DNA barcodes for two scale insect families, mealybugs (Hemiptera: Pseudococcidae) and armored scales (Hemiptera: Diaspididae)

D.-S. Park[1], S.-J. Suh[2], P.D.N. Hebert[3], H.-W. Oh[4]* and K.-J. Hong[5]*

[1]Biological Resource Center, KRIBB, Daejeon 305-806, Korea: [2]Youngnam Regional Office, National Plant Quarantine Service, Busan 600-016, Korea: [3]Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, N1G 2W1, Canada: [4]Industrial Biomaterials Research Center, KRIBB, 111 Gwahangno, Yuseong-gu, Daejeon 305-806, Korea: [5]Central Post-entry Quarantine Station, National Plant Quarantine Service, 433-1 Anyang 6-dong Manan-gu Anyang-shi, Gyeonggi Province, 430-016, Korea

Abstract

Although DNA barcode coverage has grown rapidly for many insect orders, there are some groups, such as scale insects, where sequence recovery has been difficult. However, using a recently developed primer set, we recovered barcode records from 373 specimens, providing coverage for 75 species from 31 genera in two families. Overall success was >90% for mealybugs and >80% for armored scale species. The G·C content was very low in most species, averaging just 16.3%. Sequence divergences (K2P) between congeneric species averaged 10.7%, while intra-specific divergences averaged 0.97%. However, the latter value was inflated by high intra-specific divergence in nine taxa, cases that may indicate species overlooked by current taxonomic treatments. Our study establishes the feasibility of developing a comprehensive barcode library for scale insects and indicates that its construction will both create an effective system for identifying scale insects and reveal taxonomic situations worthy of deeper analysis.

Keywords: DNA barcoding, COI, mealybug, armored scale, NUMTs, G·C content

Introduction

The scale insects (Hemiptera: Coccoidea) are a major group of pests on crops and ornamental plants worldwide. They include more than 7700 described species in 1050 genera (Ben-Dov *et al.*, 2006) belonging to 31 extant and 14 fossil families (Kondo *et al.*, 2008). Nearly all scale insects can be identified only morphologically to a species level by examining characters present in adult females. However, even when adult females are examined, it is sometimes difficult to distinguish between very similar species and determining if slight differences represent inter- or intra-specific variation. Studies of sequence variation in several gene regions have been carried out in an attempt to resolve this problem. Initial efforts examined three nuclear genes: the D2 region of 28S rDNA, elongation factor 1α (Downie & Gullan, 2004; Morse & Normark, 2006) and ITS (Beuning *et al.*, 1999; Park *et al.*, 2010b). ITS sequences were able to discriminate closely related species, but the high incidence of indels made sequence

---

*Authors for correspondence
Fax: +82-42-860-4677 (H.-W.O); +82-31-420-7606 (K.-J.H.)
E-mail: hwoh@kribb.re.kr; curcul@hanmail.net

alignment difficult except for closely allied taxa. The 28S sequences were easier to align but lacked enough variation to resolve some species. Another investigation examined sequence variation in the 16S–23S rDNA of the vertically transmitted endosymbiotic microbe, *Tremblaya princeps* (Thao *et al.*, 2002; Baumann & Baumann, 2005), as a tool for scale insect identification, but it too lacked enough variation. A more recent study tested the possibility of identifying mealybug species through the analysis of four markers: two nuclear genes, one mitochondrial gene and one gene from the endosymbiont (Malausa *et al.*, 2010). The standard barcode region of cytochrome *c* oxidase subunit I (COI) gene was targeted for analysis in this study, but the standard primer set (Folmer *et al.*, 1994) only generated an amplicon in a few species. It proved possible to recover more species with a new primer set that amplified a 491 bp segment of the barcode region, but these records are too short to qualify for recognition as formal barcodes. Nonetheless, the results were promising enough to provoke further investigation.

