

RESEARCH PAPER

DNA barcode library and its efficacy for identifying food-associated insect pests in Korea

Seo Young CHO¹, Kyong In SUH² and Yeon Jae BAE^{1,2,3}

¹ Department of Life Sciences and Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, Korea

² Korean Entomological Institute, Korea University, Seoul, Korea

³ Division of Life Sciences, College of Life Sciences and Biotechnology, Korea University, Seoul, Korea

Correspondence

Yeon Jae Bae, Division of Life Sciences,
College of Life Sciences and Biotechnology,
Korea University, 5-ga, Anam-dong,
Seongbuk-gu, Seoul 136-701, Korea.
Email: yjbae@korea.ac.kr

Received 16 July 2013;
accepted 28 July 2013.

doi: 10.1111/1748-5967.12034

Abstract

Food-associated insect pests are of great economic and hygienic importance. However, their identification requires expert knowledge and excessive time. Such pests are discovered in food as body parts or immature stages, which further complicates the identification process. In this study, we constructed a DNA barcode dataset of insect pests that can be detected in food. We also tested the efficacy of these DNA barcode sequences for identifying food-associated insect pests. A 658 bp fragment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene was analyzed from 55 species of food-associated insect pests in Korea. The results indicated that this portion of the COI gene effectively discriminated >90% of the food-associated insect pests. Mean genetic divergences among individuals belonging to one species/between species belonging to one genus of the five orders, Blattaria, Coleoptera, Hymenoptera, Lepidoptera and Diptera, were 0.59%/13.18%, 0.84%/20.10%, 0.02%/22.61%, 0.24%/3.48% and 0.17%/15.90%, respectively. In conclusion, we established the first DNA barcode dataset and confirmed its efficiency for identifying food-associated insect pests in Korea.

Key words: Cytochrome *c* oxidase subunit I gene, DNA barcode, food-associated insect pests, Korea.

Introduction

The number of claimed cases of food-associated insect pests has increased rapidly in the last few years (Kim 2011). The major pathway of pest infestation arises from invasion of insects into stored products. Therefore, pest prevention is a major concern during harvesting, storage, transportation, processing, packaging and distribution of food. In addition, the occurrence of insect particles in processed food is an important quality-control problem in the industry (Gentry *et al.* 2001).

Most insects that are found in stored products are invaders (Campbell *et al.* 2004). They invade products through existing openings rather than destroying packaging film (Highland 1991; Adler 2008). For example, typical coleopteran penetrators are *Stegobium paniceum* (Linnaeus) and *Lasioderma serricorne* (Fabricius) from the family

Anobiidae, *Prostephanus truncates* (Horn) and *Rhyzopertha dominica* (Fabricius) from the family Bostrychidae and *Oryzaephilus surinamensis* (Linnaeus) from the family Silvanidae (Highland 1991; Choi *et al.* 1996; Campbell *et al.* 2002, 2004). Additionally, the Indian meal moth *Plodia interpunctella* (Hubner) and the almond moth *Cadra cautella* (Walker) are penetrators in their search for a pupation location (Highland 1991; Choi *et al.* 1996; Campbell *et al.* 2004). Therefore, identifying pest species and understanding their ecological characteristics are required to provide invasion control methods.

Insect identification relies on traditional taxonomy, which is primarily based on external morphology. However, taxonomic keys are often prepared for only certain life stages or genders; phenotypic variations in taxonomically important traits may also cause significant difficulties in species identification (Ball & Armstrong 2006). Furthermore,

food-associated insect pests are usually found as body parts or immature stages, making it virtually impossible to identify the species based on morphological characteristics. Therefore, DNA barcode approaches have been used for associating different developmental stages and for identifying partially preserved specimens unsuitable for morphological study (Wheeler 2004; Vences *et al.* 2005; Will *et al.* 2005). The mitochondrial gene cytochrome *c* oxidase unit I (COI) is commonly used as a primary character in most animal groups (Hebert *et al.* 2003a, 2003b). Although DNA barcode data have been generated for a relatively few well-known insect groups such as butterflies, mayflies, tussock moths, fruit flies and cereal aphids (Hebert *et al.* 2004; Armstrong & Ball 2005; Ball *et al.* 2005; Janzen *et al.* 2005; Shufran & Puterka 2011; Virgilio *et al.* 2012), DNA barcode reference libraries are still largely incomplete and, hence, cannot yet be used to reliably identify most other insect groups (Virgilio *et al.* 2012).

Cho *et al.* (2011) established a species list for food-associated insect pests in Korea. The purposes of this study were to construct a DNA barcode dataset and to test the efficacy of the DNA barcode for identifying food-associated insect pests in Korea.

Materials and methods

Insect specimen collection

Eighty-one insect specimens belonging to 24 species, 19 genera, 17 families and five orders were examined for this study (Table 1; the species are referred from Cho *et al.* 2011). These were larvae, pupae, or adult specimens of insect pests that were collected from food products, food depositories and food production facilities and factories located in diverse localities throughout South Korea in 2011 and 2012. The species of each specimen was identified by taxonomic authorities of each taxon using external morphology (see Acknowledgments). Specimens for DNA extraction were preserved in 95% ethanol. Voucher specimens were deposited in the Entomological Museum of Korea University, Seoul.

