink, M. (2014). Diversity of pyrrolizidine alkaloids in the boraginaceae structures, distribution, and biological properties. *Diversity, 6*, 188–282. https://doi.org/10.3390/d6020188

European commission regulation 2020/2040, COMMISSION REGULATION (EU) 2020/2040 of 11 December 2020 amending Regulation (EC) No 1881/2006 as regards maximum levels of pyrrolizidine alkaloids in certain foodstuffs: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020R2040&from=EN [accessed September 2021].

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., … Barrett, S. C. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One, 3*, Article e2802. https://doi.org/10.1371/journal.pone.0002802

Galvin-King, P., Haughey, S. A., & Elliot, C. T. (2018). Herb and spice fraud; the drivers, challenges and detection. *Food Control., 88*, 85–97. https://doi.org/10.1016/j.foodcont.2017.12.031

Ghorbani, A., Saeedi, Y., & de Boer, H. J. (2017). Unidentifiable by morphology: DNA barcoding of plant material in local markets in Iran. *PLoS One, 12*, Article e0175722. https://doi.org/10.1371/journal.pone.0175722

Gonzalez, M. A., Baraloto, C., Engel, J., Mori, S. A., Pétronelli, P., Riéra, B., & Chave, J. (2009). Identification of Amazonian trees with DNA barcodes. *PLoS One, 4*, Article e7483. https://doi.org/10.1371/journal.pone.0007483

Hariram, U., & Labbé, R. (2015). Spore prevalence and toxigenicity of Bacillus cereus and Bacillus thuringiensis isolates from U.S. retail spices. *Journal of Food Protection, 78*, 590–596. https://doi.org/10.4315/0362-028X.JFP-14-380

Heiss, A. G., Kropf, M., Sontag, S., & Weber, A. (2011). Seed Morphology of Nigella s.l. (Ranunculaceae): identification, Diagnostic Traits, and Their Potential Phylogenetic Relevance. *International Journal of Plant Sciences, 172*, 267–284. https://doi.org/10.1086/657676

Howard, C., Lockie-Williams, C., & Slater, A. (2020). Applied barcoding: the practicalities of DNA testing for herbals. *Plants (Basel, Switzerland), 9*, 1150. https://doi.org/10.3390/plants9091150

Huang, W. J., Li, F. F., Liu, Y. J., & Long, C. L. (2015). Identification of Crocus sativus and its Adulterants from Chinese Markets by using DNA barcoding technique. *Iranian Journal of Biotechnology, 13*, 36–42. https://doi.org/10.15171/ijb.1034.

Kakar, F., Akbarian, Z., Leslie, T., Mustafa, M. L., Watson, J., van Egmond, H. P., … Mofleh, J. (2010). An outbreak of hepatic veno-occlusive disease in Western afghanistan associated with exposure to wheat flour contaminated with pyrrolizidine alkaloids. *Journal of Toxicology, 2010*, Article 313280. https://doi.org/10.1155/2010/313280

Kaltner, F., Rychlik, M., Gareis, M., & Gottschalk, C. (2020). Occurrence and risk assessment of pyrrolizidine alkaloids in spices and culinary herbs from various geographical origins. *Toxins, 12*, 155. https://doi.org/10.3390/toxins12030155

Kool, A., de Boer, H. J., Krüger, A., Rydberg, A., Abbad, A., Björk, L., & Martin, G. (2012). Molecular identification of commercialized medicinal plants in southern Morocco. *PLoS One, 7*, Article e39459. https://doi.org/10.1371/journal.pone.0039459

Kopp, T., Abdel-Tawab, M., & Mizaikoff, B. (2020). Extracting and analyzing pyrrolizidine alkaloids in medicinal plants: A review. *Toxins, 12*, 320. https://doi.org/10.3390/toxins12050320

Kowalska, G. (2021). The safety assessment of toxic metals in commonly used herbs, spices, tea, and coffee in Poland. *International Journal of Environmental Research and Public Health, 18*, 5779–5797. https://doi.org/10.3390/ijerph18115779

Kress, W. J., & Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One, 2*, Article e508. https://doi.org/10.1371/journal.pone.0000508

Krishnaraj, S., Gunaseelan, R., Arunmozhi, M., & Sumandiran, C. S. P. (2020). Supply chain perspective and logistics of spices in Indian retail industry. *Materials Today: Proceedings, 2020*. https://doi.org/10.1016/j.matpr.2020.02.681

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., … Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 2923–2928. https://doi.org/10.1073/pnas.0709936105

Li, B., Gao, F., Liu, X., Pan, J., & Liu, L. (2021). Herbal tea-induced hepatic veno-occlusive disease in China: A case description. *Quantitative Imaging in Medicine and Surgery, 11*, 3882–3889. https://doi.org/10.21037/qims-20-48.

