WILEY | Hindawi

Research Article

Combining a *COI* Mini-Barcode with Next-Generation Sequencing for Animal Origin Ingredients Identification in Processed Meat Product

Yanyi Pan ,[1] Deyi Qiu,[1] Jian Chen,[1] and Qiaoyun Yue [1,2]

[1]*Zhongshan Customs Technology Center, 2, Zhongshan 6 Road, Zhongshan 528403, Guangdong, China*
[2]*School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China*

Correspondence should be addressed to Qiaoyun Yue; 779721036@qq.com

For revealing animal species in complex or adulterated processed meat product, we presented a method combining a novel cytochrome oxidase I (*COI*) mini-barcode with next-generation sequencing (NGS), which identifies various animal species (swine, bovine, Caprinae, and some of fish, shrimp, and poultry) accurately and efficiently in processed meat products. We designed a universal primer based on 140 sequences from 51 edible animal species. A mixture of 12 species raw meat samples were identified with the clone sequencing and also with a mini-barcode- (136 bp) combined NGS method, respectively. The mini-barcode of these 12 species was 100% identical to the target species sequence by Sanger sequencing. Compared to the clone sequencing method, the NGS method is superior in accuracy, sensitivity, and detection efficiency. Various edible animal species were identified in the species level both in the mixed samples and the 7 heavily processed food products. Moreover, some *unlabeled species and dubious contamination were* detected as well.

1. Introduction

The concern of the adulteration and misbranding of food has increased since the fraudulently mislabeling of beef event occurred in Europe in 2013 despite the strict food safety supervision system [1]. Fraudulent behaviors concerning food products are not rare, especially in the fish adulteration [2, 3]. Therefore, scientific research and techniques for food authenticity verification remains challenging and should be closely monitored, as it is highly related to public health, economic interest, religion, and even lifestyle [4]. Techniques for species identification of animal-originated heavily processed food are a vital part of food authenticity although further development is needed.

In general, morphology-based traditional species identification techniques are not suitable for processed and cooked meat products. During the last two decades, DNA molecular detection methods, such as PCR gene chip and molecular fingerprint techniques, have emerged as successful approaches for species identification [5, 6]. Such techniques are based on informative DNA fragments and PCR techniques (real-time PCR, PCR-SSCP, and PCR-RAPD) and have been adopted in food authenticity practices recently [7–10]. However, these procedures may be not very suitable for the heavily processed products with multiple species ingredients. Most of them can identify the specific species but not the unknown or potential ones. In addition, the difficulty in designing specific primers or time-consuming is hard to satisfy the accurate and rapid identification needs.

New opportunities in species identification are available due to the rapid development of sequencing techniques and continuously updated and refined DNA barcodes databases. In 2003, a DNA barcode repository for all species was proposed by Paul Hebert and coworkers [11]. The 658 bp fragment of the cytochrome oxidase I (*COI*) gene located in the mitochondrial genome was selected for species identification based on seven classes and eight orders of animal

species. With the exception of Cnidaria, an average of 98% of the species studied showed 0%–2% intraspecies divergence and 11.3% interspecies divergence [12]. Hebert proposed that a single *COI* gene was sufficient to differentiate among multiple species and acted as a global bioidentification system for most fauna. In the past decade, the 658 bp *COI* fragment was widely used for the species identification of different species animals from arthropod to mammals [13, 14] and has been employed as a species identification tool for food authentication, including mislabeled food products [15].

Another advanced molecular technique, next-generation sequencing (NGS) has been widely applied in the studies of whole genome sequencing of animals and plants, transcriptome sequencing, and metagenomic investigations such as microorganism diversity [16]. One of the NGS technology amplicon sequencing combining with a gene marker shows promise in food authentication practices as it can meet the need for detecting target organism ingredients or revealing unknown ingredients. In recent years, it has been applied in the authentication of plant and animal origins in traditional Chinese medicines or in animal DNA mixture but rarely in processed meat products [17, 18].

However, the application of it in authentication of processed meat products is facing some challenges. First, utilizing longer reads is not available as the technology provided by Roche 454 was phased out in 2016. Presently, read lengths for NGS were only up to 2×300 bp [19]. Second, high conservation of universal primers should be designed for mixture meat products. Third, the target gene's accuracy of species identification should be assured [20].

For considering the first challenge, some mini-DNA markers (16S and 12S ribosomal RNA genes and cytochrome b (*Cytb*) gene) were designed [21–23]. They have been used in the detection for degraded raw mixtures of rodent samples (mammalian and avian) [18, 24] but rarely used for the species identification of processed food. The primers' universality of 16S rRNA should be further verified for more meat species since some proposed that it showed less variation than the *COI* and *Cytb* genes [25]. DNA barcode based on *COI* has been shown to be effective in the identification of edible meat species [26]. The wide use of *Cytb*'s universal primers is attributed to their early availability long before *COI* became popular [27]. The use of a mini-barcode for *COI* was first mentioned in 2008, where they were used for species identification of major eukaryote groups (animals, fungi, plants, and protists) [28]. However, this mini-barcode is limited since its priming sites are not sufficiently conserved to cover a broad range of taxa [29]. Some mini-barcodes of *COI* were explored and applied in species identification for fish [15], gut contents of metazoan [30], and marine invertebrates [31].

For considering the second and third challenges, the mini-barcodes of *COI* would be a better marker for its accuracy and universality, but so far, there has not been a standard mini-barcode adopted for diverse meat product authentication [32]. On the contrary, for detecting a unique ingredient or revealing unknown ingredients from mixture, cloning sequencing may be a good way, but various working procedures and false-negative results are hard to be avoided. The NGS (amplicon sequencing) may be a better way for its deep sequencing presenting ingredients more fully.

