

1 Species-specific InDel markers for authentication of the
2 Korean herbs *Zanthoxylum schinifolium* and *Zanthoxylum*
3 *piperitum*

4
5
6
7
8
9
10
11
12
13
14
15

Yonguk Kim¹, Jawon Shin¹, Seung-Sik Cho², Yong-Pil Hwang³, and Chulyung Choi^{1,*}

¹ Jeonnam Institute of Natural Resources Research, Jangheung-gun, Jeollanam-do 59339, Korea

² Department of Pharmacy, College of Pharmacy, Mokpo National University, Muan-gun, Jeollanam-do 58554, Korea

³ Department of Pharmaceutical Engineering, International University of Korea, Jinju-si, Gyeongsangnam-do 52833, Korea

* Correspondence: blockstar@hanmail.net; Tel.: +82-61-860-2620

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

***Corresponding Author:**

Dr. Chulyung Choi

Jeonnam Institute of Natural Resources Research, 288

Woodland-gil, Anyang-myeon, Jangheung-gun, Jeonnam, 59338, Republic of Korea

Tel : +82-61-860-2620

Fax : +82-61-864-7105

44

45 **Abstract**

46

47 **Objective:** *Zanthoxylum schinifolium* and *Zanthoxylum piperitum* are the sources of the well-
48 known traditional Korean herbal medicines “sancho” (prickly ash) and “chopi” (Korean
49 pepper), respectively. Sancho and chopi are often indiscriminately mixed due to the similar
50 appearance of the herbal materials when used as spices and herbal medicines. Moreover,
51 commercial sancho and chopi products often contain adulterants, which could compromise
52 drug efficacy and safety.

53 **Methods:** In this study, we developed hypervariable InDel markers to distinguish between
54 sancho and chopi products by comparing the complete chloroplast genome sequences of four
55 *Zanthoxylum* species deposited in NCBI GenBank.

56 **Results:** Comparative analyses of the nucleotide diversity (P_i) of these *Zanthoxylum*
57 genomes revealed four hypervariable divergent sites (*trnH-psbA*, *psbZ-trnG*, *trnfM-rps14*,
58 and *trnF-ndhK*) with $P_i > 0.025$ among 520 windows. Of these four regions, including two
59 genic and two intergenic regions, only *psbZ-trnG* yielded accurate PCR amplification results
60 between commercial sancho and chopi products from the Korean herbal medicine market. We
61 therefore selected *psbZ-trnG*, an InDel-variable locus with high discriminatory powers, as a
62 candidate DNA barcode locus.

63 **Conclusion:** This InDel marker could be used as a valuable, simple, and efficient tool for
64 identifying these medicinal herbs, thereby increasing the safety of these spices and herbal
65 materials in the commercial market.

66

67 **Keywords:** chloroplast genome; chopi; InDel markers; nucleotide diversity; sancho;
68 *Zanthoxylum piperitum*; *Zanthoxylum schinifolium*

69 Introduction

70

71 *Zanthoxylum* plants are thorny, deciduous shrubs and small trees with dense foliage, prickly
 72 trunks, and branches bearing edible fruits and leaves with a strong, pungent taste resembling
 73 lemon, anise, or mint. These plants, belonging to the Rutaceae family, comprise 250 species
 74 that are native to warm temperate and subtropical regions (Sun and Duan 1996; Epifano et al.
 75 2011). While members of *Zanthoxylum* are commonly found in the Himalayan region, they
 76 also occur in Central, South, Southeast, and East Asia, and some species are found in
 77 America and Africa (Arun and Paridhavi 2012; Patiño et al. 2008; Appelhans et al. 2018).
 78 Many of these species are used as traditional medicines to treat human and animal diseases in
 79 Africa, Asia, and South America (Negi et al. 2011; Supabphol and Tangjitjareonkun 2014).
 80 Among these, *Z. schinifolium* (“sancho” in Korean, “huajiao” in Chinese) and *Z. piperitum*
 81 (“chopi” in Korean) are important traditional medicinal herbs in Korea and China (Kim et al.
 82 2000; Ko and Han 1996; Yang 2008) that are also used as spices. *Korean Flora* describes
 83 seven *Zanthoxylum* species from Korea, including *Z. schinifolium* var. *inermis* (Nakai) T.B.
 84 Lee, *Z. piperitum* f. *pubescens* (Nakai) W.T. Lee, *Z. schinifolium* Siebold & Zucc., *Z.*
 85 *piperitum* (L.) DC., *Z. schinifolium* f. *microphyllum* (Nakai) W.T. Lee, *Z. planispinum*
 86 Siebold & Zucc., and *Z. coreanum* Nakai (Lee 2003; Chung 1957; Nakai 1930). *Z.*
 87 *schinifolium* and *Z. piperitum* are the most ancient cultivated trees in southern regions of
 88 Korea. These trees are of major agricultural importance as sources of spices and traditional
 89 medicines (Ko and Han 1996; Yang 2008). Whereas *Z. schinifolium* thorns are arranged in an
 90 alternating pattern on branches, *Z. piperitum* thorns are arranged in an opposing pattern (Lee
 91 2003). However, it is difficult to visually discriminate between dried and powdered herbal
 92 products made from *Z. schinifolium* and *Z. piperitum*. Furthermore, commercial herbal
 93 products commonly include a mixture of *Z. schinifolium* and *Z. piperitum* tissue. Since these
 94 plants are used in herbal medicines and as health food supplements, accurate methods should
 95 be developed to identify and characterize the two species.

96 Several studies have focused on developing molecular markers that can discriminate
 97 between various *Zanthoxylum* species used as traditional medicines in different countries.
 98 These molecular markers include amplified fragment length polymorphism (AFLP) markers
 99 for distinguishing between *Z. acanthopodium* and *Z. oxyphyllum* (Gupta and Mandi 2013),
 100 sequence-related amplified polymorphism (SRAP) markers (Feng et al. 2015), internal

transcribed spacer (ITS) rDNA-specific markers (Kim et al. 2015), and ISSR markers (Feng et al. 2015). However, the use of these markers is limited by population dynamics and their reproducibility as diagnostic markers in the field. With the advancement of next-generation sequencing (NGS) technology, chloroplast (cp) genome assembly has become more accessible than Sanger sequencing. The development of molecular markers has also become cost effective through comparisons of cp genomes. The complete cp genomes of several *Zanthoxylum* species have been sequenced by de novo assembly using a small amount of whole genome sequencing data (Lee et al. 2016; Liu and Wei 2017; 2017).

