MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2014) **14**, 50–59 doi: 10.1111/1755-0998.12165

Comparative studies on species identification of Noctuoidea moths in two nature reserve conservation zones (Beijing, China) using DNA barcodes and thin-film biosensor chips

F. YANG,* Z. Y. SHI,* S. L. BAI,* R. D. WARD† and A. B. ZHANG*
College of Life Sciences, Capital Normal University, Beijing 100048, China, †Wealth from Oceans Flagship, CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tasmania 7001, Australia

Abstract

Rapid and accurate identification of species is required for the biological control of pest Noctuoidea moths. DNA barcodes and thin-film biosensor chips are two molecular approaches that have gained wide attention. Here, we compare these two methods for the identification of a limited number of Noctuoidea moth species. Based on the commonly used mitochondrial gene cytochrome *c* oxidase I (the standard DNA barcode for animal species), 14 probes were designed and synthesized for 14 species shared by two national nature reserves in Beijing and Hebei, China. Probes ranged in length from 18 to 27 bp and were designed as mismatch probes to guarantee that there were at least three base differences between the probe and nontarget sequences. The results on the chip could be detected by the naked eye without needing special equipment. No cross-hybridizations were detected although we tested all probes on the 14 target and 24 nontarget Noctuoidea species. The neighbour-joining tree of the 38 species based on COI sequences gave 38 highly supported independent groups. Both DNA barcoding and thin-film biosensor chips, based on the COI gene, are able to accurately identify and discriminate the 14 targeted moth species in this study. Because of its speed, high accuracy and low cost, the thin-film biosensor chip is a very practical means of species identification. Now, a more comprehensive chip will be developed for the identification of additional Noctuoidea moths for pest control and ecological protection.

Keywords: DNA barcoding, identification, Noctuoidea, thin-film biosensor chip

Received 4 June 2013; revision received 21 August 2013; accepted 22 August 2013

Introduction

Noctuoidea are the largest superfamily within the Lepidoptera (Zahiri *et al.* 2010). A large number of Noctuoidea moths pose threats to agriculture (Mitchell *et al.* 2006). Some of them are devastating pests, for example *Spodoptera litura*, *Spodoptera exigua*, *Agrotis ipsilon*, *Athetis lepigone* and *Argyrogramma agnata*. Their main targets include the plant families Cruciferae, Solanaceae, Cucurbitaceae, Poaceae and Leguminosae (Luo *et al.* 2004; Baker & Tann 2013; Kandagal & Khetagoudar 2013; Mohamed 2013; Smigocki *et al.* 2013). Accurate and rapid identification of Noctuoidea moths is the solid foundation required for their biological control and for agriculture protection. However, most traditional identification focus on morphological observation of body parts, including the structures of wings, genitalia, larvae and eggs, requiring specialist taxonomic knowledge. The whole procedure is time-consuming and easily affected by subjective assessments, making some identification incorrect. The procedure is made more difficult by possible taxonomic inexactitudes. For example, *Spodoptera frugiperda*, a severe agriculture pest, has two different ecotypes whose feeding habits are different: one feeding on maize and the other on rice. The species has brought great distress to the agricultural economy of the western hemisphere, but while the two types cannot be distinguished morphologically, they can be distinguished by molecular characterizations of their repeated DNA sequences (Lu *et al.* 1994). The limitations of morphological identification and a reduced number of taxonomists have led to a 'taxonomy crisis', and a more integrative approach to taxonomy has been proposed for species delineation (Dayrat 2005).

With the rapid development of molecular biology in the 1990s, an increasing number of scientists applied PCR and DNA sequencing technologies to the identification and classification of species. However, initial

Correspondence: Ai-Bing Zhang, Fax: 8610-68901860;
E-mails: zhangab2008@gmail.com;
zhangab2008@mail.cnu.edu.cn

research was hampered by the lack of a unified molecular marker. In 2003, Hebert *et al.* (2003a) from Canada proposed the concept of 'DNA barcoding', which aimed to use the sequence of a single piece of mtDNA (part of the cytochrome *c* oxidase subunit I gene, COI) for the rapid, accurate and automated identification of most animal species. Moths and butterflies (Lepidoptera) have been a common target of DNA barcoding studies because of their large number of species and complicated relationships, and published results have repeatedly confirmed the validity of COI-based species identification (Hebert *et al.* 2003b, 2004; Janzen *et al.* 2005; Hajibabaei *et al.* 2006a,b; Hulcr *et al.* 2007; Moreira *et al.* 2012; Lastuvka *et al.* 2013). However, rather few studies based on COI have focused on the identification of Noctuoidea moths. Most have been concerned with detailed studies of the taxonomic status of one or two species (Schmidt 2009; Lafontaine & Poole 2010), although a few have been more wide-ranging (Schmidt & Sperling 2008; Hausmann *et al.* 2011; Joomun *et al.* 2012; Yang *et al.* 2012; Zahiri *et al.* 2013).

Most DNA sequencing approaches are limited by the requirement that the specimen sample contain a single unique DNA sequence (Tambong *et al.* 2006; Kochzius *et al.* 2008, 2010) and thus cannot be applied to discriminate mixed samples, such as moth eggs for instance. Next-generation sequencing based on high-throughput parallel sequencing can permit the discrimination of mixed DNA samples (Urich *et al.* 2008; Kochzius *et al.* 2010). However, the requirements for a well-constructed gene library, the difficult control of multienzyme systems, expensive instrumentation and high sequencing costs combine to make this technology not widely used (Shendure & Ji 2008; Kircher & Kelso 2010; Metzker 2010; Wei *et al.* 2012).

Microarray technology, another method of high-throughput detection, originated 20 years ago. Its initial application field was the construction of gene expression profiling and large-scale genotyping. It is based on the principle of the complementary pairing of nucleotides, so that a fixed probe is able to hybridize with the single-stranded form of fluorescently labelled target DNA. After a series of elution process, sequences of failed hybridization are washed off, and the successful hybridization duplex complexes can be observed by a fluorescent signal. Because of its simple operation, and high levels of repeatability and accuracy, microarray technology has been applied to species identification, for example in the detection of pathogenic microorganisms (Korimbocus *et al.* 2005; Garaizar *et al.* 2006; Zhang *et al.* 2010). Furthermore, numerous COI probes have been designed and used for species identification in the animal kingdom. Examples include six jellyfish species (Lee *et al.* 2011), the differentiation of Mediterranean fly from related species (Li *et al.* 2008), five filefishes (Kim *et al.* 2011), 17 birds (Chung *et al.* 2010) and genetically closely related voles and shrews (Pfunder *et al.* 2004).