Therefore, for revealing real animal ingredients in meat product, our study aims to develop a method of using the NGS technology with a new mini-barcode for identification of various animal species in a mixture.

2. Materials and Methods

2.1. Primer Design. Hundred forty ($n = 140$) *COI* gene sequences for animals were downloaded from BOLD (Barcode of Life Data Systems database, http://www.boldsystems.org/) and NCBI (National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/). The 140 sequences were analyzed for searching a target region that presented a smaller intraspecies gap and a bigger interspecies divergence among 51 traditional economic species including fish ($n = 25$), shrimp ($n = 6$), poultry ($n = 10$), swine ($n = 3$), bovine ($n = 3$), and Caprinae ($n = 4$). The priming sites were moderately conservative and universally enough for efficient amplification across all selected species. After multiple sequence alignments by MEGA v. 6.0.6 [33], a short fragment (136 bp) meeting the purpose was selected for further analysis.

The most frequently adopted LCO1490 [34] was selected as our forward primer and named as mini-*COI*-F:5′-GGTCAACAAATCATAAAGATATTGG-3′. The reverse primer was designed and named as mini-*COI*-R: 5′-ACTATAAAGAAGATTATTACAAAGGC-3′. The further analyses of the primers (primer dimer, hairpin structure, and specificity) were evaluated by Oligo 7.0 [35] and BLAST (Basic Local Alignment Search Tool, https://blast.ncbi.nlm.nih.gov/Blast.cgi). The species differentiation among mini-barcode was presented by neighbor-joining trees (NJ tree). We used the P-distance model in MEGA v. 6.0.6 to build the NJ trees of target DNA sequences from 51 commercial animal species.

2.2. Sample Collection. Twelve kinds of fresh meat and seven heavily processed real-food products were purchased from a local grocery market in Zhongshan, China. Fresh meat samples in the study were composed of twelve species: bovine (cattle and water buffalo), swine (domestic pig), Caprinae (sheep), gallus (domestic chicken), partridge, fish (grass carp, silver carp, blue scad, tile fish, and pomfret), and shrimp (prawn). Heavily processed real-food products were meatballs (beef ball, pork ball, fish ball, and shrimp ball), modulation beefsteak, sausage, and Chinese sausage.

In order to show the species composition of each sample clearly, samples were divided into four groups: A, B, C, and D. Group A consists of the 12 raw meat species but PCR and sequenced separately, Groups B and C are mixture of the above 12 species, Group B was sequenced with traditional clone picking, while Group C was sequenced with NGS, and Group D consists of 7 real heavily proceeded food products. Specific information of sample grouping and detection

TABLE 1: The name, composition, and sequencing methods for each sample in the four groups (A, B, C, and D).

Sample group	Sample composition	Sequencing method
Group A (S1–S12)	12 different species treated separately	PCR-direct sequencing
Group B (B1–B3)[a]	Sample of the 12 different species mixture	Cloning sequencing
Group C (C1–C3)[a]	Same as B1–B3	Next-generation sequencing
Group D (D1–D7)	7 commercial products	Next-generation sequencing

[a]B1–B3 and C1–C3 are parallels for the corresponding group.

methods is summarized in Table 1. Group A was set for confirming the animal species. 12 raw meat samples were identified separately by sequencing PCR products. For comparing results between clone sequencing and NGS method in detecting unique species in mixed samples, we set Group B and Group C. Meat species are the same in Groups A, B, and C. For Group B, the 12 raw meat samples were mixed equal-weight together and exposed to clone sequencing. Group C was prepared in the same way as Group B but treated with the NGS method (amplicon sequencing). Group D was set for testing the efficiency of the newly designed mini-barcode combining the NGS method: animal origins from 7 heavily processed real-food products were authenticated by the developed NGS method in the study.

2.3. DNA Extraction and PCR Amplification.
In Group A, 30 mg of muscle tissue was cut from the inner section of the 12 raw meat samples with 12 different knifes in order to avoid cross contamination. For Group B, samples were taken as in Group A but mixed the 12 species in equal-weight together. Preparation for Group C was identical to Group B. In Group D, 200 g of each sample was taken and minced together to paste with pure water in 1 : 2 (g/ml) proportion and 100 mg mince mixture was used for DNA extraction. Three parallel tests were prepared for each sample. Tools such as surgical scissors, meat grinder, and stirrer were sterilized with high temperature and pressure (121°C, 0.25 MPa, 20 min) and UV irradiation (for 1 hour) before use. DNA was extracted using the TIANamp Genomic DNA Kit.

PCR amplification of the target fragments in 50 μL total volume included 10 × PCR buffer (5 μL), 20 μmol/L forward primer (1 μL), 20 μmol/L reverse primer (1 μL), 2.5 mM dNTP (2 μL), 5 U/μL Ex-*Taq* (1 μL), template DNA (50–200 ng/μL and 1–3 μL), and ddH$_2$O (up to 50 μL). The optimum annealing temperature was decided through gradient temperature PCR. The primers were synthetized by Thermo Fisher Scientific Corporation (Suzhou, China).