DNA barcoding aims to identify species through the analysis of sequence variation in a standardized DNA sequences. A 658 bp segment near the 5′ end of the mitochondrial COI has been accepted as the core barcode region for the animal kingdom (Hebert & Gregory, 2005; Hebert *et al.*, 2003b). Many studies have now established that this region possesses enough sequence variation to discriminate most animal species (Hajibabaei *et al.*, 2006; Hebert *et al.*, 2003a). However, in certain groups, such as scale insects, it has proven difficult to amplify the barcode region, suggesting mismatches between the primers and the target DNA. A recent study confirmed this fact for certain scale insects, revealing two substitutions at critical positions near the 3′ binding site for the forward primer. This observation allowed the design of a new forward primer, which showed high success when tested on a few species of scale insects (Park *et al.*, 2010a). The present study examines the utility of this primer set in a more comprehensive way, by testing its performance on 88 species from the two most diverse families of scale insects, the Pseudococcidae (mealybugs) and the Diapsidae (armored scales). Aside from testing amplification success, this investigation considers both the utility of the resultant sequences in scale insect identification and certain unusual properties of these sequences.

Materials and methods

Specimens

Most of the specimens examined in this study were obtained from collections made by the National Plant Quarantine Service in South Korea. However, specimens of *Planococcus ficus* were provided by the Plant Protection Research Institute of South Africa, and some *Planococcus kraunhiae* and *Crisicoccus matsumotoi* were contributed by the Shimane Agricultural Technology Center, Japan. In total, 524 specimens were analyzed, representing 88 species from two families of scale insects. Whenever possible, at least five specimens were analyzed of each species from each of the localities where it was collected.

DNA extraction, PCR and sequencing

DNA was extracted from most individuals using a standard glass fiber extraction protocol (Ivanova *et al.*, 2006);
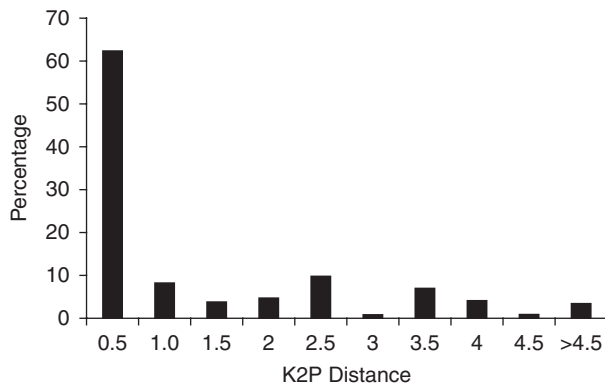

Fig. 1. Distribution of intra-specific genetic distances (K2P) at COI 5′ region for 343 individuals of 68 species of scale insects. Mean K2P distance of a species was calculated using the sequence analysis in BOLD and percentage of species within certain intra-specific value was represented.