Typically, these experiments require expensive signal detection instruments to 'read' information from the chip, together with sophisticated software programs to quantify and interpret the results (Blohm & Guiseppi-Elie 2001). To build an accurate but less expensive microarray system, Ostroff launched the first 'thin-film biosensor chip' in 1999 to detect target DNA from human serum and urine (Ostroff *et al.* 1999). When white light irradiates the silicon nitride surface of this optical biosensor chip, it initially reflects a golden colour; after enzymatic hybridization catalysis, the insoluble mass deposited on the chip increases the thickness of the thin-film surface, which alters the wavelength of reflected light. The surface colour where the mass stays on changes to blue can be perceived by the naked eye. Because of the optical characteristics of the thin-film biosensor chip surface, one of the greatest benefits of this approach is that experimental results can be visualized by the unaided human eye without the help of specific instrumentation. Shortly thereafter, this biosensor chip was used in the detection and identification of human genome SNPs (Zhong *et al.* 2003), food pathogens (Zhao *et al.* 2008; Bai *et al.* 2010a), specific nucleotide sequence variations in plants (Bai *et al.* 2007) and genetically modified crops (Bai *et al.* 2010b, 2012). Our study, based on the COI gene fragment, introduces thin-film biosensor chips for moth identification. We used this technology, together with a conventional DNA barcoding approach, to identify Noctuoidea moths collected from two national nature reserves in Beijing, China. The study aims to compare the two identification methods and to test the ability of the thin-film biosensor chip to accurately identify these moths in our study from a limited geographic region.

Materials and methods

COI sequences used in this study had been deposited in GenBank (Chi *et al.* 2012; Yang *et al.* 2012). We selected 38 species of Noctuoidea containing at least two COI sequences to obtain more information on intraspecific sequence variation. These species belong to three different families according to the latest catalogue of Noctuoidea (Zahiri *et al.* 2010). Probes were designed for 14 species shared by two nature reserve conservation zones near Beijing (China): BHS is the BaiHuaShan national nature reserve conservation zone, and WLS is the WuLingShan nature reserve conservation zone. Name and number of sequences, GenBank Accession nos., taxonomic status and collection sites are listed in Table 1.

Table 1 Species, sampling sites and corresponding GenBank Accession nos. used in this study

Species	Sequences of COI	GenBank Accession nos.*	Family†	Sampling sites
Anterastria atrata	2	JX392747, JX411865	Noctuidae	BHS and WLS
Athetis lineosa	2	JX392766, JX392767	Noctuidae	BHS
Athetis furvula	2	JX411794, JX411795	Noctuidae	WLS
Athetis pallidipennis	3	JX411796–JX411798	Noctuidae	WLS
Blasticorhinus unduligera	2	JX411840, JX411841	Noctuidae	WLS
Bryophila orthogramma	2	JX411858, JX411859	Noctuidae	WLS
Calyptra hokkaida	3	JX392801, JX411837, JX411838	Erebidae	BHS and WLS
Calyptra lata	4	JX411833–JX411836	Erebidae	WLS
Callopistria juventina	5	JX392730–JX392734	Noctuidae	BHS
Catocala columbina	2	JX392755, JX392756	Erebidae	BHS
Chrysorithrum amata	2	JX392768, JX392769	Noctuidae	BHS
Eucarta virgo	2	JX411881. JX411882	Noctuidae	WLS
Eutelia hamulatrix	3	JX411830–JX411832	Euteliidae	WLS
Flexivaleria mienshani	3	JX392789–JX392791	Noctuidae	BHS
Hadjina chinensis	21	JX411866–JX411876, JX392735–JX392744	Noctuidae	BHS and WLS
Hypena kengkalis	2	JX392785, JX392786	Erebidae	BHS
Hypena stygiana	3	JX392793, JX392794, JX411800	Erebidae	BHS and WLS
Hypena tristalis	3	JX411799, JX392780, JX392781	Erebidae	BHS and WLS
Hypocala subsatura	3	JX392745, JX392746, JX411814	Erebidae	BHS and WLS
Iambia transversa	5	JX411805, JX411890, JX411807–JX411809	Noctuidae	WLS
Lacanobia aliena	4	JX392774–JX392777	Noctuidae	BHS
Mocis ancilla	3	JX411842, JX411843, JX392761	Erebidae	BHS and WLS
Moma kolthoffi	2	JX411802, JX411804	Noctuidae	WLS
Mythimna monticola	12	JX411819–JX411829, JX392779	Noctuidae	BHS and WLS
Mythimna velutina	4	JX411815–JX411818	Noctuidae	WLS
Narangodes argyrostrigatus	5	JX411845–JX411849	Noctuidae	WLS
Niphonyx segregata	7	JX392748–JX392750, JX411861–JX411864	Noctuidae	BHS and WLS
Pangrapta disruptalis	2	JX392728–JX392729	Erebidae	BHS
Pangrapta obscurata	2	JX411888, JX411889	Erebidae	WLS
Paracolax tristalis	4	JX411884–JX411886, JX392764	Erebidae	BHS and WLS
Phyllophila obliterate	5	JX411877–JX411879, JX392796, JX392797	Noctuidae	BHS and WLS
Polypogon gryphalis	2	JX411883, JX392751	Erebidae	BHS and WLS
Pseudocosmia maculata	2	JX411855, JX411856	Noctuidae	WLS
Pyrrhia umbra	3	JX411850–JX411852	Noctuidae	WLS
Raphia peustera	3	JX411811, JX411812, JX392795	Noctuidae	BHS and WLS
Xanthomantis cornelia	2	JX392762, JX392763	Noctuidae	BHS
Xestia ditrapezium	2	JX392798, JX392799	Noctuidae	BHS
Zanclognatha lunalis	3	JX392772, JX392773, JX411887	Erebidae	BHS and WLS
Total:	141			

*Sequences in this article are shown in Appendix S1 (Chi *et al.* 2012; Yang *et al.* 2012).
†On the basis of the latest classification of Noctuoidea (Zahiri *et al.* 2010).

Sequence analysis

Multiple alignments of 141 COI sequences from 38 species of moths were performed by the program CLUSTAL W (Thompson *et al.* 1997) implemented in BIOEDIT (version 7.0.9.0) (Hall 1999). All sequences were manually edited and translated into amino acids with MEGA 4 (Tamura *et al.* 2007) to avoid sequencing errors or pseudogene sequences. Numbers of haplotypes, polymorphic nucleotides and nucleotide composition were calculated by DNASP 5.10.01 (Librado & Rozas 2009).