2.4. PCR-Direct Sequencing to Confirm Target Fragment.
First, the 12 selected animal species were identified by the normal length 658 bp DNA barcode for affirming their species [36], and then mini-barcode amplifications were applied to Groups A, B, C, and D. To verify consistency between expected species identification and actual results with mini-barcode, PCR amplification products with mini-barcode from Group A were sent to Thermo Fisher Scientific Corporation for Sanger sequencing. The obtained sequences were analyzed using Nucleotide BLAST (the website is mentioned above) to confirm their corresponding species. The sequences were used as a reference dataset for further species identification of the meat mixture by clone sequencing and NGS.

2.5. Clone Sequencing.
Purification, transformation, and clone selection were conducted to the PCR products which were obtained with mini-barcode in Group B. To detect positive clones, we utilized the white-blue plaque selection technique to identify 130 positive clones from each parallel (repeated independent experiments) before Sanger sequencing.

2.6. Next-Generation Sequencing Flow and Data Processing.
For both Groups C and D, we used an Agilent 2100 to assess the DNA concentration of the target fragments (after PCR). For those qualified samples (quality of DNA fragments >2.0 μg), the following procedures of DNA libraries preparation (PCR-free in the NGS process), cluster generation, and high throughput sequencing were conducted by BGI Genomics. The Illumina HiSeq 4000 platform and PE150 bp sequencing strategy were chosen. Additionally, we chose Q20 (>85%) to insure that over 85% of sequences were limited to a sequencing error rate of 1%.

The raw sequence data were obtained after sequencing and format conversion. The clean reads were automatically produced such as redundant reads with N base, adapter, or polybase (with the same base over ten) were filtered by readfq (a read conducting system). High quality reads were used for creating tags based on the overlap relation of the reads. Tags with high variation were produced using FLASH [37]. Overlap regions of reads should be over 15 bp with less than a 1% mismatch rate, and reads without overlap were trimmed. The clean tags within 97% identity were clustered to generate different Operational Taxonomic Units (OTUs) by USEARCH (a unique sequence analysis tool) [38]. Due to the lack of a specialized database for animal, species annotation of OTUs was achieved using BLAST in NCBI.

To demonstrate the correlations among different samples, principal component analysis (PCA) was processed in the R statistical environment (v3.0.3) [39] with the ade4 package [40]. Heatmaps were generated using the gplots package based on clustering analysis. The clustering analysis method is calculating sequence distance and sequence abundance by the Euclidean algorithm and complete algorithm.
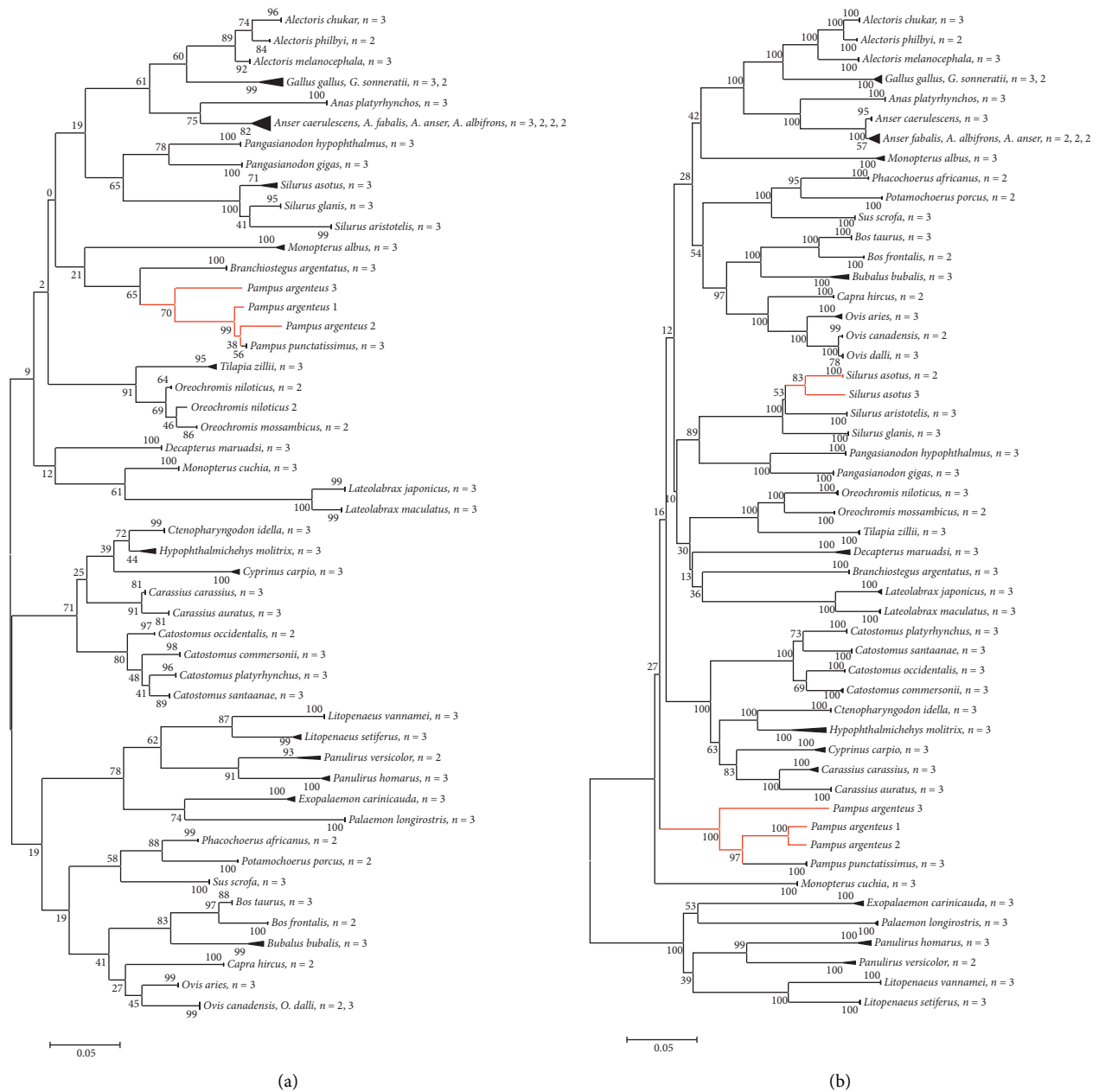
Figure 1: Phylogenetic trees of mini-barcode and standard barcode of *COI* are presented in (a) and (b), respectively.