DNA extraction and polymerase chain reaction (PCR) analyses

Adult, larval and pupal specimens of the insect pests were used for DNA extraction. The larval digestive tract was removed to reduce possible contamination and the thorax muscle was cut for DNA extraction, whereas the head and abdomen were retained as vouchers. The thorax muscle or legs of adults was used for DNA extraction. The whole body was occasionally used when only old or relatively small specimens were available. All instruments used for dissec-

tion were sterilized with 95% ethanol 5–10 times between specimen dissections to prevent transfer of DNA from one sample to another.

Total DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen Genomics Inc., Dusseldorf, Germany), according to the manufacturer's protocol. The extracted DNA was dissolved in 80 μ L ddH₂O and stored at -20° C. The mainly universal primer pair LCO1490 (forward) 5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3' and HCO2198 (reverse) 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3' was used to amplify an ~658 bp region containing part of the mitochondrial COI gene (Folmer *et al.* 1994). PCR amplification was conducted under the following conditions: initial denaturation at 94° C for 3 min; 40 cycles at 94° C (30 sec), $45-48^{\circ}$ C (1 min) and 72° C (30 sec); and a final elongation at 72° C for 10 min. PCR products were stored at 4° C. We used a 50 μ L PCR reaction that contained ultrapure water, 0.5 U Taq DNA polymerase with $1 \times$ Taq, buffer containing 20 mM Tris-HCl (pH 8.2) and 50 mM KCl, 4 mM MgCl₂, 200 μ M total dNTP and 0.5 μ M of each primer. All PCR products were tested by electrophoresis on 1.5% agarose gels using GelRed™ (Biotium Inc, Hayward, CA, USA). PCR products were sequenced on an ABI 3730 Automated Sequencer at Macrogen Co. (Seoul, Korea). Sequences were deposited at the National Center for Biotechnology Information GenBank (<http://www.ncbi.nlm.nih.gov/>).

DNA sequencing analyses

DNA sequences acquired from the specimens were confirmed and edited manually using BioEdit Sequence Alignment Editor ver. 7.1.3 (Hall 1999) and Geneious ver.5.5.7 (Biomatters Ltd, Auckland, NZ). The COI sequences were aligned using Clustal X (Larkin *et al.* 2007) and trimmed to a final length of <658 bp. They were submitted to GenBank (<http://www.ncbi.nlm.nih.gov/>) using the Sequin program.

Genetic distance analyses

Pairwise sequence divergences of 311 DNA barcode sequences, combining 81 sequences directly acquired from this study and 230 GenBank sequences from 55 species of food-associated insect pests in Korea (GenBank data are presented in Table 1), were calculated at three taxonomic levels of species, genus and family using the Kimura two parameter (K2P) model (Kimura 1980). Neighbor-joining (NJ) trees based on the calculated generic differences within and between genera and species were constructed using MEGA 5.05 (Tamura *et al.* 2011), as presented in Figure 1.

Table 1 Insect specimens used to sequence the DNA barcoding region of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene and GenBank accession numbers. We obtained sequences with >600 bp and <1% ambiguity from 311 specimens of 55 food-associated insect pest species. Target food, claim information, and reference to the pest insect species are presented in Cho *et al.* (2011). GenBank accession numbers with asterisk (*) are specimens with sequences acquired from this study and deposited in GenBank