We used three sequences belonging to Trichoptera as the outgroup and constructed a rooted neighbour-joining (NJ) tree with MEGA 4 (Tamura *et al.* 2007) using the Kimura 2-parameter distance (Kimura 1980) and unweighted transversions and transitions. As we were only interested in species clustering and not in phylogenetic analysis, we included all bases although we are aware that third codon position transitions might be saturated (Martin *et al.* 2000). Confidence in the estimated relationship was determined by bootstrap analyses using MEGA 4 (Tamura *et al.* 2007). Figure 1 shows the NJ bootstrap

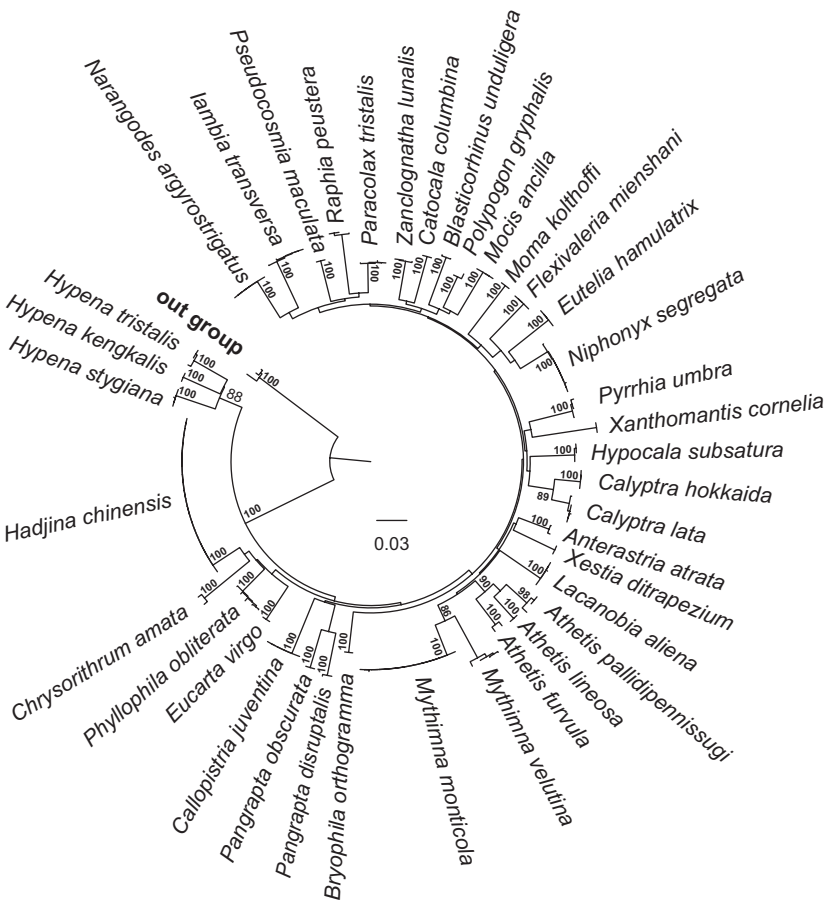

Fig. 1 Neighbours-joining tree based on COI DNA barcoding of 38 Noctuoidea moth species. The tree is assembled using K2P genetic distances (Kimura 1980) in a neighbour-joining method. Bootstrap values (1000 replicates) >50 are shown.

consensus tree from 1000 replicates, bootstrap values >50 being posted above branches.

Probe design and synthesis for the thin-film biosensor chip

Probe design depended on the previous multiple alignments of the 141 COI sequences. To design species-specific oligonucleotide probes for the 14 species shared by the two nature reserve conservation zones, we searched for diagnostic regions throughout the sequences. The length of probes should be <30 bp because shorter probes (about 25 bp) are better than longer probes (50–60 bp) for distinguishing highly homogenous sequences from different species in the same genus or family (Tomik & Hofmann 2001). The melting temperatures (T_m) of probes should fall in a narrow range to ensure the similarity of conditions among hybridizations (Liu *et al.* 2010). These are the most crucial factors that can influence the efficiency of oligonucleotide hybridization (Owczarzy *et al.* 1997; Hughes *et al.* 2001).

To reduce the possibility of cross-hybridization events and to enhance the discriminating power of probes, we selected perfectly matched candidate probes that exhibited no <3 base differences from nontarget sequences. When perfectly matched probes could not be found, artificial mismatch bases were introduced into probes to increase the probes' selective characteristics (Lee *et al.* 2004). To maximize efficiency, the mismatch base should avoid changing A or T into G because the mismatches [G,T] and [G,A] resulting from the mutation prove to be stable duplexes (Ikuta *et al.* 1986; Werntges *et al.* 1986; Lee *et al.* 2004). It has been previously demonstrated that a 50-mer probe which shares a 15-base or longer stretch with nontarget DNA could result in significant cross-hybridization (Kane *et al.* 2000) and cross-hybridization increases as the length of the shared stretch increases (He *et al.* 2005). Thus, this cut-off was set as the threshold in our study despite the fact that our probes were much shorter. Each probe had an aldehyde modification at the 5′ end for connection to the chip surface, followed by 12 deoxyadenosine residues which serve as a 'spacer'. The positive signal probe utilized 20 continuous deoxyadenosine residues, which would not hybridize with any COI sequence in this experiment. The 3′ end of the positive signal probe had a biotin modification. Details about the probes and primers are given in Table 2.