3. Results

3.1. The Species Differentiation between Mini-Barcode.
The NJ trees of the mini-barcode of *COI* (136 bp) and the standard barcode (658 bp) from 51 species are shown in Figures 1(a) and 1(b). According to the mini-barcode NJ tree (see Figure 1(a)), 36 species can be differentiated on species level (the interspecies divergence was over 2%), while the others (15 species from the genus of *Gallus, Anser, Oreochromis, Carassius*, and *Ovis*) can only be differentiated on the genus level (the interspecies divergence was less than 2% in the same genus). Except for *Pampus argenteus* (marked with bold branches), the intraspecies divergence of the others was within 2%. According to the standard barcode NJ tree (see Figure 1(b)), 41 species can be differentiated on species level.

The others (10 species from the genus of *Gallus, Anser*, and *Alectoris*) can only be differentiated on the genus level. The intraspecies divergence of three species of fish (marked with bold branches) was more than 2%.

Comparing NJ trees from Figures 1(a) with 1(b), the species sequence divergence of mini-barcode (136 bp) is less than the standard one, such as 5 species cannot be identified by the mini-barcode, but the mini-barcode still retains the basic species divergence information of the standard barcode for most species were identified to species level successfully.

3.2. Optimization of Amplification Condition and Verification of Target Fragments.
In order to detect the target species as comprehensive as possible, PCR amplification conditions

Table 2: Detection of organisms in prescribed meat mixture.

Method	Cloning sequencing			NGS		
Sample ID	B1	B2	B3	C1	C2	C3
Total no. of detected clones/tags	127	122	125	2,012,052	2,054,514	2,042,428
Pangasianodon hypophthalmus (S1)	√	√	√	√	√	√
Alectoris chukar (S2)	√	√	√	√	√	√
Bos taurus (S3)	√	√	√	√	√	√
Bubalus bubalis (S4)		√		√	√	√
Litopenaeus vannamei (S5)	√	√		√	√	√
Gallus gallus (S6)				√	√	√
Ctenopharyngodon idella (S7)	√		√	√	√	√
Ovis aries (S8)	√	√	√	√	√	√
Decapterus maruadsi (S9)	√			√	√	√
Sus scrofa (S10)	√			√	√	√
Branchiostegus argentatus (S11)				√	√	√
Pampus chinensis (S12)				√	√	√
*Ovis canadensis**	√		√	√	√	√
*Anser canagica**				√	√	√
*Sarcophaga josephi**				√	√	√
*Sarcophaga antilope**				√		
*Homo sapiens**				√	√	√
*Opisthoplatia orientalis**					√	√
*Blattella germanica**					√	
*Hypophthalmichthys nobilis**				√	√	

*The species were detected beyond expectation.

were optimized. 12 animal originated meat species were amplified separately and subjected to gradient temperature PCR. We found that the best annealing temperature for the species in Group A was 48°C with high quality fragments, while S12 (*Pampus chinensis*) amplified best at 43°C. Sometimes, 48°C and 45°C would lead to negative results for S12 in agarose gel electrophoresis. Annealing temperature at 43°C could not assure the identification accuracy for all the studied species. Thus, to guarantee identification accuracy and test detection sensitivity of the NGS, the annealing temperature for mixture was set at 48°C. Finally, amplification was conducted under the following conditions: initial denaturation at 94°C for 3 min, followed by 30 cycles of 98°C for 10 s, 48°C for 30 s, and 68°C for 1 min, and a final extension at 68°C for 7 mins.

All species in Group A were amplified successfully by the optimized amplification protocol, and the amplified fragments were verified through PCR-direct sequencing. Sequencing quality was well with minute miscellaneous peaks in the sequencing signal map. After alignment in BLAST, the mini-barcode (136 bp) of each species showed 100% identity with the target species sequence.

3.3. Detection of Raw Meat Mixture by Clone Sequencing.
95.9% of clones for each parallel test were sequenced successfully and identified to species level. The species identified on a level of species were shown in Table 2 (see in Group B). Totally 9 species were detected accurately, while three species (S6, S11, and S12) were undetected. The negative detection of S12 was expected since 48°C was not suitable for it. However, the undetection of S6 (*Gallus gallus*) and S11 (*Branchiostegus argentatus*) was unexpected. Moreover, other three species

(S4, S9, and S10) were detected with only one clone each. Three species (S1, S2, and S3) were detected with 10–77 clones in each parallel test. Despite the equal quality of muscle tissue for the different species, sequence abundance among different species varied significantly. In addition, we unexpectedly found two clone sequences were from nuclear gene (43U chromosome 9 sequence) and 99% identical with *Ovis canadensis* sequence (CP011894.1) in the GenBank database. Perhaps, there is a certain conservative fragment between the nuclear sequence of *O. canadensis* and the primers that cause a false amplification.