In the current study, we compared the published cp genome sequences of four *Zanthoxylum* species and searched for species-specific regions with hypervariable nucleotide diversity among members of *Zanthoxylum*. Our aim was to develop interspecies insertion and deletion (InDel) markers that could be used to discriminate between *Z. schinifolium* and *Z. piperitum* and prevent the indiscriminate mixing of these two materials in commercial products.

Results

Comparative analysis of the cp genomes of various *Zanthoxylum* species

To investigate the level of sequence divergence among *Zanthoxylum* cp genomes, we performed a comparative analysis of three *Zanthoxylum* cp genomes with mVISTA using the annotated *Z. piperitum* sequence as a reference (Figure 1). The cp genomes of the *Zanthoxylum* species showed high sequence similarity, with identities of <90% in only a few regions, pointing to a high level of conservation among the cp genomes. However, sliding window analysis using DnaSP detected highly variable regions in the *Zanthoxylum* cp genomes (Figure 2). We calculated the nucleotide diversity (P_i) values in the four *Zanthoxylum* cp genomes as an indicator of divergence at the sequence level. The large single-copy (LSC) regions and small single-copy (SSC) regions were more divergent than the inverted repeat (IR) regions.

Among the 520 windows examined, IR regions were more conserved than LSC and SSC regions, with average P_i values of 0.0015 and 0.0012 for IRa and IRb, respectively (for regions other than those with a P_i value = 0). The P_i values for the LSC regions averaged 0.0082, whereas the SSC regions had a P_i value of 0.0103, and the average P_i value for all regions was 0.00609. Four mutational hotspots in the cp genomes showed markedly higher P_i values (>0.025), including two intergenic regions (*trnH-psbA*, 0.02833; *psbZ-trnG*, 0.05067) and two genic regions (*trnfM-rps14*, 0.05367; *trnF-ndhK*, 0.02667 and 0.03133) (Figure 2). Although the SSC and IR regions were generally highly conserved, the four regions located in the LSC regions were particularly divergent.

Development of InDel markers to discriminate between *Z. schinifolium* and *Z. piperitum*

Based on multiple alignments of complete cp genome sequences, we selected the four most highly variable InDel loci as candidate DNA markers (Table 1). We confirmed these four InDel regions by PCR amplification and sequencing and investigated their utility for discriminating between *Z. schinifolium* and *Z. piperitum* (Figure 3). We produced four markers (ZanID1, ZanID2, ZanID3, and ZanID4) that were specific to *Z. schinifolium* and *Z. piperitum* and were derived from long InDels in the intergenic regions *psbZ-trnG*, *trnfM-rps14*, and *trnF-ndhK*, respectively. ZanID1 was derived from 11-, 3-, and 6-bp InDels in the *trnH-psbA* locus and was specific to *Z. schinifolium* and *Z. piperitum* (Figure 3). ZanID2 was

derived from 19-, 28-, and 5-bp InDels and 40- and 23-bp tandem repeats (TRs) in the *psbZ-trnG* locus. ZanID3 was derived from 19-, 28-, 39-, and 6-bp InDels and a 23-bp TR in *trnM-rps14*. ZanID4 was derived from 32-, 6-, 59-, and 50-bp InDels and an 8-bp TR in *trnF-ndhK*.

Utilization of InDel markers in the Korean food market

To validate the utility of our newly developed markers to identify commercial dried herbal materials, we extracted genomic DNA from powdered or dried samples of seven sancho and six chopi products and amplified them by PCR using the newly developed primers (Table 1). The banding patterns of the *ZanID1* marker revealed that lanes 1, 2, 5, 6, 8, and 10 contained sancho samples, while lanes 3, 4, 9, and 11–16 were identified as chopi (Figure 4A). Different banding patterns were obtained using *ZanID2*, with lanes 2, 3, 5, 7, 8, and 10 containing sancho samples and the other samples identified as chopi (Figure 4B). Interestingly, samples 2, 5, and 7 clearly produced double bands in both species (Figure 4B). Furthermore, analysis of the banding patterns of the *ZanID3* and *ZanID4* markers revealed that lanes 2, 3, 5, 6, 7, 8, and 10 contained sancho samples, while the other samples were identified as chopi (Figure 4C and 4D). These three markers produced considerably different banding patterns, making it difficult to discriminate between sancho and chopi.

To select markers that could accurately discriminate between sancho and chopi, we performed PCR-RFLP analysis of the 13 samples by developing a PCR RFLP test to identify sancho samples. Many taxonomic studies of land plants have been based on the *trnH-psbA* region of cpDNA, as this DNA barcode exhibits high rates of sequence divergence among species. Based on the partial sequences of *trnH-psbA* in the cp genome that are shared between *Z. schinifolium* and *Z. piperitum*, we predicted that the *PleI* restriction enzyme would produce species-specific RFLP patterns and could therefore be used to identify *Z. schinifolium* based on the *trnH-psbA* locus. As shown in Figure 5, the fragment sizes for the two species were as follows: In *Z. schinifolium*, *PleI* produced two fragments (436 bp and 107 bp) from PCR products (lanes 2, 3, 5, 7, 8, and 10) of *trnH-psbA* (543 bp); in *Z. piperitum*, this enzyme did not digest the PCR products of *trnH-psbA* (562 bp). These results indicate that the *psbZ-trnG* marker is suitable for use as a reliable DNA barcoding tool to discriminate between *Z. schinifolium* and *Z. piperitum*.

Discussion

Zanthoxylum formed a phylogenetic group in previous molecular phylogenetic studies (Appelhans et al. 2018; Feng et al. 2015; Kim et al. 2015); however, these studies did not sufficiently resolve the relationships among some of its taxa. These studies were based on the ITS sequences of nuclear ribosomal DNA and the *trnL-trnF*, *matK-trnK*, *atpB*, *atp-rbcL*, and *rbcL* sequences of the cp genome (Sun et al. 2010; Feng et al. 2016; Zhao et al. 2018). Although these regions were considered to be universal DNA barcodes for higher plant, they did not allow to assess usefulness of these loci in barcoding of some taxa. Therefore, advances in NGS technologies have made it possible to sequence whole cp genomes and identify molecular markers. Highly variable markers derived from the cp genomes of different species at the genus level have uncovered many loci that are informative for systematic botany and DNA barcoding research (Dong et al. 2012; Li et al. 2015).