Table 2 Probes shared by moths from two nature reserve conservation zones

Probe name	Species	Sequence (5′-to 3′, 5′-aldehyde)	GC (%)	T_m (°C)*	Location
AA	*Anterastria atrata*	ALD-A(12) GTACCCTCCTCTT**A**CTTCT	29	50	284–302
CH	*Calyptra hokkaida*	ALD-A(12) TTCTCTACATTTA**C**CAGGAATC	23.5	49.7	344–365
HC	*Hadjina chinensis*	ALD-A(12) GGTTTT**T**CCCCCCCCTATCTTA**T**	32.3	53.3	281–302
HS	*Hypocala subsatura*	ALD-A(12) TGCATTCCTACTTT**C**ATTATCTT	20.0	48.8	473–495
HST	*Hypena stygiana*	ALD-A(12)TCTACTACTCCTTTCCCTACCT	29.4	52.1	479–500
HT	*Hypena tristalis*	ALD-A(12) T**C**ACTTCTTTCTCTTCCTGTA	24.2	49.4	483–503
MA	*Mocis ancilla*	ALD-A(12) ATTACATCTAGCA**T**GAATC	19.4	46.1	347–365
MM	*Mythimna monticola*	ALD-A(12) TAGATCT**C**TTGATCTAGCTA	21.9	47.8	320–339
NS	*Niphonyx segregata*	ALD-A(12) TTACTTCCACCTTCTATTA**G**TTTAT	18.9	49.4	210–234
PG	*Polypogon gryphalis*	ALD-A(12) TCATGGCGGAA**C**TTCTGTCGATT	31.4	53.5	311–333
PO	*Phyllophila obliterate*	ALD-A(12) TTACTA**G**CCCCATCATTAACATTAT	21.6	50.5	210–234
PT	*Paracolax tristalis*	ALD-A(12) TTA**G**CCCCATCTTTAACCCTTCTT**C**TT	28.2	54.2	213–239
RP	*Raphia peustera*	ALD-A(12) CCCC**A**CTCTTACCCTTCTT	32.3	51.3	218–236
ZL	*Zanclognatha lunalis*	ALD-A(12) A**T**TCTACCCCCCACTATCA**A**CA	29.4	52.1	281–302
PS	Positive signal	ALD-A(20)-BIO	–	–	–

*Salt concentration is 0.05 M; Bases of Bold black are the mismatch bases.

PCR amplification

The barcode COI gene fragment of mitochondrial DNA was amplified for the hybridization procedures. Fifty microlitre reaction mixtures contained DNA 3 μL, ddH2O 18 μL,2× mastermix 27 μL, 10 μM former primer LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) 1 μL, 10 μM reverse primer HCO2198 (5′-TAAACTTC-AGGGTGACCAAAAAATCA-3′) 1 μL. The reverse primers for PCR were synthesized with a biotin group at 5′ ends for subsequent detection. Amplification procedures in an ABI 2720 Thermal Cycler (Applied Biosystems) were as follows: heating at 94 °C for 5 min for predenaturation, 45 cycles of 30 s at 94 °C, 50 °C for 30 s, and 72 °C for 30 s, followed by the last step of 5 min at 72 °C. PCR amplicons were analysed using 1% agarose gels.

Preparation of the optical thin-film biosensor chips

The biosensor chips used here were commercially available from Inverness Medical-Biostar (Louisville, CO, USA). To facilitate covalent attachment of aldehyde-labelled probe, the wafers were coated with T-structure aminoalkylpolydimethylsiloxane (TSPS) and poly (Phe-Lys) (PPL). Biosensors were prepared following the procedure described by Zhong *et al.* (2003).

Standard assay procedures

Aldehyde-labelled probes, at a concentration of 2 μM, were spotted by manual pipetting or robotic pipetting (Biodot AD3200), 300 or 40 nL per spot, respectively, onto biosensor chips. PCR amplicon targets 650 bp in length, at concentrations of 0.1 μM per 100 μL reaction, were denatured and hybridized on the chips for 10 min at 50 °C in the hybridization buffer [5× saline–sodium citrate (SSC) and 5 mg/mL acid-treated casein (ATC)]. After being washed three times in 0.1× SSC, chips were incubated with an antibiotin immunoglobulin G (IgG)–horseradish peroxidase (HRP) conjugate (Jackson ImmunoResearch; 1:1000 dilution from a 1 mg/mL stock in a buffer containing 5× SSC, 5 mg/mL ATC, and 10% glycerol) for 5 min in hybridization buffer. Chips were washed three times with 0.1× SSC, and then 100 μL of tetramethylbenzidine (TMB) formulation from BioFx Laboratories (Owings Mills, MD, USA) was added to the chips and incubated for 5 min at room temperature. Chips were then rinsed in ddH2O, air-dried and scored by eye. Results were imaged using a simple digital camera (Panasonic, DMC-ZS1).

Sensitivity and specificity detection of biosensor chips

To define sensitivity and specificity of biosensor chips, we randomly chose two probes, RP and PG, and spotted them on the chip manually. The volume of each spot was 300 nL (see Fig. 2, left panel). PCR amplicon targets of *Raphia peustera* were diluted to 0.0001, 0.001, 0.01, and 0.1 μM in 100 μL and added to the chip (see Fig. 2, right panel). Processes of hybridization and elution are given above.

Results and discussion

DNA barcoding and sequence analysis

One hundred and forty-one cytochrome c oxidase I gene sequences of 38 different Noctuoidea species from three different families were aligned, resulting in 69 unique

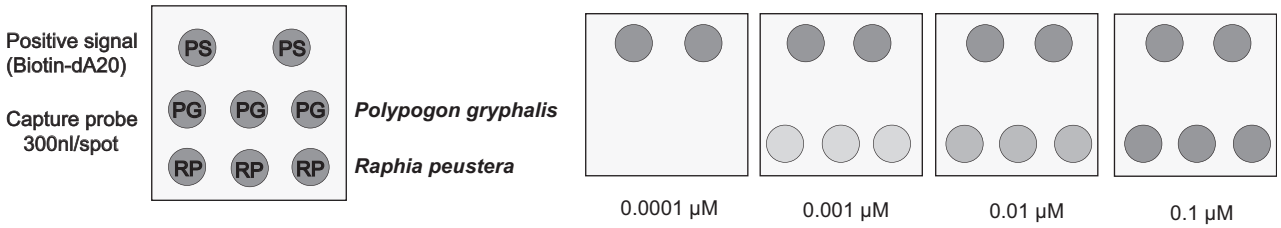

Fig. 2 Schematic drawing of specificity and sensitivity for the detection of Noctuoidea moths on the thin-film chip. Left panel: capture probes of PG (*Polypogon gryphalis*) and RP (*Raphia peustera*) spotted by hand at a volume of 300 nL and concentrations of 2 μM (left panel), positive signal PS Biotin-dA20 at 2 μM. Right panel: PCR amplicons of *R. peustera* at concentrations of 0.0001, 0.001, 0.01 and 0.1 μM, respectively, in 100 μL reaction solution hybridized to four identical chips.

haplotypes. All sequences correctly translated into amino acids; there were no sequencing errors nor pseudogenes. Final used sequences were 594 nucleotides long, of which 220 were variable. Haplotype diversity (Hd) was 0.966, and nucleotide diversity (Pi) was 0.102. The sequences comprised an average of A: 30.1%, T: 39.1%, C: 16.3%, and G: 14.6%.