3.4. Detection of Raw Meat Mixture by NGS.
After reads quality filtering, 96.56% (2.39 million clean reads) were retained for each sample. 98.97% OTU sequences were over 98% identical to the target species sequence from the GenBank database.

The species identified on a level of species are presented in Table 2 (see in Group C). All 12 species were detected though NGS, while only 9 species were detected with the clone sequencing. S12 (*Pampus chinensis*) was detected with a low sequence abundance of 3 to 50 (<0.01% abundance of total OTU sequences). For remaining species, the abundance rate range was from 0.04% to 28.37%. Four species (S1, S2, S3, and S4) were detected, respectively, with 28.37%, 28.20%, 27.12%, and 8.97% abundance. The unexpected species, *O. canadensis,* was also detected with 0.22% abundance.

In the case of deep sequencing (4-5 billion of single reads for each test), several unexpected organisms were detected such as duck, bighead carp, fly, cockroach, and human. The abundance rate for duck was 0.08%, and the remaining species was less than 0.01%.

3.5. Detection of Processed Meat Products by NGS. After sequencing, at least 1.37 million clean reads pairs (the qualified reads before splicing) were obtained with the read utilization ratio of 94.16% for a sample. The labeled ingredients for each sample were detected (see Table 3). Most organisms were identified to the species level (was over 98% identical to the target species sequence) except for two species of fish (*Upeneus* and Trachystoma petardi) and one species of shrimp (Metapenaeopsis). Moreover, unlabeled ingredients were found meanwhile (the species in bold in Table 3). The unlabeled species of *Sus scrofa* from the YW sample (minced fillet) was detected with 28.3% abundance of total OTU sequences, and each of the remaining unlabeled species was about 0.002%–5.2%. Two dubious contamination ingredients (*Bos taurus* and *Trachysalambria curvirostris*) were found in most samples with low abundance rates between 0.01 and 0.08%.

Since dubious contaminants existed among the samples, principal component analysis and clustering analysis were used to evaluate the independence between samples. Based on the OTU sequences and species annotation, the PCA plot and heatmap executed in R (v3.0.3) are presented in Figure 2. In the PCA plot, two principal components (PC1 and PC2) were obtained, and PC1 accounted for 46.52% of global variable information and PC2 for 21.64%. In heatmap, main ingredients of each sample were presented over 1% relative abundances and shown in blue color. The distinguishment of the 21 tests' compositions was shown by the PCA plot and a cluster tree (vertical clustering) in heatmap.

The PC1 (*x*-coordinate) intuitively reflecting XW (shrimp balls), NJW (beef tendon balls), and NP (modulation beefsteak) samples was clearly distinguished. The relative gathering of tests in LC (Chinese sausage), SNNY (Pork balls), XC (sausage), and YW (fish balls) was likely attributed to the pork ingredient. The cluster tree in heatmap showed that all parallel tests within each sample grouped together. Therefore, there is a high repeatability of detection results, and relative independence of each sample was distinct for each sample.

4. Discussion

Generally, the mini-barcode showed similar identification between species to the standard *COI* barcode based on the analyses of 51 species. Some related species (*Bos taurus and Bubalus bubalis*) would be distinguishable from each other with the mini-barcode. However, it was difficult to differentiate between species within the same genus, for example, *Ovis*. The alignment of their mini-barcode showed 99.26% identity of *O. aries* to the other two species and 98.53% identity between the other two species. However, there was also high identity (97.27% on average) among each of them by the *COI* standard barcode (658 bp), but the identity was still lower than intraspecies divergence [11]. To distinguish related species well, a combined utilization of gene markers may be a better choice to get more reference information.

For the detection of unexpected sequence of *Ovis canadensis* (with 0.22% abundance rate) from nuclear genes, it

was most likely that the mini-barcode was too short to be strictly specific for the *COI* gene. Since *O. canadensis* is native to arid environments of western North America and was not considered a meat animal, it was impossible to get this species in the market from Zhongshan, China. This situation only occurred for the sheep species in the study, and *O. aries* was detected with 1.39% abundance rates.

According to the clone sequencing results of the mixed samples, the negative detection of S12 was likely related to low amplification efficiency at a higher annealing temperature of 48°C. Interestingly, some species (S4, S5, S7, and S8) which exhibited high amplification efficiency in independent amplification tests showed low detection rates. Though major organisms of mixed samples could be detected by clone sequencing, false-negative or low detection rates for some species were frequent. The variation in detection results of clone sequencing appeared to be related to the PCR amplification bias [18, 41], the random selection of positive clones, and relative clone abundance for detection.

The NGS methods showed high sensitivity in detecting individual species in mixture samples. The S12 species which lacked favorable amplification efficiency and some unexpected ingredients or dubious contaminants would be identified to species level. Of the dubious contaminants, duck and bighead carp were possibly incidentals attached to the other meat samples when bought in the market, whereas the fly and cockroach contaminants were likely to be obtained from laboratory working environment or processing facility because we are amplifying thousands of flies and cockroaches genes each year in the lab. The detection of human may be attributed to lack of attention to operation procedures or safety protocol, occasionally contaminated with the operator's saliva for talking while operation. Though contaminants could be detected at high sensitivity, there was a distinct difference in the detection abundance between most of the prescribed species (0.04% to 28.37% abundance rate) and the contaminative organisms (less than 0.01%).