Here, we retrieved the complete cp genome sequences of four *Zanthoxylum* species from the NCBI database and compared species-specific cp diversity in *Zanthoxylum*. We confirmed the variation among species, with an average nucleotide diversity value (P_i) of 0.00609 among the four *Zanthoxylum* species. Although the average P_i value for the SSC region was relatively high (0.0103), high sequence divergence was detected at loci *trnH-psbA* (0.02833), *psbZ-trnG* (0.05067), *trnfM-rps14* (0.05367), and *trnF-ndhK* (0.02667 and 0.03133) in the LSC (0.0082). Indeed, the *trnH-psbA* locus is highly variable in most plants and is known as a universal DNA barcoding region (Kress et al. 2007; Whitlock et al. 2010). We used PCR amplification and sequencing to validate four hypervariable markers to distinguish between *Z. schinifolium* and *Z. piperitum* and to discriminate between sancho and chopi spice materials consumed in the online food market. However, PCR amplification of the four markers produced variable banding patterns, making it difficult to discriminate between sancho and chopi. Therefore, we designed a PCR-RFLP method using *PleI* digestion of the *trnH-psbA* DNA barcoding region, resulting in the production of two fragments (436 bp and 107 bp) only in *Z. schinifolium*. This marker, the ZanID2 marker from the *psbZ-trnG* region, is therefore suitable for discriminating between *Z. schinifolium* and *Z. piperitum* in sancho and chopi. The ZanID2 marker shows high sensitivity and specificity for detecting both sancho and chopi samples. Among the 10 sancho samples examined, three were successfully detected as sancho (30%), whereas three other samples (30%) produced a double band pattern that was clearly detected in both sancho and chopi samples. As expected, these

results confirm the notion that products labeled as sancho that are sold in the spice and herbal medicine market often contain a mixture of sancho and chopi.

Although our results confirm that our newly developed InDel markers can be used to authenticate *Z. schinifolium* and *Z. piperitum* based on available cp genome data, more complete cp genome sequences are needed to comprehensively evaluate these InDel markers in the *Zanthoxylum* genus.

Materials and methods

Comparison of cp genomes and identification of hypervariable loci

All cp genome sequences in plants of the *Zanthoxylum* genus with complete genome sequence information were downloaded from GenBank (Table 1). The sequences were aligned using the Clustal W algorithm from MEGA 7.0 (Kumar et al. 2016) and CLC viewer 8.0 software. The gene distribution patterns and similarities in the *Zanthoxylum* cp genomes were compared and visualized using mVISTA software (<http://genome.lbl.gov/vista/mvista/submit.shtml>) in Shuffle-LAGAN mode with the annotated *Z. piperitum* KT153018 cp genome as a reference. The variability of the aligned genomes was evaluated using the sliding window method with DnaSP ver. 5.0. software (Librado and Rozas 2009). The window size was set to 600 base pairs (bp), the typical length of DNA markers. The step size was set to 300 bp for relatively accurate positioning of hypervariable InDels. Only regions with a nucleotide diversity (P_i) value of > 0.025 were considered. Hypervariable sites and genetic distance in the cp genome were calculated using MEGA 7.0. The InDel events were checked manually based on the aligned sequence matrix.

Sample collection and genomic DNA isolation

Table 2 lists the *Z. schinifolium* and *Z. piperitum* collections from the National Institute of Biological Resources used in this study. Thirteen spice and powdered herbal materials (six sancho and six chopi samples) were purchased from verified local market sources in Korea and China (Table 4). Species identification was performed by the National Institute of Biological Resources, Korea. Prior to total genomic DNA extraction, 50 mg (dry weight) of each sample was added to a tube filled with stainless steel beads (2.38 mm in diameter) from a PowerPlant Pro DNA Isolation Kit (Qiagen, Valencia, CA), and the mixture was

homogenized in a Precellys® Evolution homogenizer (Bertin Technologies). Genomic DNA was extracted from the collected samples using the PowerPlant Pro DNA Isolation Kit according to the manufacturer's instructions.

Development and validation of the InDel molecular marker

To validate interspecies polymorphisms within the cp genomes and to develop DNA genetic markers for identifying these four *Zanthoxylum* species, primers were designed using Primer 3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>), and NCBI Primer-BLAST was performed based on the mutational hotspot regions (hypervariable regions) found in these *Zanthoxylum* cp genomes. PCR amplifications were performed in a reaction volume of 50 µL containing 5 µL 10x Ex Taq buffer (with MgCl₂), 4 µL dNTP mixture (each 2.5 mM), Ex Taq (5 U/µL) (TaKaRa, Japan), 10 ng genomic DNA templates, and 1 µL (10 pM) forward and reverse primers. The mixtures were denatured at 95°C for 5 min and amplified for 40 cycles at 95°C for 30 s, 55°C for 20 s, and 72°C for 30 s, with a final extension at 72°C for 5 min. To detect PCR amplicons, the PCR products were separated by capillary electrophoresis (QIAxcel, Qiagen) using a QIAxcel DNA High Resolution Kit via the 0M500 method (Qiagen). The target DNA was extracted and purified using a MinElute PCR Purification Kit (Qiagen). Purified PCR products were sent to CosmoGenetech for sequencing (Seoul, Korea) with both forward and reverse primers. The sequencing results were analyzed by BLAST searches of the GenBank database. Sequence alignment and data visualization were carried out with CLC sequence viewer 8.0.