The mean interspecies distance was 10.75 ± 1.20%. There was a clear DNA barcoding gap between the greatest intraspecies (1.02 ± 0.39%, *Athetis furvula*) and minimum interspecies genetic distances (4.6 ± 0.9%) (Appendix S4). The NJ tree (Fig. 1) showed that the 141 sequences could be clustered into 38 highly supported monophyletic groups, each constituting a single species. This COI barcoding approach can reliably identify the 141 individuals to species.

Design of probes

Fourteen species-specific oligonucleotide probes were designed and synthesized for 14 species of Noctuoidea shared by two nature reserve conservation zones. Probe lengths were 19–27 bp, and T_m values were 46.1 °C ~54.4 °C. There were 3–13 mismatches between the probes and nontarget sequences, and the maximum similarity between them was 86% (probe HC and *Catocala columbina*). The diagnostic regions (probes) of the 13 species were changed at one or two sites (Table 2), and we found that introducing one or two mismatch bases did not hinder the hybridization between probes and the sequences of target species. All probes based on COI sequences are accurate and valid to identify the moths under study. Detailed information about the probes is given in Table 2.

Coding regions are more conserved than noncoding regions and show a high degree of similarity with the coding regions of other closely related genes. Thus, it has been claimed that probes selected from coding regions are more susceptible to cross-hybridization events (Liu *et al.* 2010). However, oligonucleotide probes based on coding genes like COI and *cytb* have given unambiguous signals and have been used in several studies for species identification (Pfunder *et al.* 2004; Hajibabaei *et al.* 2007; Li *et al.* 2008; Chung *et al.* 2010; Kim *et al.* 2011; Lee *et al.* 2011). Furthermore, unlike many noncoding genes, the COI gene can be amplified by universal primers (Hebert *et al.* 2003b) and lacks insertions or deletions (Doyle & Gaut 2000), facilitating PCR experiments.

Defining sensitivity and specificity of the thin-film biosensor chips

Sensitivity and specificity of the thin-film biosensor chips are shown in Fig. 2. As shown in the left panel, 300 nL of each spot at concentrations of 2 μM of probes PG (*Polypogon gryphalis*) and RP (*Raphia peustera*) were manually spotted on the chips' surfaces. Four identical chips were then hybridized with *R. peustera* PCR products at concentrations of 0.0001, 0.001, 0.01 and 0.1 μM in 100 μL total reaction volume. As shown, only the spots containing the RP probe (to the exclusion of the PG probe) were detected, indicating a remarkable specificity. Although the signal intensity decreased as the target sequence concentration was lowered, as little as 0.001 μM of target (0.1 pmol DNA) was detected by the probe. Sequencing of the barcode COI gene (about 650 bp) generally needs not <600 ng DNA, but in this assay, 50 ng DNA is enough for hybridization.

Diagnostic ability of the thin-film biosensor chips and target species discrimination

Hybridization experiments were performed with 14 target and 24 nontarget Noctuoidea moth species. Results from 14 target moth species of 13 genera (*Hypena stygiana* and *Hypena tristalis* belong to the same genus) are shown in Fig. 3. As shown, every target species probe has two repetitions on the chip. All signals could be clearly detected by the human eye without the help of instruments. No nontarget species lit up any target probes. However, results from the thin-film biosensor chips, analogous to the results of various kinds of