Taxa				
Order	Family	Species	GenBank accession number (COI)	Specimen no.
Blattaria	Blattellidae	<i>Blattella germanica</i> Linnaeus	*JQ350728, *KC407709, AY176057, EU253828, HM996892, NC012901, S72627, GBMH6255, GBMH6993	1–9
Blattaria	Blattidae	<i>Periplaneta americana</i> Linnaeus	*JQ350707, AM114927, AY165646, GU947663, NC016956, GBMH4121, SBGB018	10–16
Blattaria	Blattidae	<i>Periplaneta fuliginosa</i> Serville	*JQ350729, AB126004, NC006076, GBMH0026, GBMH1942	17–21
Blattaria	Blattidae	<i>Periplaneta japonica</i> Karny	*JQ350708, *KC407710, *KC407711, AM114929, GBMH4119	22–26
Coleoptera	Dermestidae	<i>Anthrenus verbascie</i> (Linnaeus)	*KC407712, *KC407713, *KC407714	27–29
Coleoptera	Dermestidae	<i>Dermestes tessellatocollis</i> Motschulsky	*KC407715, *KC407716	30–31
Coleoptera	Bostrichidae	<i>Rhizopertha dominica</i> (Fabricius)	*JQ989165, *KC407717, *KC407718	32–34
Coleoptera	Anobiidae	<i>Stegobium paniceum</i> (Linnaeus)	*KC407719-KC407724	35–40
Coleoptera	Silvanidae	<i>Oryzaephilus surinamensis</i> (Linnaeus)	*JQ350709, *KC407725, *KC407726, *KC407727, FM877921, AAF0496	41–46
Coleoptera	Tenebrionidae	<i>Tenebrio molitor</i> Linnaeus	*KC407728-KC407741	47–60
Coleoptera	Tenebrionidae	<i>Tribolium castaneum</i> (Herbst)	GBCL2711, GBCL4148	61–62
Coleoptera	Tenebrionidae	<i>Tribolium confusum</i> Jacquelin Du Val.	*JQ350711, *KC407742, *KC407743, FJ743725	63–66
Coleoptera	Rhynchophoridae	<i>Sitophilus oryzae</i> Linnaeus	*JQ350733, GU196317, GU196318, AY131099	67–70
Hymenoptera	Formicidae	<i>Monomorium pharaonis</i> Linnaeus	*JQ350713, *KC407744, *KC407745	71–73
Hymenoptera	Formicidae	<i>Monomorium cryptobium</i> (Santschi)	GU709857-GU709859, GU709862, GU709864, GU709865, GU709867	74–80
Hymenoptera	Formicidae	<i>Monomorium destructor</i> (Jerdon)	GU709851, GU709852, GU709854-GU709856	81–85
Hymenoptera	Formicidae	<i>Monomorium floricola</i> (Jerdon)	GU709843, GU709845-GU709850	86–92
Hymenoptera	Formicidae	<i>Monomorium latinode</i> (Mayr)	GU709833, GU709835, GU709837-GU709842, GU709844	93–101
Lepidoptera	Noctuidae	<i>Helicoverpa assulta</i> (Guenee)	*JQ350724, GQ892841, GQ892856-GQ892861, GQ892863, JX156330, EU768937	102–112
Lepidoptera	Noctuidae	<i>Helicoverpa armigera</i> (Hübner)	GQ995237, GQ995238, GQ995241, GQ892842, GQ892849, GU686955, GU654969, HM854928, HM854930, EU768936, JX156326, JF415782	113–124
Lepidoptera	Noctuidae	<i>Helicoverpa punctigera</i> (Wallengren)	JQ240198, EU768941	125–126
Lepidoptera	Noctuidae	<i>Helicoverpa zea</i> (Boddie)	GU087831, GU090470, JQ577648, JQ578528, EU768942, JX156328, JX156329, JF854710	127–134
Lepidoptera	Noctuidae	<i>Helicoverpa hawaiiensis</i> (Quaintance & Brues)	EU768939	135
Lepidoptera	Noctuidae	<i>Helicoverpa pallida</i> Hardwick	EU768940	136
Lepidoptera	Noctuidae	<i>Helicoverpa gelotopoeon</i> (Dyar)	EU768938	137
Lepidoptera	Pyrilidae	<i>Plodia interpunctella</i> (Hubner)	*JQ350723, *KC407746-KC407753, MECB221-MECB225	138–151