In comparing results between Group B and Group C, the NGS method in the study shows a superior ability to detect unique species in mixed samples. The advantages of this method include (1) simultaneously identifying multiple organisms in heavily processed products, (2) discovering unexpected ingredients or contaminants, (3) highly efficient, and (4) high accuracy of sequencing and high utilization of data. On the contrary, some defects of the NGS (amplicon sequencing) were found in the study as well: (1) it was unable to avoid the amplification efficiency disparity among the species in a mixture, but it is worth mentioning that other NGS technologies such as the whole genome shotgun can get around of this problem [42]; (2) in a mixed sample, there was no linear relation between detected abundance and mass of ingredient. Considering PCR amplification bias and copies of multicopy gene disparity in different organisms [29, 43], modifications of the amplicon sequencing or use of other NGS methods could be better considered for further quantifying ingredients [44].

All the labeled species and even some unlabeled species were detected in 7 heavily processed meat products in Group

TABLE 3: The detected ingredients and the labeled ingredients of the seven commercial samples.

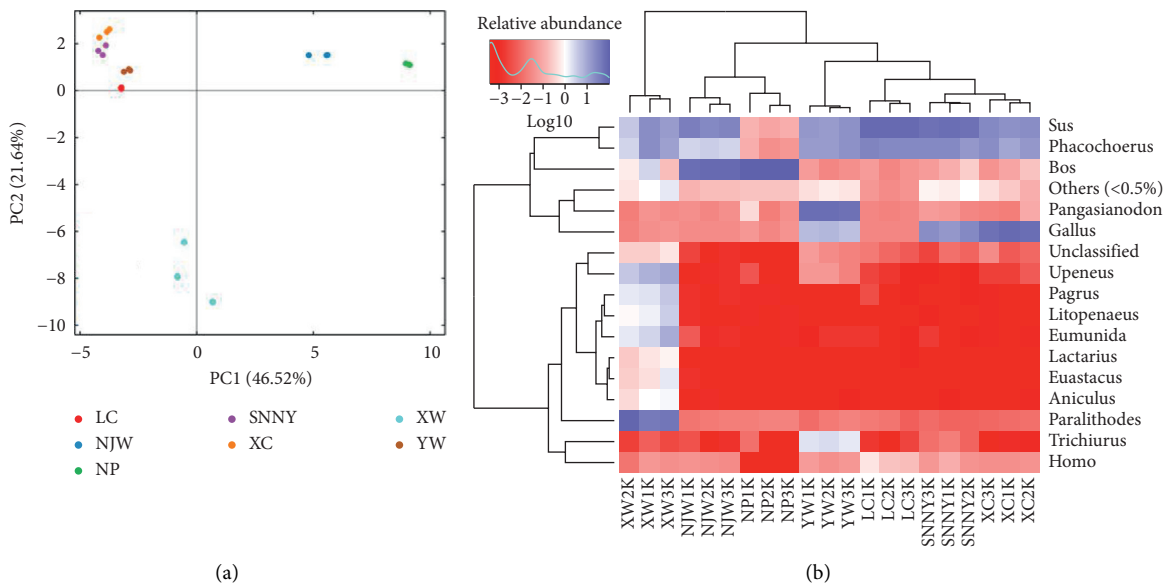| Sample ID	Sample name	Labeled ingredients	Detected ingredients	Detected species
NJW (1, 2, 3)	Beef tendon balls	Beef Pork	Beef (66.2%) Pork (29.4%)	*Bos taurus, Bubalus bubalis* *Sus scrofa*
XW (1, 2, 3)	Shrimp balls	Minced fillet	Fish (6.8%)	Upeneus, *Parargyrops edita, Trachystoma petardi, Nemipterus japonicus, Lactarius lactarius, Atule mate, Konosirus punctatus, Plotosus lineatus*
		Shrimp mince	Shrimp (76.3%)	*Trachysalambria curvirostris, Litopenaeus vannamei, Metapenaeopsis*
SNNY (1, 2, 3)	Pork balls	Pork Chicken Duck	Pork (74.7%) Chicken (20.9%) Duck (0.04%)	*S. scrofa* *Gallus gallus* *Anas poecilorhyncha*
		Minced fillet	Fish (2.9%)	*T. petardi, Pangasianodon hypophthalmus, Trichiurus japonicus, Neobythites sivicola*
YW (1, 2, 3)	Fish balls	Minced fillet	Fish (65.4%) Pork (28.3%) Chicken (5.2%) Duck (<0.1%)	*P. hypophthalmus, Trichiurus japonicus, T. petardi, Scomber japonicus, T. curvirostris* *S. scrofa* *G. gallus* *A. poecilorhyncha*
XC (1, 2, 3)	Sausage	Chicken	Chicken (57.5%) Pork (32.7%) *Fish (1.1%)* *Chukar partridge (0.1%)*	*G. gallus* *S. scrofa* *T. petardi, P. hypophthalmus* *Alectoris chukar*
LC (1, 2, 3)	Chinese sausage	Pork	Pork (92.3%) *Fish (6.0%)* *Chicken (<0.1%)*	*S. scrofa* *T. petardi, P. hypophthalmus* *G. gallus*
NP (1, 2, 3)	Modulation beefsteak	Beef	Beef (99.5%) *Pork (0.2%)* *Fish (0.2%)* *Chicken (<0.1%)*	*B. taurus, B. bubalis* *S. scrofa* *P. hypophthalmus* *G. gallus*

FIGURE 2: Organism components and abundance differentiated the seven commercial samples.