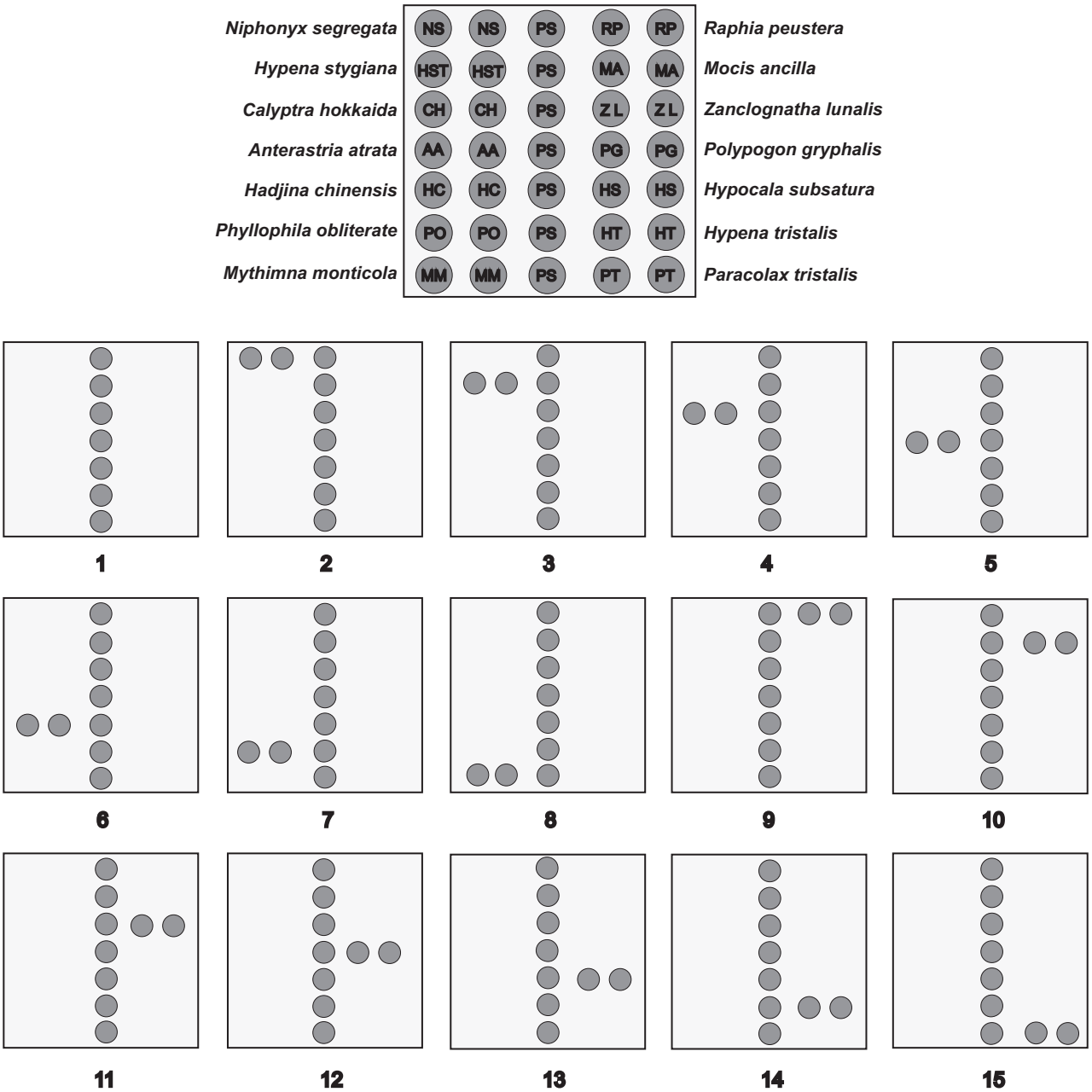

Fig. 3 Schematic drawing of 14 target species of Noctuoidea moths detected by the thin-film chip. Capture probes of 14 moth species were spotted by a computer-controlled robot arm at a volume of 40 nL and concentrations of 2 μM. The order probes were printed is shown in the upper panel. Hybridization results are shown in the lower panel: 1 PS (positive signal); 2 NS (*Niphonyx segregata*); 3 HST (*Hypena stygiana*); 4 CH (*Calyptra hokkaida*); 5 AA (*Anterastria atrata*); 6 HC (*Hadjina chinensis*); 7 PO (*Phyllophila obliterate*); 8 MM (*Mythimna monticola*); 9 RP (*Raphia peustera*); 10 MA (*Mocis ancilla*); 11 ZL (*Zanclognatha lunalis*); 12 PG (*Polypogon gryphalis*); 13 HS (*Hypocala subsatura*); 14 HT (*Hypena tristalis*); 15 PT (*Paracolax tristalis*).

microarrays used in previous studies (Peplies *et al.* 2003; Warsen *et al.* 2004; Rønning *et al.* 2005; Shi *et al.* 2006; Tobler *et al.* 2006), were compromised by the fluorescence intensities of true-positive hybridization being heterogeneous. A large number of parameters are able to influence hybridization signal: sequence characteristics, secondary structure, chip quality, solution concentration, the relative positions of the fluorescence signal and probe binding site, steric hindrance and so on (Southern *et al.* 1999; Zhang *et al.* 2005; Kochzius *et al.* 2008, 2010). It is thus hard to point to any specific reason for the observed diversity of signal intensities. However, this heterogeneity was not a concern here because all hybridization events were successful and

the chip signals were all clear and accurate, without any false-positive hybridization.

Comparison of thin-film biosensor chips and DNA barcoding

In this study, both thin-film biosensor chips and DNA barcoding provide accurate identifications of 14 Noctuoidea species. Compared with DNA barcoding which is a sequencing-based technology, the procedure for using a thin-film biosensor chip – a technological development from microarrays – is straightforward and rapid, yet provides high sensitivity and specificity when identifying moths in this study. The whole process needs only 4 h to get the final results, which will remain visible for several years on the chip surface. All necessary chemical agents in the experiment are readily available and inexpensive. An important advantage of the thin-film biosensor chip is the optical characteristic of the chip surface, permitting the hybridization results to be visualized by the unaided human eye. This technology eliminates dependence on large and expensive instrumentation and reduces the cost of experimental budgets: it can be readily distributed to any individual research laboratory for the identification of moths or other species within a particular geographic range.

The main limitation of the array-based approach is that it requires prior knowledge of sequences from target species (Hajibabaei *et al.* 2007); thin-film biosensor chips that do not include sequences of the target species cannot work. Under these circumstances, however, we can still use the DNA barcoding method to identify or flag species by comparing them with known sequences. That means that good taxon coverage of the database is crucial in the successful application of thin-film biosensor chips because changing taxon coverage will influence the designed probe's diagnostic capability, and any false-positive signals detected by the human eye can give misleading results.

Our study focused on a limited number of species from a rather small geographic region and the sequences used in this study possess low intraspecific and high interspecific variation, so it proved possible to identify successfully every targeted species by only one visually chosen probe on our low-density chip. When the number of the sequences increases with greater geographic coverage and higher taxonomic diversity, chips will possess higher practicality but the difficulty of designing species-specific probes increases. Hence, most studies using array-based technology have only focused on a limited number of species. However, designing more probes with different gene markers for every species and introducing mismatch bases properly will help to increase a chip's ability to identify species. For instance, Hajibabaei *et al.* (2007) designed three probes per species for 121 species of mammals from a large geographic region using two gene markers (COI and *Cytb*), with a resulting high degree (~95%) of identification success.

In conclusion, we have shown that we could accurately identify 14 Noctuoidea moth species using thin-film biosensor chips and confirm that both this and the DNA barcoding approach are valuable molecular technologies for species identification. Oligonucleotide probes based on COI sequences are accurate and are able to discriminate moths at the species level. Thin-film biosensor chips possess high sensitivity and specificity, and their low expense and high accuracy make them suitable for species identification. Now, we plan to use additional and taxonomically broader sequences to construct a more practical and useful chip that we can use for identifying more species from a wider geographic range. Such a chip covering a wide range of Lepidoptera (and perhaps other insect species) will facilitate species identification at any life history stage and will provide valuable assistance for biological pest control and plant protection programmes.

Acknowledgements

This study was supported by Natural Science Foundation of China (to Zhang, Grant No. 31071963 and 31272340), by Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality (to Zhang, Grant No. PHR201107120), by The Research Fund for the Doctoral Program of Higher Education (to Zhang, Grant No. 20101108120002).

Conflict of interests

The authors declare that they have no conflict of interests.

References

Bai SL, Zhong XB, Ma LG *et al.* (2007) A simple and reliable assay for detecting specific nucleotide sequences in plants using optical thin-film biosensor chips. *Plant Journal*, **49**, 354–366.

Bai SL, Zhao JY, Zhang YC *et al.* (2010a) Rapid and reliable detection of 11 food-borne pathogens using thin-film biosensor chips. *Applied Microbiology and Biotechnology*, **86**, 983–990.

Bai SL, Zhang J, Li SC *et al.* (2010b) Detection of six genetically modified maize lines using optical thin-film biosensor chips. *Journal of Agricultural and Food Chemistry*, **58**, 8490–8494.

Bai SL, Xu L, Guo XM *et al.* (2012) A rapid approach for detecting seven genetically modified canola events using optical thin-film biosensor chips. *Journal of Biochips & Tissue Chips*, **2**, 101.

Baker GH, Tann CR (2013) Mating of *Helicoverpa armigera* (Lepidoptera: Noctuidae) moths and their host plant origins as larvae within Australian cotton farming systems. *Bulletin of Entomological Research*, **103**, 171–181.

Blohm D, Guiseppi-Elie A (2001) New developments in microarray technology. *Current Opinion in Biotechnology*, **12**, 41–47.