Table 1 *Continued*

Taxa				
Order	Family	Species	GenBank accession number (COI)	Specimen no.
Diptera	Culicidae	<i>Culex pipiens pallens</i> Coquillett	*JQ350727, *KC407754, FN395181-FN395183, FN395185, FN395187, FN395206, HQ724614-HQ724616	152–162
Diptera	Culicidae	<i>Culex pipiens molestus</i> Forska	*KC407755-KC407759, FN395171-FN395173	163–170
Diptera	Culicidae	<i>Culex torrentium</i> Martini	JQ253808, JQ253810, JQ253818, JQ253820, FN395191, FN395195, FN395196, FN395198, FN395200	171–179
Diptera	Chironomidae	<i>Chironomus dorsalis</i> (Meigen)	*KC407760-KC407764, JN887046-JN887048	180–187
Diptera	Chironomidae	<i>Chironomus flaviplumus</i> (Tokunaga)	JF412075-JF412077	188–190
Diptera	Chironomidae	<i>Chironomus javanus</i> (Kieffer)	JF412082, JF412083, JF412085	191–193
Diptera	Chironomidae	<i>Chironomus kiiensis</i> Tokunaga	*KC407765, JF412086-JF412089, COTW008	194–199
Diptera	Chironomidae	<i>Chironomus plumosus</i> (Linnaeus)	*KC407766-KC407771, JF412098-JF412100, JN887054	200–209
Diptera	Chironomidae	<i>Chironomus nipponensis</i> Tokunaga	JN887051-JN887053, JF412090-JF412092, JF412097	210–216
Diptera	Chironomidae	<i>Chironomus riparius</i> Meigen	HM137887-HM137895	217–225
Diptera	Drosophilidae	<i>Drosophilla melanogaster</i> Meigen	*JQ350715, *KC407772, HQ979010, HQ979116, GQ229519, JQ686693, JQ686694, JQ686697, FJ190106, FJ190109, FJ190110	226–236
Diptera	Drosophilidae	<i>Drosophilla Simulans</i> Sturtevant	AY518671, AY518674, AF200844	237–239
Diptera	Drosophilidae	<i>Drosophila mauritiana</i> (Tsacas and David)	M57912, AF200831	240–241
Diptera	Drosophilidae	<i>Drosophilla sechellia</i> (Tsacas & Baechli)	M57908, AF200832	242–243
Diptera	Calliphoridae	<i>Lucilia sericata</i> (Linnaeus)	*JQ350713, AJ417712, AJ417714, AJ417715, AJ417717, HQ978732, HQ979099, HQ979101, FJ614824-FJ614826, EU880208, EU880210, EU880212	244–257
Diptera	Calliphoridae	<i>Lucilia caesar</i> (Linnaeus)	EU880194-EU880196	258–261
Diptera	Calliphoridae	<i>Lucilia cuprina</i> (Wiedemann)	DQ453495, DQ453496, JX187387-JX187390, AJ417704, AJ417705	262–269
Diptera	Calliphoridae	<i>Lucilia illustris</i> (Meigen)	EU880198, EU880200, EU880201, EU880205, FJ614827, FJ614828, L14945	270–276
Diptera	Calliphoridae	<i>Lucilia porphyrina</i> Walker	FJ614829, FJ614830, AY097336	277–279
Diptera	Calliphoridae	<i>Lucilia thatuna</i> Shannon	DQ453489	280
Diptera	Calliphoridae	<i>Lucilia silvarum</i> Meigen	FR719175	281
Diptera	Muscidae	<i>Musca domestica</i> (Meigen)	*JQ350716, AF259518, GQ465784, AB479528, AB479529, AY526196, EU154477, HM389238, HM389239	282–290
Diptera	Muscidae	<i>Musca bezzii</i> Patton & Cragg	AB479532, AB479533	291–292
Diptera	Muscidae	<i>Musca crassirostris</i> Stein in Becker	AB479530, AB479531	293–294
Diptera	Muscidae	<i>Musca confisicata</i> Speiser	EU627698	295
Diptera	Phoridae	<i>Megaselia scalaris</i> (Loew)	*KC407773, *KC407774, JN896297, JN896298, GU075400	296–300
Diptera	Phoridae	<i>Megaselia longicostalis</i> Wood	JN896281-JN896283	301–303
Diptera	Phoridae	<i>Megaselia rufipes</i> (Meigen)	GU075403-GU075406, JN896279, JN896280	304–309
Diptera	Phoridae	<i>Megaselia subtumida</i> (Wood)	JN896284, JN896285	310–311

Results

Constructing DNA barcode dataset of food-associated insect pests in Korea

A DNA barcode dataset was constructed on the basis of 81 sequences from 24 species of food-associated insect pests in Korea. These species belong to 19 genera, 17 families and 5 orders of Blattaria, Coleoptera, Hymenoptera, Lepidoptera and Diptera. The sequences are deposited in GenBank (accession numbers are listed in Table 1 with an asterisk). Also, we made a dataset of a total of 311 sequences combining 230 sequences of 55 species from GenBank (listed in Table 1). All the species belong to the species listed in the food-associated insect pests in Korea (Cho *et al.* 2011).

The results are summarized as pairwise sequence divergences according to diverse taxonomic levels such as species, genus and family (Table 2). Species with more than one sample formed a tight cluster in the 5 orders (except for Hymenoptera in comparison between families of genetic divergence). In comparison with Coleoptera and Diptera, mean divergence among individuals belonging to one species/between species belonging to one genus were 0.84%/20.10% and 0.20%/15.90%, respectively. In the 5 orders, Blattaria and Coleoptera were clearly distinguished by genetic divergences among individuals belonging to one species, between species belonging to one genus and between genera belonging to one family. The mean frequency of divergence among individuals belonging to one species was very low (mean 0.23%, range 0.00–1.49%). The sequence distances between different families always exceed 19% (mean 21.08%, range 19.10–22.54%) (Fig. 2).

In addition, COI sequences of *Rhizopertha dominica* (Fabricius), *Anthrenus verbascie* (Linnaeus), *Dermestes tessellatocollis* Motschulsky and *Stegobium paniceum* (Linnaeus) were submitted to GenBank for the first time.

Sequence divergence patterns

As indicated in the NJ tree (Fig. 1), members within a family formed a coherent cluster. In the order Blattaria, *Periplanta americana* showed higher variation in genetic divergences (mean, 1.3%; range, 0.0–2.5%; standard error [SE], 0.4%) than that of other species in the genus *Periplanta* (e.g. *Periplaneta fuliginosa* and *Periplaneta japonica*). The tree including Hymenoptera was clearly separated because no genetic variations among individuals belonging to one species were detected in the genus *Monomorium*.

Plodia interpunctella, in the family Pyralidae in Lepidoptera, is a widespread insect pest that is commonly found in food and grains and relatively low COI gene sequence divergences were shown among the 14 specimen pairs with a genetic pattern that showed a mixture of Korean and Canadian specimens (mean, 0.4%; range, 0.0–0.9%; SE, 0.2%). All species belonging to the genus *Helicoverpa* formed a tight cluster, despite relatively low genetic divergence.