RFLP analysis to identify suitable InDel markers

PCR was performed using universal primers for DNA barcodes within *trnH-psbA*: forward primer 5'-GTTATGCATGAACGTAATGCTC-3' and reverse primer 5'-CGCGCATGGTGGATTCACAATCC-3' (approximate product size: 401 bp). The PCR was conducted in a 50 µL reaction mixture containing 5 µL 10x Ex Taq buffer (with MgCl₂), 4 µL dNTP mixture (each 2.5 mM), Ex Taq® (5 U/µL) (TaKaRa, Japan), 10 ng genomic DNA template, and 1 µL (10 pM) forward and reverse primers. The mixtures were denatured at 94°C for 5 min and amplified for 30 cycles of 94°C for 60 s, 55°C for 60 s, and 72°C for 90 s, with a final extension at 72°C for 7 min. The PCR-amplified products were separated by electrophoresis on 1.5% agarose gels for 30 min.

A 2- μ L aliquot of each PCR-amplified *trnH-psbA* product (concentration 0.6 to 1 μ g/ μ L) was digested in 2 μ L of 10x CutSmart buffer, 2 units of PstI restriction enzyme (New England Biolabs, Ipswich, MA; NEB), and 15.8 μ L distilled H₂O in a final volume of 20 μ L, followed by incubation at 37°C for 20 min and inactivation at 65°C for 20 min. The digested fragments were separated by electrophoresis on 1.5% agarose gels stained with ethidium bromide, and the fragment patterns were visualized under UV light.

Abbreviations

InDel: Insert and Deletion; RFLP: Restriction Fragment Length Polymorphism; CP:Chloroplast; Large single-copy: LSC; Small single-copy: SSC; IR: Inverted repeat; PI: pairwise distance

Acknowledgments

This research was supported by the Support Program for Creative Industry Institutes (Commercial Biotechnology Sophistication Platform Construction Program, R0003950), funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

Competing interests

The authors declare that they have no competing interests.

Authors' Contributions

Y.K. and C.C. conceived and designed the experiments; S.-S.C. and Y.-P.H. collected the data; Y.K. and J.S. performed the experiments; Y.K. analyzed the data and wrote the manuscript. All authors read and approved the manuscript.

Funding

Not applicable

References

- Appelhans MS, Reichelt N, Groppo M, Paetzold C, Wen J (2018) Phylogeny and biogeography of the pantropical genus *Zanthoxylum* and its closest relatives in the proto-Rutaceae group (Rutaceae). *Mol Phylogenet Evol* 126: 31-44
- Arun K, Paridhavi M (2012) An ethno botanical phytochemical and pharmacological utilization of widely distributed species *Zanthoxylum*: a comprehensive overview. *Int J Pharm Invent* 2: 24-35
- Chung TH (1957) Korean flora. *Shinjisa* Seoul 1025.
- Dong W, Liu J, Yu J, Wang L, Zhou S (2012) Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PloS one* 7: e35071
- Epifano F, Curini M, Carla Marcotullio M, Genovese S (2011) Searching for novel cancer chemopreventive plants and their products: the genus *Zanthoxylum*. *Curr Drug Targets* 12: 1895-1902
- Feng S, Yang T, Liu Z, Chen L, Hou N, Wang Y, Wei A (2015) Genetic diversity and relationships of wild and cultivated *Zanthoxylum* germplasms based on sequence-related amplified polymorphism (SRAP) markers. *Genet Resour Crop Evol* 62: 1193-1204
- Feng S, Yang T, Li X, Chen L, Liu Z, Wei A (2015) Genetic relationships of Chinese prickly ash as revealed by ISSR markers. *Biologia* 70: 45-51
- Feng S, Liu Z, Chen L, Hou N, Yang T, Wei A (2016) Phylogenetic relationships among cultivated *Zanthoxylum* species in China based on cpDNA markers. *Tree Genet Genomes* 12: 45
- Gupta DD, Mandi SS (2013) Species specific AFLP markers for authentication of *Zanthoxylum acanthopodium* & *Zanthoxylum oxyphyllum*. *J Med Plants* 6: 1-9

- Kim J, Jeong CH, Bae YI, Shim KH (2000) Chemical components of *Zanthoxylum schinifolium* and *Zanthoxylum piperitum* leaves. *Korean J Food Preserv* 7:189-194
- Kim WJ, Ji Y, Lee YM, Kang YM, Choi G, Moon BC (2015) Development of Molecular Markers for the authentication of *Zanthoxylum Pericarpium* by the analysis of rDNA-ITS DNA barcode regions. *Korea J Herbol* 30: 41-47
- Kress WJ, Erickson DL (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: e508
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33: 1870-1874
- Ko YS, Han HJ (1996) Chemical constituents of Korean chopi (*Zanthoxylum piperitum*) and sancho (*Zanthoxylum schinifolium*). *Korean J Food Sci Technol* 28: 19-27
- Lee TB (2003) Coloured flora of Korea. Vol. I. *Hyangmunsa*: Seoul, South Korea
- Lee J, Lee HJ, Kim K, Lee SC, Sung SH, Yang TJ (2016) The complete chloroplast genome sequence of *Zanthoxylum piperitum*. *Mitochondrial DNA Part A* 27: 3525-3526
- Lee HJ, Koo HJ, Lee J, Lee SC, Lee DY, Giang VNL, Kim M, Shim H, Park JY, Yoo KO, Sung SH, Yang TJ (2017) Authentication of *Zanthoxylum* species based on integrated analysis of complete chloroplast genome sequences and metabolite profiles. *J Agric Food Chem* 65: 10350-10359
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452
- Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. *Biol Rev* 90: 157-166