the remainder (9.4%) were extracted using a DNeasy kit (Qiagen, Hilden, Germany) following the manufacturer's protocol, except that the final elution step was performed with 60 µl of distilled water instead of 200 µl buffer. Specimen cuticles were recovered after DNA extraction and then stored in ethanol until slide mounted. All of the vouchers from this study are deposited in the scale insect collection of National Plant Quarantine Service (NPQS), Busan, Korea. A 649-bp segment of the barcode region was amplified from most specimens using the primer PcoF1-5′CCTTCAACTAATCA-TAAAAATATYAG3′ (Park *et al.*, 2010a) and LepR1-5′ TAAACTTCTGGATGTCCAAAAAATCA3′. A few specimens were amplified using two alternate primer sets (LepF1/LepR or C_tRWF/LepR1) before the PcoF1 primer was developed. PCRs were performed using a Maxime® PCR PreMix (iNtRON Biotechnology, Seongnam, Korea) with 2.0 ρmol of each primer and 2–50 ng of template DNA in a 20 µl reaction, or a standard 12.5 µl reaction developed by Ivanova *et al.* (2006). PCR thermocycling was done under the following conditions: 2 min at 95°C; five cycles of 40 s at 94°C; 40 s at 45°C, 60 s at 72°C; 40 cycles of 40 s at 94°C; 40 s at 51°C; 60 s at 72°C; 5 min at 72°C; held at 4°C. PCR products were visualized in a 2% agarose gel stained with ethidium bromide and bidirectionally sequenced using a BigDye Terminator ver. 3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster, CA) on an ABI 3730XL capillary sequencer. Contigs were assembled using CodonCode Aligner ver 3.5.6 (CodonCode Co., Dedham, MA) and were subsequently aligned using the same software or MEGA (Tamura *et al.*, 2007). All sequences were deposited in GenBank (Accession numbers HM474068–HM474409), while the collection data are available in the DBSI project file in BOLD (http://www.barcodinglife.org). Sequence divergences were calculated using the Kimura 2-parameter (K2P) model (Kimura, 1980), and a neighbour-joining (NJ) tree was generated by MEGA or by the Taxon ID tree function in BOLD (Ratnasingham & Hebert, 2007).

Results

A DNA barcode record, almost 649 bp in length, was recovered from 343 of 524 specimens, a 65% success rate. These sequences were derived from 75 of the 88 species and included

Table 1. Genetic divergences (K2P) among members of a species and between congeneric species of scale insects.

	Taxa	Comparisons	Min. Dist (%)	Mean Dist (%)	Max. Dist (%)	SE Dist (%)
Within Species	57	1474	0	0.97	5.98	0.04
Within Genus	30	2263	1.93	10.09	23.29	0.08

Table 2. Variation in G·C content for the barcode region of COI in two families of scale insects.

	Min.		Mean		Max.		SE	
	D	P	D	P	D	P	D	P
Total G·C%	15.5	12.9	18.2	14.9	33.5	17.8	0.24	0.09
G·C% at Pos. 1	14.6	12.2	18.2	15.9	31.6	19.4	0.22	0.09
G·C% at Pos. 2	26.2	24.8	28.4	26.1	31.4	28.6	0.12	0.04
G·C% at Pos. 3	1.4	0	8.2	4.2	38.6	10.4	0.48	0.2

D, diaspididae; P, pseudococcidae.

members of 32 genera. So, the success in barcode recovery per species level exceeded 90% for mealybug and 80% for armored scales. We recovered 30 additional sequences from seven species of mealybugs that possessed a 1 or 2 bp indel, suggestive of pseudogenes. These cases, which involved the deletion of one or two thymine residue deletions, were detected in every individual of seven mealybug species (*Dysmicoccus wistariae*, *Ferrisia malvastra*, *Ferrisia virgata*, *Palmicultor lumpurensis*, *Pseudococcus calceolariae*, *Pseudococcus jackbeardsleyi*, *Pseudococcus viburni* and an unidentified *Pseudococcus* sp.). These deletions all occurred at the same position (149–150) in the standard COI barcode region although an additional deletion was also present at position 476 in all seven individuals of *P. jackbeardsleyi* (S2). These sequences are included in the compressed phylogenetic tree (fig. 1) and were submitted to GenBank with a pseudogene notation.

The G·C content of the COI gene regions from the scale insects examined in this study are very low, averaging just 16.4%. Species of mealybugs have a lower G·C content (14.9%) than those of armored scales (18.2%), but there was significant variation in G·C content among the members of each family (table 2). The mealybugs *Atrococcus paludinus* and *Paracoccus marginatus* have the lowest values (12.6% and 12.9%, respectively), while the armored scales (*Odonaspis secrata* and *Pseudaonidia duplex*) have relatively high G·C content (30.2% and 33.5%, respectively). Mean G·C content was 18.2% and 28.4% in armored scales and 15.9% and 26.1% in mealybugs at codon positions 1 and 2, respectively, but just 8.2% in armored scales and 4.2% in mealybugs at position 3.