Although most congeneric species tended to be grouped together and were separated from each other by the COI sequences, some dipteran genera such as *Lucilia* were not easily distinguished because of low genetic divergence, as between *Lucilia sericata* and *Lucilia cuprina* (mean, 0.32%; range, 0.30–0.34%; SE, 0.29%) and between *Lucilia illustris* and *Lucilia caesar* (mean, 0.47%; range, 0.34–0.58%; SE, 0.39%). In the mosquito subspecies, virtually no genetic divergences were found between *Culex pipiens pallens* and *Culex pipiens molestus*: the mean intrasubspecies K2P divergences were 0.17% (range, 0.00–0.46%; SE, 0.13%) and 0.00% (range, 0.00%; SE, 0.00%), respectively, and the mean intersubspecies divergence was 0.17% (range, 0.00–0.46%; SE, 0.13%).

As evidenced herein, most of the examined species showed relatively low genetic variations (<1%) within species from diverse geographic areas (e.g. *Blatta* and *Periplanta* in Blattaria, *Tribolium* in Coleoptera,

Table 2 Genetic divergence for 311 mitochondrial cytochrome *c* oxidase subunit I (COI) sequences from 55 species of food-associated insect pests according to three taxonomic levels in five orders: Blattaria, Coleoptera, Hymenoptera, Lepidoptera, and Diptera

Order	Intraspecies				Interspecies				Interfamilies			
	Mean (%)	Range	SE (%)	No. of specimens	Mean (%)	Range	SE (%)	No. of species	Mean (%)	Range	SE (%)	No. of families
Blattaria	0.59	0.00–2.49	0.20	26	13.18	10.49–17.00	1.60	17	20.04	18.42–22.16	2.07	2
Coleoptera	0.84	0.00–2.54	0.28	44	20.10	19.89–20.31	1.90	6	25.30	16.98–29.93	2.24	6
Hymenoptera	0.02	0.00–0.00	0.00	31	22.61	18.81–26.28	2.10	31				
Lepidoptera	0.24	0.00–0.92	0.10	50	3.48	1.70–15.18	0.90	36	13.91	13.15–15.18	1.54	2
Diptera	0.20	0.00–2.38	0.16	160	15.90	11.50–19.70	1.70	46	19.40	10.79–27.26	1.88	4