- 378 Liu Y, Wei A (2017) The complete chloroplast genome sequence of an economically
379 important plant, *Zanthoxylum bungeanum* (Rutaceae). *Conserv Genet Resour* 9 : 25-27
380
- 381 Nakai T (1930) Notulæ and Plantas Japoniæ& Koreæ XXXIX. *Shokubutsugaku Zasshi* 44:
382 507-537
383
- 384 Negi JS, Bisht VK, Bh AK, Singh P, Sundriyal RC (2011) Chemical constituents and
385 biological activities of the genus *Zanthoxylum*: a review. *Afr J Pure Appl Chem* 5: 412-416
386
- 387 Patiño LOJ, Prieto RJA, Cuca SLE (2008) *Zanthoxylum* genus as potential source of
388 bioactive compounds. *Bioact Compd Phytomed* 185-218
389
- 390 Sun XW, Duan ZX (1996) Progress in the studies on medicinal plants of the genus
391 *Zanthoxylum* Linn. *Yao xue xue bao= Acta pharm Sin B* 31: 231-240
392
- 393 Sun Y-L, Park W-G, Kwon O-W, Hong S-K (2010) The internal transcribed spacer rDNA
394 specific markers for identification of *Zanthoxylum piperitum*. *Afr J Biotechnol* 9: 6027-6039
395
- 396 Supabphol R, Tangjitjareonkun J (2014) Chemical constituents and biological activities of
397 *Zanthoxylum limonella* (Rutaceae): A review. *Trop J Pharm Res* 13: 2119-2130
398
- 399 Whitlock BA, Hale AM, Groff PA (2010) Intraspecific inversions pose a challenge for the
400 *trnH-psbA* plant DNA barcode. *PloS One* 5: e11533
401
- 402 Yang X (2008) Aroma constituents and alkylamides of red and green huajiao (*Zanthoxylum*
403 *bungeanum* and *Zanthoxylum schinifolium*). *J Agri Food Chem* 56: 1689-1696
404
- 405 Zhao LL, Feng SJ, Tian JY, Wei AZ, Yang T X (2018) Internal transcribed spacer 2 (ITS 2)
406 barcodes: A useful tool for identifying Chinese *Zanthoxylum*. *Appl Plant Sci* 6: e01157
407
408

Figure legends

Figure 1. Comparison of the four *Zanthoxylum* cp genomes using mVISTA. The complete cp genomes of four *Zanthoxylum* species obtained from GenBank were compared, with *Z. piperitum* used as a reference. Purple blocks: conserved genes; pink blocks: conserved non-coding sequences (CNS). White represents regions with high levels of sequence variation among the four *Zanthoxylum* species.

Figure 2. Sliding window analysis of the whole cp genomes of four *Zanthoxylum* species (window size: 600 bp; step size: 300 bp). X-axis: position of the midpoint of a window, y-axis: nucleotide diversity (P_i) of each window.

Figure 3. Gel profiles of fragments amplified from seven *Z. schinifolium* and six *Z. piperitum* samples using four pairs of primers derived from the *trnH-psbA* (A), *psbZ-trnG* (B), *trnfM-rps14* (C), and *trnF-ndhK* (D) loci in the four *Zanthoxylum* cp genomes. M: 100 bp ladder; 1–7: amplicons from *Z. schinifolium* DNA; 8–13: amplicons from *Z. piperitum* DNA.

Figure 4. PCR identification of 16 commercial sancho and chopi products comprising dried or powdered seeds and pericarps. Lanes on 1.5% agarose gel; M: 100 bp ladder; 1–16: purchased commercial sancho and chopi products comprising dried or powdered seeds and pericarps (see Table 4 for details); 16 samples were amplified by PCR using four primer pairs; *ZanID1_F* and R (A), *ZanID2_F* and R (B), *ZanID3_F* and R (C), and *ZanID4_F* and R (D).

Figure 5. PCR-RFLP profiles of partial regions of *trnH-psbA* from all samples digested with *PleI*. Numbers indicate sancho and chopi samples, as described in Table 4.

Table legends

Table 1. Primers used to amplify and sequence four highly variable loci

Table 2. Complete cp genomes of the four *Zanthoxylum* species used in this study

Table 3. *Z. schinifolium* and *Z. piperitum* samples used in this study

Table 4. The 16 commercial sancho and chopi spices and herbal products used in this study

Figure 1.