Chi MY, Han HL, Gao Q *et al.* (2012) Effects of different evolution models on DNA barcoding evaluated with the Noctuidae from Wuling Mountain, Hebei, northern China. *Acta Entomologica Sinica*, **55**, 1193–1204.

Chung IH, Yoo HS, Eah JY *et al.* (2010) A DNA microarray for identification of selected Korean birds based on mitochondrial cytochrome *c* oxidase i gene sequences. *Molecules and Cells*, **30**, 295–301.

Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**, 407–415.

Doyle JJ, Gaut BS (2000) Evolution of genes and taxa: a primer. *Plant Molecular Biology*, **42**, 1–23.

Garaizar J, Rementeria A, Porwollik S (2006) DNA microarray technology: a newtool for the epidemiological typing of bacterial pathogens? *FEMS Immunology and Medical Microbiology*, **47**, 178–189.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006a) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences*, **103**, 968–971.

Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, Hebert PDN (2006b) A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes*, **6**, 959–964.

Hajibabaei M, Singer GAC, Clare EL, Hebert PDN (2007) Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. *BMC Biology*, **5**, 24.

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.

Hausmann A, Haszprunar G, Segerer AH, Speidel W, Behounek G, Hebert PDN (2011) Now DNA-barcoded: the butterflies and larger moths of Germany. *Spixiana*, **34**, 47–58.

He ZL, Wu LY, Li XY, Fields MW, Zhou JZ (2005) Empirical establishment of oligonucleotide probe design criteria. *Applied and Environmental Microbiology*, **71**, 3753–3760.

Hebert PDN, Ratnasingham S, Dewaard JR (2003a) Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. *Proceedings of the Royal Society B (Supply)*, **270**, s96–s99.

Hebert PDN, Cywinska A, Ball SL, Dewaard JR (2003b) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B*, **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*, **101**, 14812–14817.

Hughes TR, Mao M, Jones AR *et al.* (2001) Expression profiling using microarrays fabricated by an ink-jet oligonucleotide Synthesizer. *Nature Biotechnology*, **19**, 342–347.

Hulcr J, Miller SE, Setliff GP *et al.* (2007) DNA barcoding confirms polyphagy in a generalist moth, *Homona mermerodes* (Lepidoptera: Tortricidae). *Molecular Ecology Notes*, **7**, 549–557.

Ikuta S, Takagi K, Wallace RB, Itakura K (1986) Dissociation kinetics of 19 base paired oligonucleotide-DNA duplexes containing different single mismatched base pairs. *Nucleic Acids Research*, **15**, 797–811.

Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E, Hebert PDN (2005) Wedding biodiversity inventory of a large and complex *Lepidoptera fauna* with DNA barcoding. *Proceedings of the Royal Society B*, **360**, 1835–1845.

Joomun N, Ganeshan S, Dookun-Saumtally A (2012) Identification of three armyworm species (Lepidoptera: Noctuidae) using DNA barcodes and restriction enzyme digestion. *International Sugar Journal*, **114**, 344–349.

Kandagal AS, Khetagoudar MC (2013) Study on larvicidal activity of weed extracts against *Spodoptera litura*. *Journal of Environmental Biology*, **34**, 253–257.

Kane MD, Jatkoe TA, Stumpf CR, Lu J, Madore SJ (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Research*, **28**, 4552–4557.

Kim JH, Park JY, Jung JW *et al.* (2011) Species identification of filefishes (Monacanthidae) using DNA microarray in Korean marketplace. *Biochip Journal*, **5**, 229–235.

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.

Kircher M, Kelso J (2010) High-throughput DNA sequencing – concepts and limitations. *BioEssays*, **32**, 524–536.

Kochzius M, Nölte M, Weber H *et al.* (2008) DNA microarrays for identifying fishes. *Marine Biotechnology*, **10**, 207–217.

Kochzius M, Seidel C, Antoniou A *et al.* (2010) Identifying fishes through DNA barcodes and microarrays. *PLoS One*, **5**, e12620.

Korimbocus J, Scaramozzino N, Lacroix B, Crance JM, Garin D, Vernet G (2005) DNA probe array for the simultaneous identification of herpesviruses, enteroviruses, and flaviviruses. *Journal of Clinical Microbiology*, **43**, 3779–3787.

Lafontaine JD, Poole RW (2010) Review of the new world genera of the subfamily Acontiinae (Lepidoptera, Noctuidae). *ZooKeys*, **39**, 137–160.

Lastuvka Z, Lastuvka A, Lopez-Vaamonde C (2013) A revision of the *Phyllonorycter ulicicolella* species group with description of a new species (Lepidoptera: Gracillariidae). *SHILAP – Revista De Lepidopterologia*, **41**, 251–265.

Lee I, Dombkowski AA, Athey BD (2004) Guidelines for incorporating non-perfectly matched oligonucleotides into target-specific hybridization probes for a DNA microarray. *Nucleic Acids Research*, **32**, 681–690.

Lee G, Park SY, Hwang J *et al.* (2011) Development of DNA chip for jellyfish verification from South Korea. *Biochip Journal*, **5**, 375–382.

Li WF, Yu DJ, Yan HM, Li JG, Xu L, Ren LF (2008) The gene chip detection technique for the Mediterranean fly, *Ceratitis capitata*, and its related species (Diptera:Tephritidae). *Acta Entomologica Sinica*, **51**, 61–67.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Liu HF, Bebu I, Li X (2010) Microarray probes and probe sets. *Frontiers in Bioscience (Elite Edition)*, **2**, 325–338.

Lu YJ, Kochert GD, Isenhour DJ, Adang MJ (1994) Molecular characterization of a strain-specific repeated DNA sequence in the fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Insect Molecular Biology*, **3**, 123–130.

Luo KJ, Gu DX, Zhang GR, Chen ZQ (2004) Parasitoids of four noctuidae pests in cruciferous vegetable. *Chinese Journal of Biological Control*, **20**, 211–214.

Martin Y, Gerlach G, Schlötterer C, Meyer A (2000) Molecular phylogeny of European muroid rodents based on complete cytochrome b sequences. *Molecular Phylogenetics and Evolution*, **16**, 37–47.

Metzker ML (2010) Sequencing technologies-the next generation. *Nature Reviews Genetics*, **11**, 31–46.

Mitchell A, Mitter C, Regier JC (2006) Systematics and evolution of the cutworm moths (Lepidoptera: Noctuidae): evidence from two protein-coding nuclear genes. *Systematic Entomology*, **31**, 21–46.