Figure 1 provides a compressed K2P/NJ tree of the patterns of sequence divergence among the 373 specimens with a barcode record while the entire tree is available in S3. Sequence divergences (K2P distance) between congeneric species averaged 10.09%, a value which was tenfold higher than the mean intra-specific variation (table 1). Some overlap was observed in the distribution of K2P distances among conspecific individuals (range 0–5.98%) and among congeneric species (range 1.93–23.29%). Although most species showed less than 1.0% divergence, 25% of them

showed over 2.0% divergence (fig. 2), cases that suggest several overlooked species. Particularly deep divergences (2.5–4.6%) were found within three armored scale species (*Lopholeucaspis japonica*, *Aspidiotus excises* and *Pseudaulacaspis pentagona*). Six species of mealybug (*Crisicoccus matsumotoi*, *Maconellicoccus hirsutus*, *Phenacoccus aceris*, *Phenacoccus solani*, *Planococcus ficus* and *Pseodococcus longispinus*) also showed deep divergences (2.5–5.8%) (fig. 3).

Discussion

DNA barcoding has gained increasing usage for the discrimination of animal species, but its application has proven difficult for certain groups, such as scale insects, because of inconsistent PCR amplification of the target COI 5′ region. Park *et al*. (2010a) recently developed a new forward primer (PcoF1), which, when combined with a standard reverse primer (LepR1), successfully amplified several scale insect species. This study involved a larger-scale test examining barcode recovery in 88 species of mealybugs and armored scales. Barcodes were recovered from 343 of 524 specimens (71%), a lower success rate than in the earlier study (92%) (Park *et al*., 2010a). We suspect that this difference reflects the use of alternate DNA extraction methods because we had 58% success for 190 armored scale specimens and 76% success for 285 mealybug specimens using a Silica-based extraction method (Ivanova *et al*., 2006) but 95% success when DNA was extracted using the DNeasy kit (ten of ten armored scales and 37 of 39 mealybugs), a success rate similar to that reported by Park *et al*. (2010a).