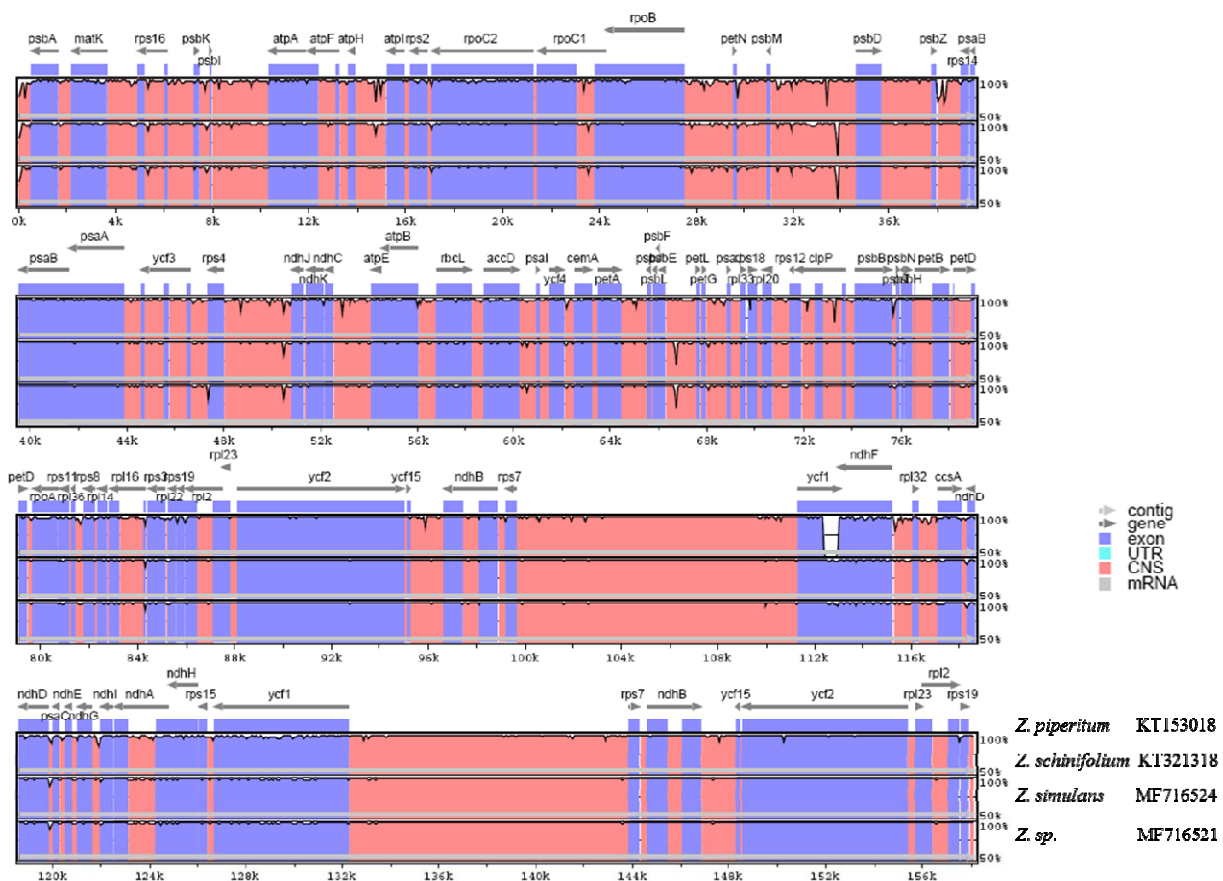
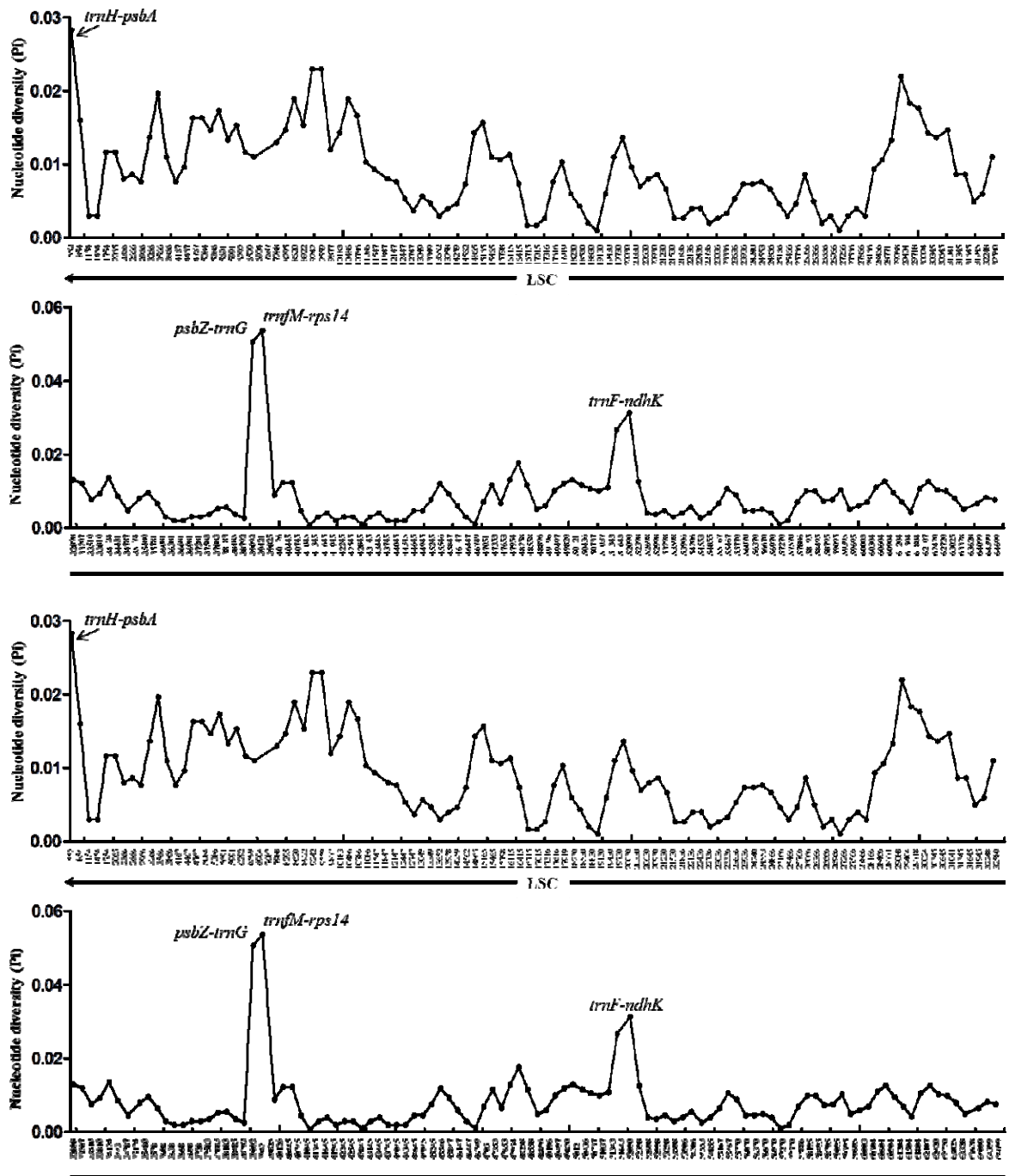


Figure 2.



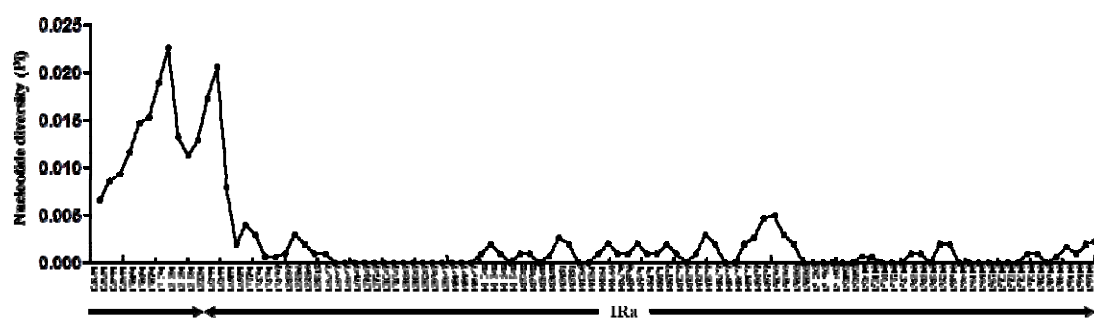


Figure 3.