D by the NGS method. Some ingredients with abundances rates between 0.01% and 0.08% were likely to be contaminated. Some species with low amplification would also be detected with low abundance. Therefore, to differentiate between the contamination and the ingredients added in the products, conducting a control sample with known species

from Group D could be better to avoid biased results from the contamination. Because it is invalid to conduct a blank control sample, sequences cannot be amplified for a small amount of DNA concentration from environment. Moreover, we found that to prevent cross contamination, the common method of avoiding contamination was not enough. Rather, meticulous attention to detail during operation and clean utensils should be assured. Utensils were only used after cleaning with a DNA scavenger solution, and any DNA in the environment should be degraded by ultraviolet before starting the preparation of each sample.

5. Conclusion

We developed a method of combining a *COI* mini-barcode with NGS for species identification of animal-originated species from raw meat mixture and heavily processed meat products. By using the novel mini-barcode, edible species (fishes, shrimps, birds, and mammals) could be identified accurately. The identification of the known species and unexpected species presented the favorable accuracy, sensitivity, and efficiency of the method. In order to avoid food contamination, cautious operation, clean utensil, and noncontaminated environments should be assured. In order to distinguish between closely related species and reduce the impact of competitive efficiency, a combined utilization of different DNA markers should be considered. The application of this method offers advantages to food safety procedures by offering *a better protocol for identifying animal-originated ingredients and composition of processed meat products.*

Data Availability

All sample data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] H. Lehr and Syntesa, "Communicating food safety, authenticity and consumer choice. Field experiences," *Recent Patents on Food, Nutrition & Agriculture*, vol. 5, no. 1, pp. 19–34, 2013.

[2] A. Galal-Khallaf, A. Ardura, K. Mohammed-Geba, Y. J. Borrell, and E. Garcia-Vazquez, "DNA barcoding reveals a high level of mislabeling in Egyptian fish fillets," *Food Control*, vol. 46, pp. 441–445, 2014.

[3] B. A. Maralit, R. D. Aguila, M. F. H. Ventolero, S. K. L. Perez, D. A. Willette, and M. D. Santos, "Detection of mislabeled commercial fishery by-products in the Philippines using DNA barcodes and its implications to food traceability and safety," *Food Control*, vol. 33, no. 1, pp. 119–125, 2013.

[4] N. Z. Ballin, F. K. Vogensen, and A. H. Karlsson, "Species determination-can we detect and quantify meat adulteration?" *Meat Science*, vol. 83, no. 2, pp. 165–174, 2009.

[5] N. Z. Ballin, "Authentication of meat and meat products," *Meat Science*, vol. 86, no. 3, pp. 577–587, 2010.

[6] A. K. Lockley and R. G. Bardsley, "DNA-based methods for food authentication," *Trends in Food Science & Technology*, vol. 11, no. 2, pp. 67–77, 2000.

[7] N. Haider, I. Nabulsi, and B. Al-Safadi, "Identification of meat species by PCR-RFLP of the mitochondrial *COI* gene," *Meat Science*, vol. 90, no. 90, pp. 490–493, 2012.

[8] J. H. Kuo, H. W. Tsuei, Z. L. Jia et al., "Identification of ingredient in mullet roe products by the real-time PCR method," *Food Analytical Methods*, vol. 11, no. 4, pp. 1–9, 2017.

[9] H. Özpınar, G. Tezmen, E. I. Gökç, and I. H. Tekiner, "Detection of animal species in some meat and meat products by comparatively using DNA microarray and real time PCR methods," *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, vol. 19, no. 2, pp. 245–252, 2013.

[10] C. Sarri, C. Stamatis, T. Sarafidou et al., "A new set of 16S rRNA universal primers for identification of animal species," *Food Control*, vol. 43, no. 5, pp. 35–41, 2014.

[11] P. D. N. Hebert, A. Cywinska, S. L. Ball, and J. R. Dewaard, "Biological identifications through DNA barcodes," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. 1512, pp. 313–321, 2003.

[12] P. D. N. Hebert, S. Ratnasingham, and J. R. De Waard, "Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. suppl_1, pp. S96–S99, 2003.

[13] P. D. N. Hebert, M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis, "Identification of birds through DNA barcodes," *PLoS Biology*, vol. 2, no. 10, p. e312, 2004.

[14] D. Qiu, C. E. Cook, Q. Yue et al., "Species-level identification of the blowfly *Chrysomya megacephala* and other Diptera in China by DNA barcoding," *Genome*, vol. 60, no. 2, pp. 158–168, 2017.

[15] S. Shokralla, R. S. Hellberg, S. M. Handy, I. King, and M. Hajibabaei, "A DNA mini-barcoding system for authentication of processed fish products," 2015, https://www.nature.com/articles/srep15894.

[16] Y. Ge, S. Yu, W. Shi, F. Yang, and Q. Yin, "Next-generation sequencing for discovery of the microorganism and pathogenic bacteria in the source water in winter," *Research Journal of Biotechnology*, vol. 12, no. 9, pp. 66–74, 2017.

[17] X. Cheng, X. Su, X. Chen et al., "Biological ingredient analysis of traditional chinese medicine preparation based on high-throughput sequencing: the story for liuwei dihuang wan," *Scientific Reports*, vol. 4, no. 1, 2015.

[18] F. Bertolini, M. C. Ghiond, E. D'Alessandro, C. Geraci, V. Chiofalo, and L. Fontanesi, "A next generation semiconductor based sequencing approach for the identification of meat species in DNA mixtures," *PLoS One*, vol. 10, no. 4, Article ID e0121701, 2015.

[19] M. Quail, M. E. Smith, P. Coupland et al., "A tale of three next generation sequencing platforms: comparison of ion torrent, pacific biosciences and illumina MiSeq sequencers," *BMC Genomics*, vol. 13, no. 1, p. 341, 2012.