Park *et al*. (2010a) reported that scale insects have a characteristic deletion of three amino acids in the barcode region. The same deletion was detected in all 75 species examined in this study, suggesting that this deletion is shared by many, if not all, species in at least these two families of scale insects. The present study revealed one other anomaly; seven of the 40 mealybug species belonging to three different genera had a deletion of 1–2 thymine residues at position 149–150 in the COI 5′ barcode region (fig. 2). Because these deletions induced a frame shift mutation which disrupts translation of the COI protein, we expect that these sequences reflect the recovery of a NUMT. If so, the conserved location of the deletion suggests that all cases reflect an ancient transposition event from the mitochondrion to the nucleus that occurred before diversification of the three genera that now carry it. Given this shared ancestry of lineages carrying the NUMTs, one would expect them to form a cohesive cluster in the NJ tree, but they do not (fig. 2). Alternatively, deletions at this position may have evolved recurrently after nuclear transposition, reflecting an inherent tendency toward the loss of a thymine residue in a poly-T region of this gene in scale insects. Interestingly, a thymine insertion was detected in a species of Coelostomidiidae, another family of scale insects, at the same position (Kaur, M., personal communication). However, it
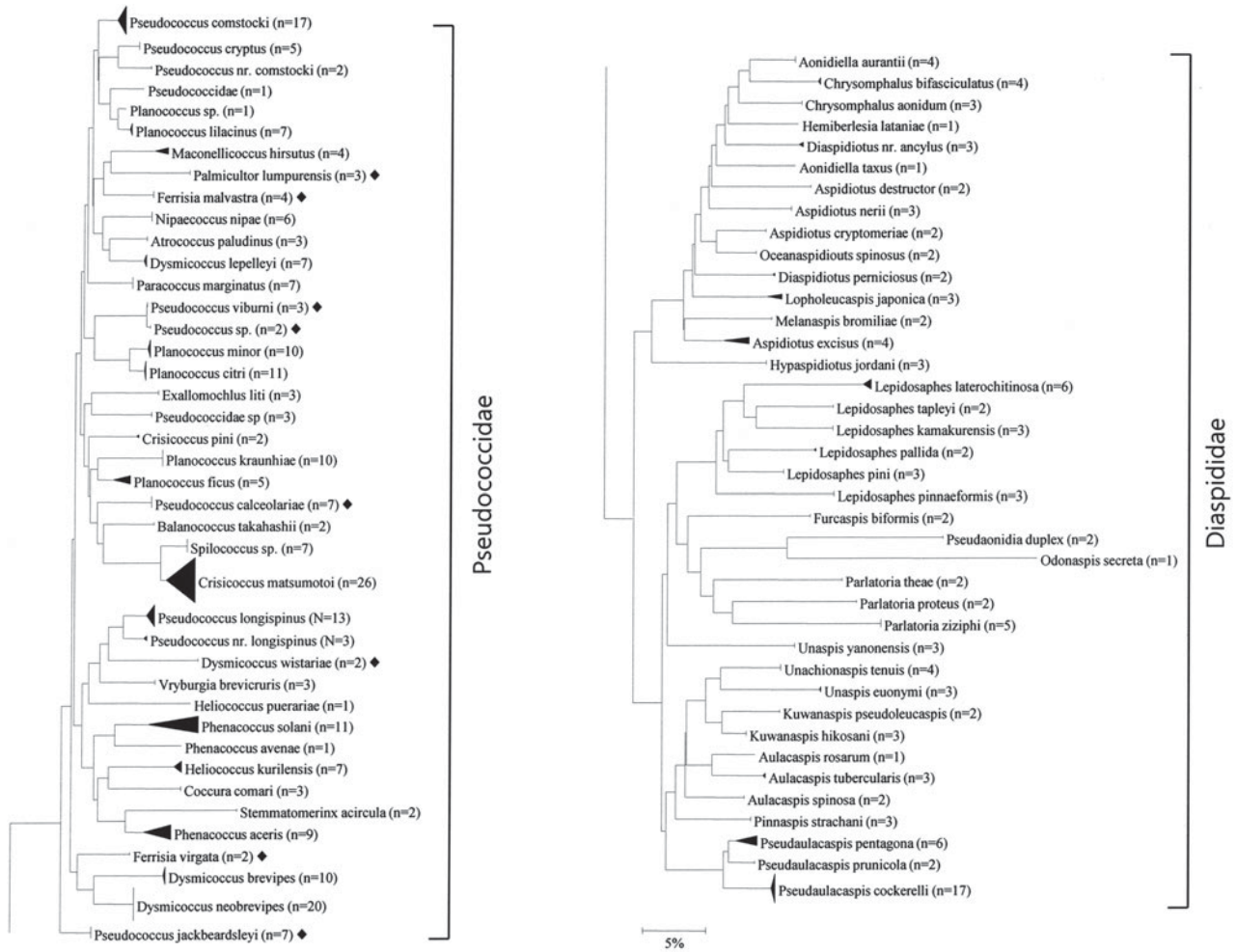
Fig. 2. A compressed NJ tree of COI sequence divergences (K2P) in 75 mealybug and armored scale species. The number of specimens is indicated following each species name. ♦, sequences which have internal thymine residue deletion.

remains important to rule out the possibility that these sequences represent the authentic mitochondrial sequence with the translation mechanism adjusting the reading frame. The recovery of a full mitochondrial genome sequence from these species is required to rule out the possibility that these deletions occur in the mitochondrial genome itself.