Mädge, I., Gehling, M., Schöne, C., Winterhalter, P., & These, A. (2020). Pyrrolizidine alkaloid profiling of four Boraginaceae species from Northern Germany and implications for the analytical scope proposed for monitoring of maximum levels. *Food Additives & Contaminants Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment, 37*, 1339–1358. https://doi.org/10.1080/19440049.2020.1757166

Man, A., Mare, A., Toma, F., Curticăpean, A., & Santacroce, L. (2016). Health Threats from contamination of spices commercialized in romania: risks of fungal and bacterial infections. *Endocrine, Metabolic & Immune Disorders Drug Targets, 16*, 197–204. https://doi.org/10.2174/1871530316666160823145817

Mathot, A. G., Postollec, F., & Leguerinel, I. (2021). Bacterial spores in spices and dried herbs: The risks for processed food. *Comprehensive Reviews in Food Science and Food Safety, 20*, 840–862. https://doi.org/10.1111/1541-4337.12690

Meiklejohn, K. A., Damaso, N., & Robertson, J. M. (2019). Assessment of BOLD and GenBank - Their accuracy and reliability for the identification of biological materials. *PloS One, 14*, Article e0217084. https://doi.org/10.1371/journal.pone.0217084

Modupalli, N., Naik, M., Sunil, C. K., & Natarajan, V. (2021). Emerging non-destructive methods for quality and safety monitoring of spices. *Trends in Food Science & Technology, 108*, 133–147. https://doi.org/10.1016/j.tifs.2020.12.021

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics (Oxford, England), 24*, 1757–1764. https://doi.org/10.1093/bioinformatics/btn322

Newmaster S. G., Fazekas A. J., Ragupathy S. (2006). DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. Canadian Journal of Botany, 84, 335–44. https://doi.org/10.1139/b06-047.

Newmaster, S. G., Grguric, M., Shanmughanandhan, D., Ramalingam, S., & Ragupathy, S. (2013). DNA barcoding detects contamination and substitution in North American herbal products. *BMC Medicine, 11*, 222. https://doi.org/10.1186/1741-7015-11-222

Pang, X., Liu, C., Shi, L., Liu, R., Liang, D., Li, H., … Chen, S. (2012). Utility of the trnH-psbA intergenic spacer region and its combinations as plant DNA barcodes: A meta-analysis. *PLoS One, 7*, Article e48833. https://doi.org/10.1371/journal.pone.0048833

Pantano, L., La Scala, L., Olibrio, F., Galluzzo, F. G., Bongiorno, C., Buscemi, M. D., & Vella, A. (2021). QuEChERS LC-MS/MS screening method for mycotoxin detection in cereal products and spices. *International Journal of Environmental Research and Public Health, 18*, 3774. https://doi.org/10.3390/ijer.ph18073774

Parvathy, V. A., Swetha, V. P., Sheeja, T. E., Leela, N. K., Chempakam, B., & Sasikumar, B. (2014). DNA barcoding to detect chilli adulteration in traded black pepper powder. *Food Biotechnology, 28*, 25–40. https://doi.org/10.1080/08905436.2013.870078

Parvathy, V. A., Swetha, V. P., Sheeja, T. E., & Sasikumar, B. (2015). Detection of plant-based adulterants in turmeric powder using DNA barcoding. *Pharmaceutical Biology, 53*, 1774–1779. https://doi.org/10.3109/13880209.2015.1005756

Pickova, D., Ostry, V., Malir, J., Toman, J., & Malir, F. (2020). A review on mycotoxins and microfungi in spices in the light of the last five years. *Toxins, 12*, 789. https://doi.org/10.3390/toxins12120789