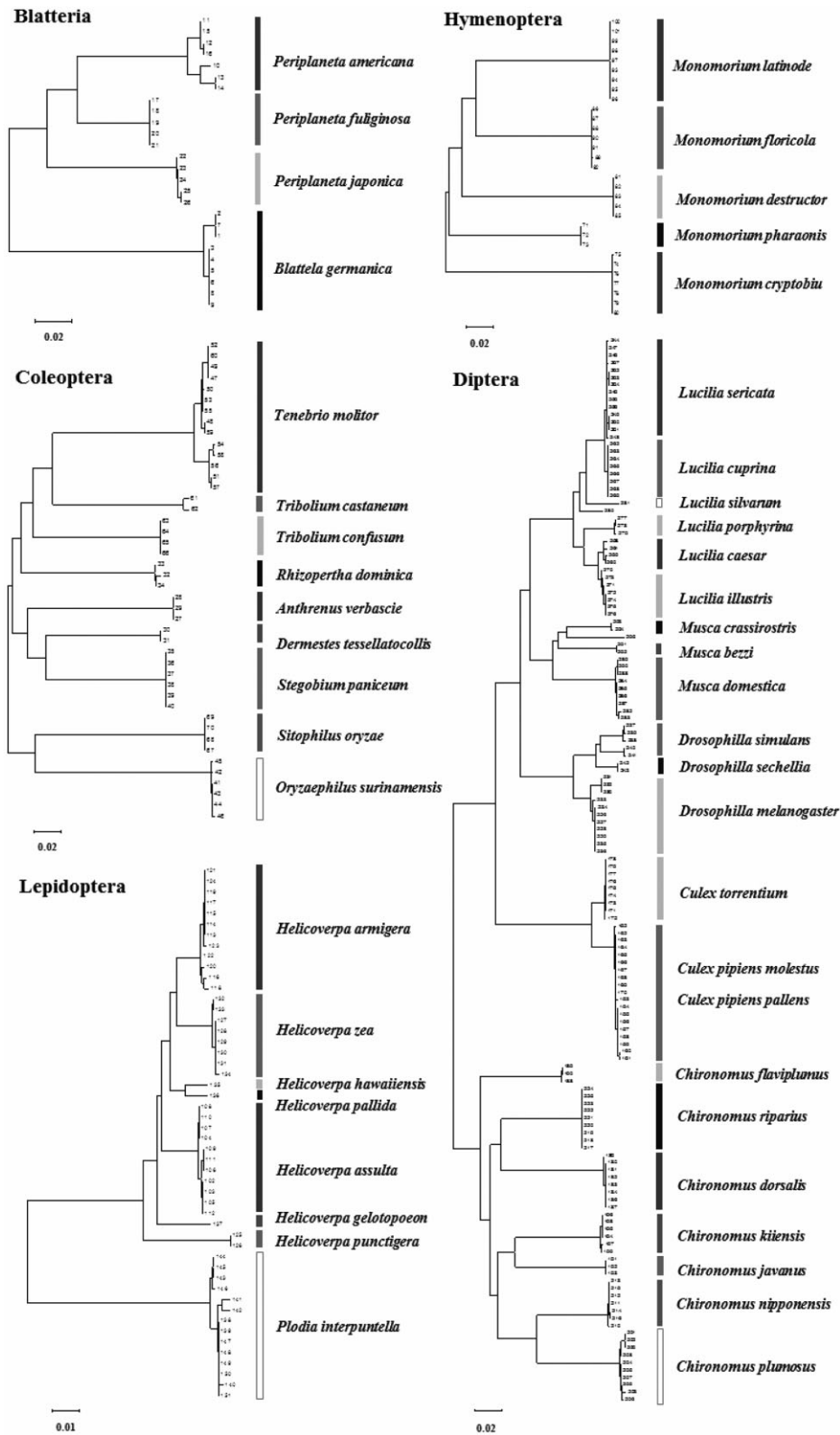


Figure 1 A neighbor-joining (NJ) tree of Kimura two parameter distance based on mitochondrial DNA mitochondrial cytochrome c oxidase subunit I (COI) reference libraries for 311 specimens from 55 species of food-associated insect pests belonging to five orders: Blattaria, Coleoptera, Hymenoptera, Lepidoptera and Diptera.

Figure 2 Genetic divergence (Kimura two parameter distance) between mitochondrial cytochrome c oxidase subunit I (COI) sequences for diverse taxonomic levels of food-associated insect pests. Frequency of pairwise divergence among individuals belonging to one species, between species belonging to one genus, and between genera belonging to one family. ■, Among individuals belong to one species; ▒, Between species belonging to one genus; □, Between genera belonging to one family.

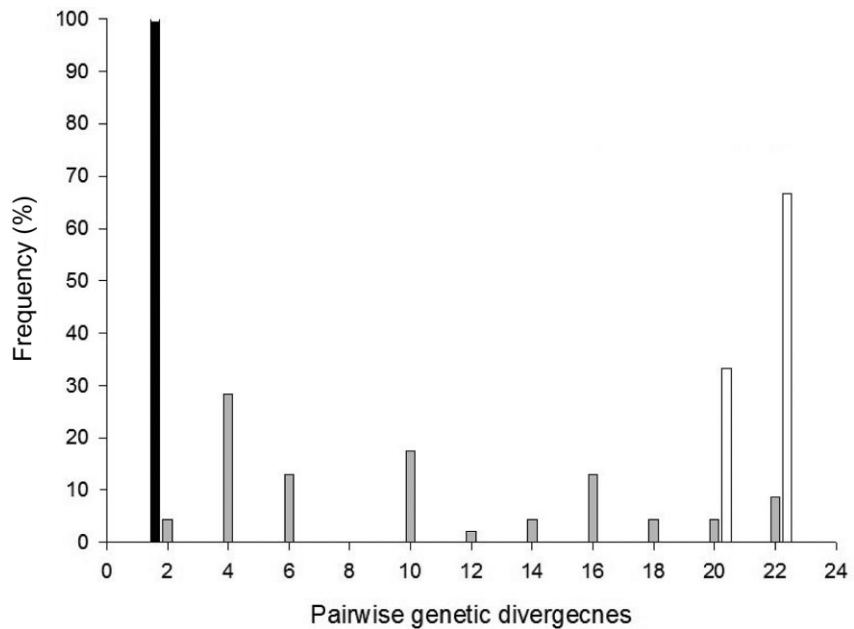
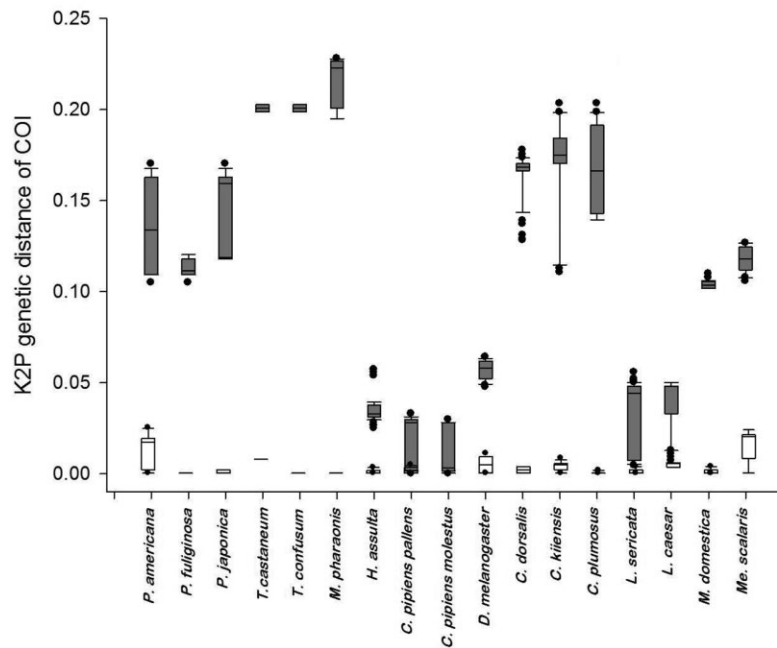