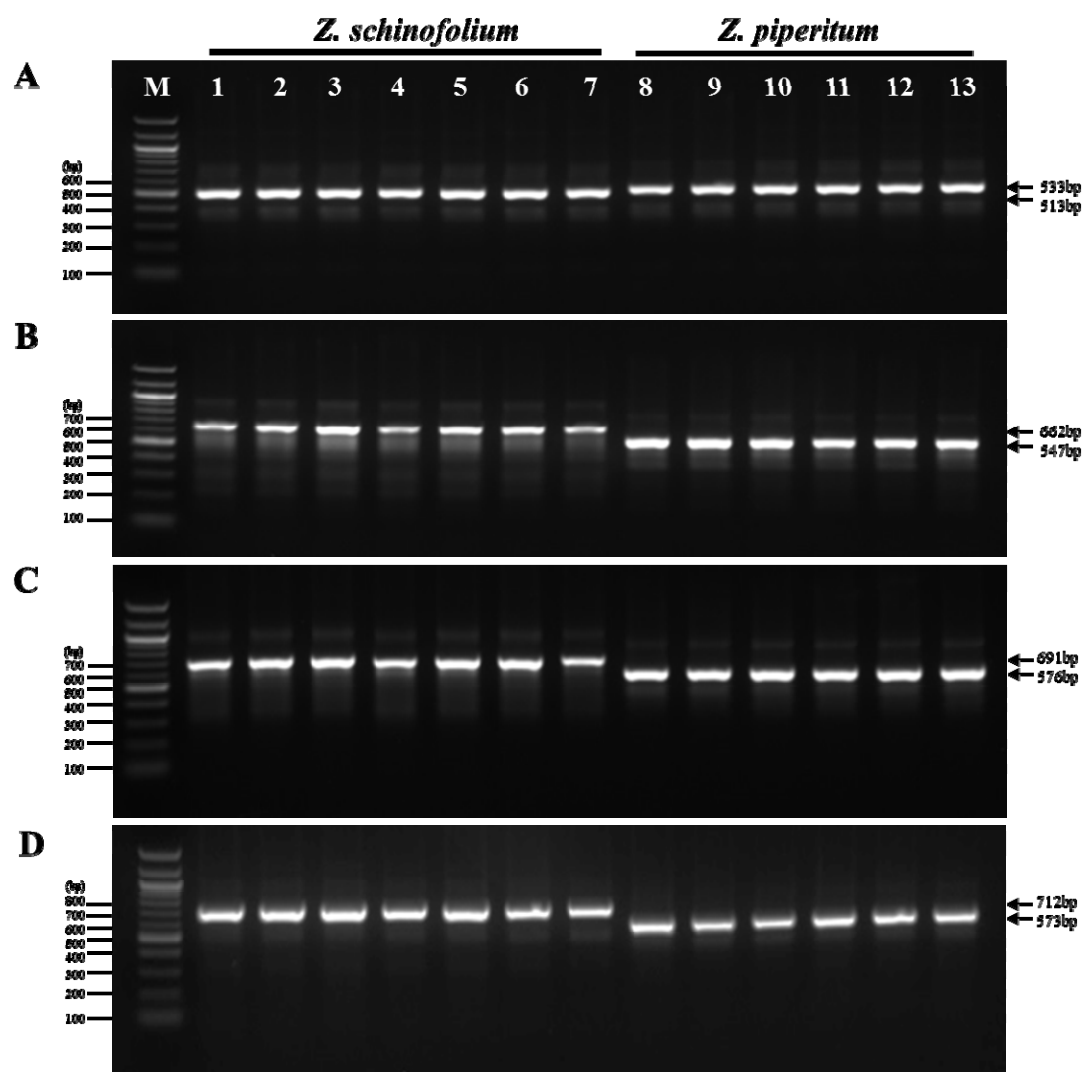
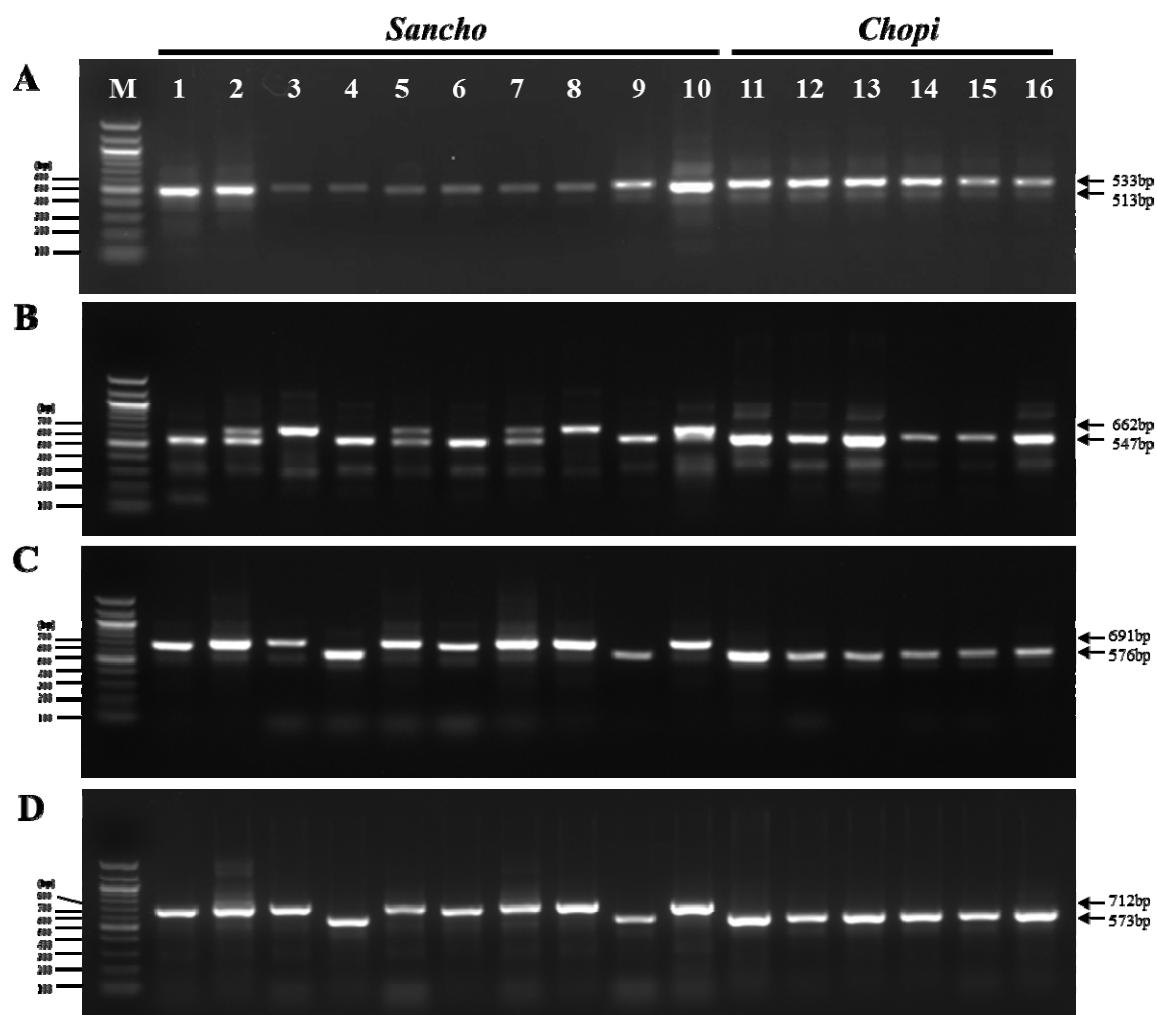
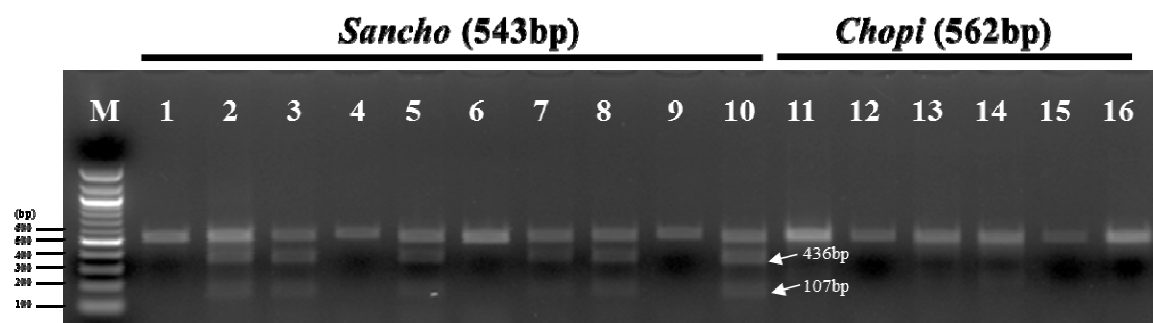