Picron, J. F., Philippe, F., Dubrulle, N., Van Hoeck, E., Giraud, N., Goscinny, S., & Vanhee, C. (2021). Targeted LC-MS/MS combined with multilocus DNA metabarcoding as a combinatory approach to determine the amount and the source of pyrrolizidine alkaloids contamination in popular cooking herbs, seeds, spices and leafy vegetables. *Food Additives & Contaminants Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment, 38*, 962–977. https://doi.org/10.1080/19440049.2021.1889043

Raclariu, A. C., Heinrich, M., Ichim, M. C., & de Boer, H. (2018). Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication. *Phytochemical Analysis: PCA, 29*, 123–128. https://doi.org/10.1002/pca.2732

Reinholds, I., Pugajeva, I., Bavrins, K., Kuckovska, G., & Bartkevics, V. (2017). Mycotoxins, pesticides and toxic metals in commercial spices and herbs. *Food Additives & Contaminants Part B, Surveillance, 10*, 5–14. https://doi.org/10.1080/19393210.2016.1210244

Rivera-Pérez, A., Romero-González, R., & Garrido Frenich, A. (2021). Persistent organic pollutants (PCBs and PCDD/Fs), PAHs, and plasticizers in spices, herbs, and tea - A review of chromatographic methods from the last decade (2010–2020). *Critical Reviews in Food Science and Nutrition, 1–21*. https://doi.org/10.1080/10408398.2021.1883546

Royal decree of 29 August 1997 on the manufacture and trade of foods composed of or containing plants or plant preparations (M.B. 21.XI.1997). https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/consolidated_version_rd_29_august_1997_v22-12-2020_fr.pdf [accessed 18/08/2021].

Schramm, S., Köhler, N., & Rozhon, W. (2019). Pyrrolizidine alkaloids: Biosynthesis, biological activities and occurrence in crop plants. *Molecules (Basel, Switzerland), 24*, 498. https://doi.org/10.3390/molecules24030498

Schrenk, D., Gao, L., Lin, G., Mahony, C., Mulder, P., Peijnenburg, A., … These, A. (2020). Pyrrolizidine alkaloids in food and phytomedicine: Occurrence, exposure, toxicity, mechanisms, and risk assessment - A review. *Food and Chemical Toxicology: An International Journal published for the British Industrial Biological Research Association, 136*, Article 111107. https://doi.org/10.1016/j.fct.2019.111107

Shi, Q. W., Su, X. H., & Kiyota, H. (2008). Chemical and pharmacological research of the plants in genus Euphorbia. *Chemical Reviews, 108*, 4295–4327. https://doi.org/10.1021/cr078350s

Tel-zur, N., Abbo, S., Myslabodski, S., & Mizrahi, D. Y. (1999). Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera hylocereus and selenicereus (Cactaceae). *Plant Molecular Biology Reporter, 17*, 249–254. https://doi.org/10.1023/A:1007656315275

Timpano, E. K., Scheible, M., & Meiklejohn, K. A. (2020). Optimization of the second internal transcribed spacer (ITS2) for characterizing land plants from soil. *PLoS One, 15*, Article e0231436. https://doi.org/10.1371/journal.pone.0231436

Todd, F. G., Stermitz, F. R., Schultheis, P., Knight, A. P., & Traub-Dargatz, J. (1995). Tropane alkaloids and toxicity of Convolvulus arvensis. *Phytochemistry, 39*, 301–303. https://doi.org/10.1016/0031-9422(94)00969-z

Wallis, Y., & Morrell, N. (2011). Automated DNA Sequencing. In B. Theophilus, & R. Rapley (Eds.), *PCR Mutation Detection Protocols* (vol 688). Humana Press. https://doi.org/10.1007/978-1-60761-947-5_12.

Waśkiewicz, A., Beszterda, M., Bocianowski, J., & Goliński, P. (2013). Natural occurrence of fumonisins and ochratoxin A in some herbs and spices commercialized in Poland analyzed by UPLC-MS/MS method. *Food Microbiology, 36*, 426–431. https://doi.org/10.1016/j.fm.2013.07.006

Zhu, L., Zhang, C. Y., Li, D. P., Chen, H. B., Ma, J., Gao, H., … Lin, G. (2021). Tu-San-Qi (Gynura japonica): The culprit behind pyrrolizidine alkaloid-induced liver injury in China. *Acta Pharmacologica Sinica, 42*, 1212–1222. https://doi.org/10.1038/s41401-020-00553-9