Mohamed HF (2013) Bioenergetics growth model for the effect of gamma irradiation and plant extract (Barnoof) on the progeny of Black cutworm, *Agrotis ipsilon* (Hufnagel). *Applied Radiation and Isotopes*, **73**, 101–108.

Moreira GRP, Goncalves GL, Eltz RP, San Blas G, Davis DR (2012) Revalidation of *Oliera Brethes* (Lepidoptera: Cecidosidae) based on a redescription of *O. argentinana* and DNA analysis of Neotropical cecidosids. *Zootaxa*, **3557**, 1–19.

Ostroff RM, Hopkins D, Haeberli AB, Baouchi W, Polisky B (1999) Thin film biosensor for rapid visual detection of nucleic acid targets. *Clinical Chemistry*, **45**, 1659–1664.

Owczarzy R, Vallone PM, Gallo FJ (1997) Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers*, **44**, 217–239.

Peplies J, Glöckner FO, Amann R (2003) Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA targeting oligonucleotide probes. *Applied and Environmental Microbiology*, **69**, 1397–1407.

Pfunder M, Holzgang O, Frey JE (2004) Development of microarray-based diagnostics of voles and shrews for use in biodiversity monitoring studies, and evaluation of mitochondrial cytochrome oxidase I vs. cytochrome b as genetic markers. *Molecular Ecology*, **13**, 1277–1286.

Rønning SB, Rudi K, Berdal KG, Holst-Jensen A (2005) Differentiation of important and closely related cereal plant species (Poaceae) in food by hybridisation to an oligonucleotide array. *Journal of Agricultural and Food Chemistry*, **53**, 8874–8880.

Schmidt BC (2009) Taxonomic revision of the genus *Grammia* Rambur (Lepidoptera: Noctuidae: Arctiinae). *Zoological Journal of the Linnean Society*, **156**, 507–597.

Schmidt BC, Sperling FAH (2008) Widespread decoupling of mtDNA variation and species integrity in *Grammia* tiger moths (Lepidoptera: Noctuidae). *Systematic Entomology*, **33**, 613–634.

Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nature Biotechnology*, **26**, 1135–1145.

Shi L, Reid L, Jones WD, Shippy R (2006) The microarray quality control (MAQC) project shows inter- and intra-platform reproducibility of gene expression measurements. *Nature Biology*, **24**, 1151–1161.

Smigocki AC, Ivic-Haymes S, Li HY, Savic J (2013) Pest protection conferred by a beta vulgaris serine proteinase inhibitor gene. *PLoS One*, **8**, e57303.

Southern E, Mir K, Shchepinov M (1999) Molecular interactions on microarrays. *Nature Genetics Supplement*, **21**, 5–9.

Tambong JT, De Cock AW, Tinker NA, Le'Vesque CA (2006) Oligonucleotide array for identification and detection of *pythium* species. *Applied and Environmental Microbiology*, **72**, 2691–2706.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

Tobler NE, Pfunder M, Herzog K, Frey JE, Altwegg M (2006) Rapid detection and species identification of *Mycobacterium* spp. Using real-time PCR and DNA-microarray. *Journal of Microbiol Methods*, **66**, 116–124.

Tomik S, Hofmann K (2001) Microarray probe selection strategies. *Briefings in Bioinformatics*, **2**, 329–340.

Urich T, Lanzén A, Qi J, Huson DH, Schleper C, Schuster SC (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS One*, **3**, e2527.

Warsen AE, Krug MK, Lafrentz S, Stanek DR, Loge F, Call DR (2004) Simultaneous discrimination between 15 fish pathogens by using 16S ribosomal DNA PCR and DNA microarrays. *Applied and Environmental Microbiology*, **70**, 4216–4221.

Wei GJ, Zou BJ, Chen ZY, Song QX, Li YZ, Zhou GH (2012) Advances in next-generation sequencing technologies. *Progress in Modern Biomedicine*, **12**, 3789–3727.

Werntges H, Steger G, Riesner D, Fritz HJ (1986) Mismatches in DNA double strands: thermodynamic parameters and their correlation to repair efficiencies. *Nucleic Acids Research*, **14**, 3773–3790.

Yang CH, Han HL, Chi MY *et al*. (2012) Species identification of Noctuidae moths (Insecta: Lepidoptera) from Baihuashan. *Acta Entomologica Sinica*, **55**, 1082–1092.

Zahiri R, Kitching IJ, Lafontaine JD, Mutanen M, Kaila L, Holloway DJ, Wahlberg N (2010) A new molecular phylogeny offers hope for a stable family level classification of the Noctuoidea (Lepidoptera). *Zoologica Scripta*, **40**, 158–173.

Zahiri R, Lafontaine D, Schmidt C *et al*. (2013) Relationships among the basal lineages of Noctuidae (Lepidoptera, Noctuoidea) based on eight gene regions. *Zoologica Scripta*, **42**, 488–507.

Zhang L, Hurek T, Reinhold-Hurek B (2005) Position of the fluorescent label is a crucial factor determining signal intensity in microarray hybridizations. *Nucleic Acids Research*, **33**, e166.

Zhang YJ, Yin J, Li GF, Li MF, Huang X, Chen HJ, Zhao WJ, Zhu SF (2010) Oligonucleotide microarray with a minimal number of probes for the detection and identification of thirteen genera of plant viruses. *Journal of Virological Methods*, **167**, 53–60.

Zhao JY, Bai SL, Huang WS, Chen Y (2008) Detection and identification of food-borne pathogen by using thin-film biosensor chips. *Food and Fermentation Industries*, **34**, 141–144.

Zhong XB, Reynolds R, Kidd JR *et al*. (2003) Single-nucleotide polymorphism genotyping on optical thin-film biosensor chips. *Proceedings of the National Academy of Sciences*, **100**, 11559–11564.

A.B.Z. and F.Y. designed the study, F.Y. performed molecular work, F.Y., Z.Y.S., S.L.B. analysed data, F.Y., R.D.W. and A.B.Z. wrote the draft of the manuscript.

Data accessibility

All sequences used in this study can be accessed on the NCBI database. The detailed information about the sequences is listed on Table 1. Appendix S1 comprises an alignment file of the 144 sequences used in this article.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 An alignment file of 144 sequences used in this article.

Appendix S2 The digital imaged taken by camera (Panasonic, DMC-ZS1) of Fig. 2.

Appendix S3 The digital imaged taken by camera (Panasonic, DMC-ZS1) of Fig. 3.

Appendix S4 Intraspecific distance of 38 species used in study.