Figure 4.



459 **Figure 5.**



460

461

462 **Table 1.**

Marker name	Locus	Location	Forward primer (5' to 3')	Reverse primer (5' to 3')	Product range (bp)	Mean pairwise distance
ZanID1	<i>trnH-psbA</i>	35...578	GATTCACAATCCACT GCCTTGAT	GCTCATAACTTCCC TCTAGACCT	513-533	0.02833
ZanID2	<i>psbZ-trnG</i>	38882... 39545	GTGGGTATCCTTAAT TCTCTCATC	GATACTCTCTTCAG GGTAATTCCA	547-662	0.05067
ZanID3	<i>trnJ-M-rps14</i>	39040... 39730	GGAGGGATCAAACT TCTGGAAC	AGAACCACTATACT ATCACGGTCA	576-691	0.05367
ZanID4	<i>trnF-ndhK</i>	51439... 52158	GATTTGAGCAAGGA ATCCCCATTG	TCTTCATTTTGACC CGATATTCAA	573-712	0.02667 0.03133

463

464 **Table 2.**

No.	Species	Research group	GenBank accession number	Sequence length (bp)
1	<i>Z. schinifolium</i>	Crop Science and Biotechnology, Department of Plant Sciences, Seoul National University, Korea	KT321318	158,963
2	<i>Z. piperitum</i>		KT153018	158,154
3	<i>Z. simulans</i>	College of Forestry, Northwest A&F University, China	MF716524	158,461
4	<i>Z. sp.</i>		MF716521	158,572

465

466

467 **Table 3.**

No.	Scientific name	Common name	Collection site	Collection number
1	<i>Zanthoxylum schinifolium</i>	Sanchonamu	Hyogok-ri, Ganjeon-myeon, Gurye-gun, Jeollanam-do, Korea	NIBRVP0000278863
2			Galcheon-ri, Seo-myeon, Yangyang-gun, Gangwon-do, Korea	NIBRVP0000448009
3			Daebubuk-dong, Danwon-gu, Ansan-si, Gyeonggi-do, Korea	NIBRVP0000538396
4			Wondeok-ri, Jincheon-eup, Jincheon-gun, Chungcheongbuk-do, Korea	NIBRVP0000538545
5			Sucheol-ri, Punggi-eup, Yeongju-si, Gyeongsangbuk-do, Korea	NIBRVP0000538696
6			Cheongsu-ri, Hangeong-myeon, Jeju-si, Jeju-do, Korea	NIBRVP0000538748
7			Gujora-ri, Irun-myeon, Geoje-si, Gyeongsangnam-do, Korea	NIBRVP0000538812
8	<i>Zanthoxylum piperitum</i>	Chopinamu	Hangeong-myeon, Jeju-si, Jeju-do, Korea	NIBRVP0000354345
9			Guseong-ri, Jugwang-myeon, Goseong-gun, Gangwon-do, Korea	NIBRVP0000470134
10			Seongu-ri, Onjeong-myeon, Uljin-gun, Gyeongsangbuk-do, Korea	NIBRVP0000538333
11			Jungjang-ri, Anmyeon-eup, Taean-gun, Chungcheongnam-do, Korea	NIBRVP0000538652
12			Cheongsu-ri, Hangeong-myeon, Jeju-si, Jeju-do, Korea	NIBRVP0000538652
13			Husapo-ri, Bubuk-myeon, Miryang-si, Gyeongsangnam-do, Korea	NIBRVP0000538646

468

469 **Table 4.**

No.	Markets	Material component	Origin
1	A	Dried or powdered seeds and pericarps: sancho products (50–300 g)	China
2	B		China
3	C		China
4	D		Korea
5	E		China
6	F		China
7	G		China
8	H		China
9	I		Korea
10	J		Korea
11	K	Dried or powdered seeds and pericarps: chopi products (20–300 g)	Korea
12	L		Korea
13	M		Korea
14	N		Korea
15	O		Korea
16	P		Korea

470