Figure 3 Box plots depicting differences between the minimum and maximum variations using the Kimura two parameter distance model. Gray box plots indicate differences between species belonging to one genus, and white box plots indicate differences among individuals belonging to one species. Black circles indicate 5th to 95th percentile for outliers. Bars indicate genetic range of the species boundary. ▒, Between species belonging to one genus; □, Among individuals belonging to one species.



Helicoverpa and *Plodia interpunctella* in Lepidoptera and *Lucilia*, *Musca* and *Drosophila* in Diptera) because these pest insects are widespread throughout the world and easily transported by human activities such as food trade.

Efficacy of the DNA barcode

We tested the efficacy of the DNA barcode to identify food-associated insect pests using 17 species belonging to 10

genera (Fig. 3). The barcode gaps clearly distinguished taxa above the species level in the majority of the tested species. As shown in the genus *Helicoverpa*, a relatively small barcode gap was observed within species (mean, 0.12%; range, 0.00–0.21%; SE, 0.12%), whereas a considerable barcode gap was found between species (mean, 3.5%; range, 2.49–5.72%; SE, 0.74%). A large barcode gap was observed between congeners in *Tribolium castaneum* and *Tribolium confusum* despite their morphological similarity.

Periplaneta americana and *Megaselia scalaris* showed relatively higher intraspecific genetic variations, whereas *Monomorium pharaonis* showed the largest interspecific genetic variations (Fig. 3). The results showed that the DNA barcode for food-associated insect pests generally enabled effective species discrimination.

Discussion

The main purpose of our study was to construct a DNA barcode dataset to accurately identify insects associated with food, as well as insects that can be detected in food. DNA barcoding cannot be used for identifying species if no DNA barcode data for related species or taxa are available in a reference dataset. Constructing a DNA reference dataset is more difficult if the target organism groups are taxonomically heterogeneous, as shown for food-associated insect pests. These limits are evident in insects due to a lack of reference DNA barcodes for approximately 90% of described species (Ratnasingham & Hebert 2007). Specimens not represented in the reference library will be erroneously assigned to the most similar heterospecific DNA barcode in the library (Virgilio *et al.* 2012). Although the major insect pests in food are widespread worldwide, only a few studies have been conducted on the DNA barcodes for these species. Therefore, this study is the first to attempt construction of a DNA reference dataset using the mitochondrial COI gene from food-associated insect species. This dataset can be effectively used to identify food-associated insect pests that are currently important in commercial food markets.

Based on our DNA barcode data, genetic divergences between species belonging to one dipteran genus were relatively small. In the case of the family Muscidae, which is a major food pest (Highland 1984, 1991; Campbell *et al.* 2004), mean genetic divergences among individuals belonging to one species and between species belonging to one genus were 0.1% and 9.9%, respectively. Another study on Muscidae showed a similar pattern (<0.6% among individuals belonging to one species, 9.7–14.4% between species belonging to one genus) (Yu *et al.* 2007). Another study also showed a relatively low success rate (<70%) for identifying dipteran species when using a tree-based barcode. Misidentification is due to the wide overlap between intraspecific and interspecific genetic variability. Even when two COI sequences are identical, there is a 6% chance that they belong to different dipteran species (Meier *et al.* 2006, 2008). Constructing a comprehensive DNA barcode library will improve the probability of accurately identifying the Diptera species.

Our results indicate that COI-based identification was effective for identifying food insect pests (Fig. 3). High genetic variation among individuals belonging to one species and between species belonging to one genus will

improve the efficacy of taxonomic discrimination. However, our data on the genus *Lucilia* (Calliphoridae in Diptera) showed limited success (maximum divergence value among individuals belonging to one species was 0.4% and the minimum value between species belong to one genus was 0.5%) (Figs 1,3). Whitworth *et al.* (2007) demonstrated that this very low success of the barcoding approach is due to non-monophyly in the taxa including the genus *Lucilia* in Diptera. A similar pattern of lower intraspecific and interspecific divergences (<0.2% among individuals belonging to one species and a range of 0.7–2.4% between species belonging to one genus) reported in other studies has been fundamentally attributed to the common occurrence of paraphyly and polyphyly among closely related species (Funk & Omland 2003).