[20] V. Elbrecht, F. Leese, and M. Hajibabaei, "Can DNA-based ecosystem Assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an

innovative metabarcoding protocol," *PLoS One*, vol. 10, no. 7, Article ID e0130324, 2015.

[21] M. Galan, M. Pagès, and J. F. Cosson, "Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples," *PLoS One*, vol. 7, no. 11, Article ID e48374, 2012.

[22] A. O. Karlsson and G. Holmlund, "Identification of mammal species using species-specific DNA pyrosequencing," *Forensic Science International*, vol. 173, no. 1, pp. 16–20, 2007.

[23] T. Kitano, K. Umetsu, W. Tian, and M. Osawa, "Two universal primer sets for species identification among vertebrates," *International Journal of Legal Medicine*, vol. 121, no. 5, pp. 423–427, 2007.

[24] A. O. Tillmar, B. Dell'Amico, J. Welander, and G. Holmlund, "A universal method for species identification of mammals utilizing next generation sequencing for the analysis of DNA mixtures," *PLoS One*, vol. 8, no. 12, Article ID e83761, 2013.

[25] S. S. Tobe and A. Linacre, "DNA typing in wildlife crime: recent developments in species identification," *Forensic Science, Medicine, and Pathology*, vol. 6, no. 3, pp. 195–206, 2010.

[26] A. Galimberti, F. De Mattia, A. Losa et al., "DNA barcoding as a new tool for food traceability," *Food Research International*, vol. 50, no. 1, pp. 55–63, 2013.

[27] T. D. Kocher, W. K. Thomas, A. Meyer et al., "Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers," *Proceedings of the National Academy of Sciences*, vol. 86, no. 16, pp. 6196–6200, 1989.

[28] I. Meusnier, G. A. Singer, J. F. Landry, D. A. Hickey, P. D. N. Hebert, and M. Hajibabaei, "A universal DNA mini-barcode for biodiversity analysis," *BMC Genomics*, vol. 9, no. 1, p. 214, 2008.

[29] J. Piñol, G. Mir, P. Gomez-Polo, and N. Agustí, "Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods," *Molecular Ecology Resources*, vol. 15, no. 4, pp. 819–830, 2015.

[30] M. Leray, J. Y. Yang, C. P. Meyer et al., "A new versatile primer set targeting a short fragment of the mitochondrial *COI* region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents," *Frontiers in Zoology*, vol. 10, no. 1, p. 34, 2013.

[31] J. Machida, C. Meyer, M. Parker, and H. Hawk, "Redesign of PCR primers for mitochondrial cytochromecoxidase subunit I for marine invertebrates and application in all-taxa biotic surveys," *Molecular Ecology Resources*, vol. 13, no. 5, pp. 851–861, 2013.

[32] M. Staats, A. J. Arulandhu, B. Gravendeel et al., "Advances in DNA metabarcoding for food and wildlife forensic species identification," *Analytical and Bioanalytical Chemistry*, vol. 408, no. 17, pp. 4615–4630, 2016.

[33] S. Kumar, M. Nei, J. Dudley, and K. Tamura, "MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences," *Briefings in Bioinformatics*, vol. 9, no. 4, pp. 299–306, 2008.

[34] O. Folmer, M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek, "DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates," *Molecular Marine Biology and Biotechnology*, vol. 3, no. 3, pp. 294–299, 1994.

[35] W. Rychlik, "OLIGO 7 primer analysis software," in *PCR Primer Design. Methods in Molecular Biology*™, A. Yuryev, Ed., vol. 402, pp. 35–59, Humana Press, Totowa, NJ, USA, 2007.

[36] M. Hajibabaei, G. A. C. Singer, P. D. N. Hebert, and D. A. Hickey, "DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics," *Trends in Genetics*, vol. 23, no. 4, pp. 167–172, 2007.

[37] T. Magoc and S. L. Salzberg, "FLASH: fast length adjustment of short reads to improve genome assemblies," *Bioinformatics*, vol. 27, no. 21, pp. 2957–2963, 2011.

[38] R. C. Edgar, "UPARSE: highly accurate OTU sequences from microbial amplicon reads," *Nature Methods*, vol. 10, no. 10, pp. 996–998, 2013.

[39] R. B. Dessau and C. B. Pipper, ""R"—project for statistical computing," *Ugeskrift for Laeger*, vol. 170, no. 5, pp. 328–330, 2008.

[40] S. Dray and A. Dufour, "The ade4 package: implementing the duality diagram for ecologists," *Journal of Statistical Software*, vol. 22, no. 4, pp. 1–20, 2007.

[41] Q. Peng, R. V. Satya, M. Lewis, P. Randad, and Y. Wang, "Reducing amplification artifacts in high multiplex amplicon sequencing by using molecular barcodes," *BMC Genomics*, vol. 16, no. 1, p. 589, 2015.

[42] F. Ripp, C. Krombholz, Y. Liu et al., "All-food-seq (AFS): a quantifiable screen for species in biological samples by deep DNA sequencing," *BMC Genomics*, vol. 15, no. 1, p. 639, 2014.

[43] S. W. Kembel, M. Wu, J. A. Eisen, and J. L. Green, "Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance," *PLoS Computational Biology*, vol. 8, no. 10, Article ID e1002743, 2012.

[44] A. C. Thomas, B. E. Deagle, J. P. Eveson, C. H. Harsch, and A. W. Trites, "Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material," *Molecular Ecology Resources*, vol. 16, no. 3, pp. 714–726, 2016.