Despite the many phylogenetic studies, the phylogeny of *Culex* remains unknown, and its classification is problematic (Harbach 2011). Diverse taxonomic groups in *Culex* were based exclusively on morphological similarities that are interpreted by traditional taxonomic methods to represent natural groupings of the species, despite their very confusing external morphology except for genital characters. *Culex pipiens pallens* and *C. pipiens molestus* were not only morphologically similar but also showed high genetic similarity. As shown in our examinations, morphological re-examination and further taxonomic studies are needed for groups with very low genetic difference such as some taxa of Diptera, including *Lucilia caesar* and *Lucilia sericata*, and the subspecies of *Culex pipiens*.

Acknowledgments

We thank the Animal Resources Team researchers at the National Institute of Biological Resources (NIBR, Incheon) for providing voucher specimens and lab facilities. We thank Sang Eon Shin (Sung Shin University, Seoul; Coleoptera), Boo Hee Jung (Korea University, Seoul; Coleoptera), Jong Wook Lee (Yeungnam University, Gyeongsan; Coleoptera and Hymenoptera), and Hye Woo Byun (NIBR, Incheon; Diptera) for identifying voucher specimens. This study was supported by a grant (10162KFDA995) from the Korea Food & Drug Administration in 2012.

References

- Adler C (2008) Insect-proof packaging to avoid stored product insects. *Integrated Protection of Stored Products IOBC/wprs Bulletin* **40**: 363–369.
- Armstrong KF, Ball SL (2005) DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**: 1813–1823.
- Ball SL, Armstrong KF (2006) DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera:

- Lymantriidae). *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* **36**: 337–350.
- Ball SL, Hebert PDN, Burian SK, Webb JM (2005) Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *Journal of the North American Benthological Society* **24**: 508–524.
- Campbell JF, Arthur FH, Mullen MA (2004) Insect management in food processing facilities. *Advances in Food and Nutrition Research* **48**: 1043–4526.
- Campbell JF, Mullen MA, Dowdy AK (2002) Monitoring stored-product pests in food processing plants with pheromone trapping, contour mapping, and mark-recapture. *Journal of Economic Entomology* **95**: 1089–1101.
- Cho SY, Suh KI, Yoon TJ, Na JH, Bae YJ (2011) Checklist of insect pests of foods in Korea. *Entomological Research Bulletin* **27**: 19–30.
- Choi KM, Lee MH, Han MJ, Ahn SB, Hong KJ (1996) *Stored Product Insect Pests with Pictorial Key to Larvae*. National Institute of Agricultural Science and Technology, Suwon. pp 1–227.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**: 294–299.
- Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics* **34**: 397–423.
- Gentry JW, Harris KL, Luce JT (2001) History and importance of qualitative microentomological determinations. *Cereal Foods World* **46**: 162–163.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95–98.
- Harbach RE (2011) Classification within the cosmopolitan genus *Culex* (Diptera: Culicidae): the foundation for molecular systematics and phylogenetic research. *Acta Tropica* **120**: 1–14.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**: 313–321.
- Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 14812–14817.
- Hebert PDN, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**: S96–S99.
- Highland HA (1984) Insect infestation of packages. In: Baur FJ (ed.) *Insect Management for Food Storage and Processing*, pp 311–320. American Association of Cereal Chemists, St. Paul.
- Highland HA (1991) Protecting packages against insects. In: Gorham JR (ed.) *Ecology and Management of Food-industry Pests*, pp 345–350. FDA Technical Bulletin 4, Association of Official Analytical Chemists, Arlington, Virginia.
- Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E (2005) Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**: 1835–1845.
- Kim JM (2011) *Understanding and Prevention of Foreign Matter Contamination in Foods*. Korea Food Safety Research Institute, Seoul. pp 19–29.
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *Journal of Molecular Evolution* **16**: 111–120.
- Larkin MA, Blackshields G, Brown NP *et al.* (2007) Clustal W and clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.
- Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology* **55**: 715–728.
- Meier R, Zhang GY, Ali F (2008) The use of mean instead of smallest interspecific distances exaggerates the size of the “Barcoding Gap” and leads to misidentification. *Systematic Biology* **57**: 809–813.
- Shufran KA, Puterka GJ (2011) DNA barcoding to identify all life stages of holocyclic cereal aphids (Hemiptera: Aphididae) on wheat and other Poaceae. *Annals of the Entomological Society of America* **104**: 39–42.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731–2739.
- Vences M, Thomas M, Bonett RM, Vieites DR (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**: 1859–1868.
- Virgilio M, Jordaens K, Breman FC, Bacheljau T, DeMeyer M (2012) Identifying insects with incomplete DNA barcode libraries, African fruit flies (Diptera: Tephritidae) as a test case. *PLoS ONE* **7**: e31581.
- Wheeler QD (2004) Taxonomic triage and the poverty of phylogeny. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359**: 571–583.
- Whitworth TL, Dawson RD, Magalon H, Baudry E (2007) DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). *Proceedings of the Royal Society B-Biological Sciences* **274**: 1731–1739.
- Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* **54**: 844–851.
- Yu DH, Park SH, Zhang Y *et al.* (2007) Molecular identification of four Muscidae species collected in Korea using the DNA barcode region of mitochondrial cytochrome oxidase subunit I. *Korean Journal of Legal Medicine* **31**: 45–50.