The G·C content of the barcode region of scale insects is very low, averaging just 16.4% versus typical values of 33–53% in other animal lineages (Min & Hickey, 2007). As expected, the low G·C frequency is more pronounced at the third codon position (8.2% for armored scales, 4.2% for mealybugs), than at the first and second positions (18.2% and 28.4% in armored scales versus 15.9% and 26.1% in mealybugs) (table 2). The marsh mealybug, *A. paludinus*, has no G·C at the third position in the standard barcode region, while the papaya mealybug, *P. marginatus*, has only a single G·C base. The average G·C content (14.9%) for the barcode region in mealybugs (Pseudococcidae) is the lowest value for any insect family, while the 12.6% G·C content for *A. paludinus* is the lowest value known for any animal species. These values are extreme even when viewed from the context of bacteria genomes whose G·C contents range from 17–75%. The lowest bacterial

values occur in species with small genome sizes (Andersson & Kurland, 1998), especially the endosymbionts of insects, such as aphids and scale insects, which possess G·C contents of 17–33% (Moran & Wernegreen, 2000; Perez-Brocal *et al.*, 2006; Moran *et al.*, 2008). Interestingly, a low G·C content (23–27%) also exists in the barcode region of aphids, animals that feed, like scale insects, on plant sap, a diet which is very deficient in organic nitrogen. Might the low G·C content in aphids, scale insects and their endosymbiotic bacteria represent an evolutionary adaptation to nitrogen scarcity because less nitrogen is required for an A·T than G·C pair?

We detected deep sequence divergence (>2.0%) in nine of the 75 species that we examined (fig. 2). Our analyses revealed that specimens of a single species from varied geographic regions often showed substantial genetic divergence, possibly reflecting cryptic species overlooked by current taxonomic treatments (fig. 3). However, we also detected deep intra-specific divergences in some cases where specimens were collected in close proximity. For example, *P. aceris* included two distinctive clusters with 5.6% divergence, and these specimens were all collected on south-western islands in the Korean Peninsula (fig. 3b). Interestingly, the specimens in one

a *C. matsumotoi*

```
2008-0116| South Korea
k20080159| South Korea
2007-0229-2| South Korea
2007-0229-1| South Korea
20080153-2| South Korea
2008-0153-2| South Korea
2008-0153-1| South Korea
k20080116-1| South Korea
2008-0319| South Korea
k20080152| South Korea
k337-1| South Korea
k169-2| South Korea
k169-1| South Korea
P1064-2| Japan
P1064-1| Japan
P1018-5-1| Japan
k173-1| South Korea
P1052| Canada
2008-0119-5| South Korea
2008-0119-4| South Korea
2008-0119-3| South Korea
2008-0119-2| South Korea
2008-0119-1| South Korea
K173-01| South Korea
k20080116-2| South Korea
```

b *P. aceris*

```
K20090019-1| South Korea; Bogil-do
K20090019| South Korea; Bogil-do
K20090019-2| South Korea; Bogil-do
K20090019-3| South Korea; Bogil-do
Mb35| South Korea; Wan-do
Mb34| South Korea: Wan-do
2008-0072-1| South Korea; Wan-do
2008-0058-2| South Korea; Wan-do
2008-0058-1| South Korea; Wan-do
```

c *P. solani*

```
P1175-1| Thailand
P1175| Thailand
P1175-2| Thailand
P1175-3| Thailand
P1107-2| Italy
P1107-1| Italy
P1107| Italy
P1110-1| Japan
P1110-2| Japan
P1110| Japan
```

d *P. ficus*

```
P1201-2| South Africa
P1201-1| South Africa
P1218-3| USA
P1218-2| USA
P1218-1| USA
```

e *P. pentagona*

```
D573-1| China
D573 | China
D573-2| China
D406-2| Indonesia
D406| Indonesia
D401-2| China
```

f *L. japonica*

```
K20090025-1| South Korea
K20090025| South Korea
D568-2| Japan
```

g *A. excisus*

```
D578-3| Thailand
D578| Thailand
D629-1| Philippines
D629-2| Philippines
```
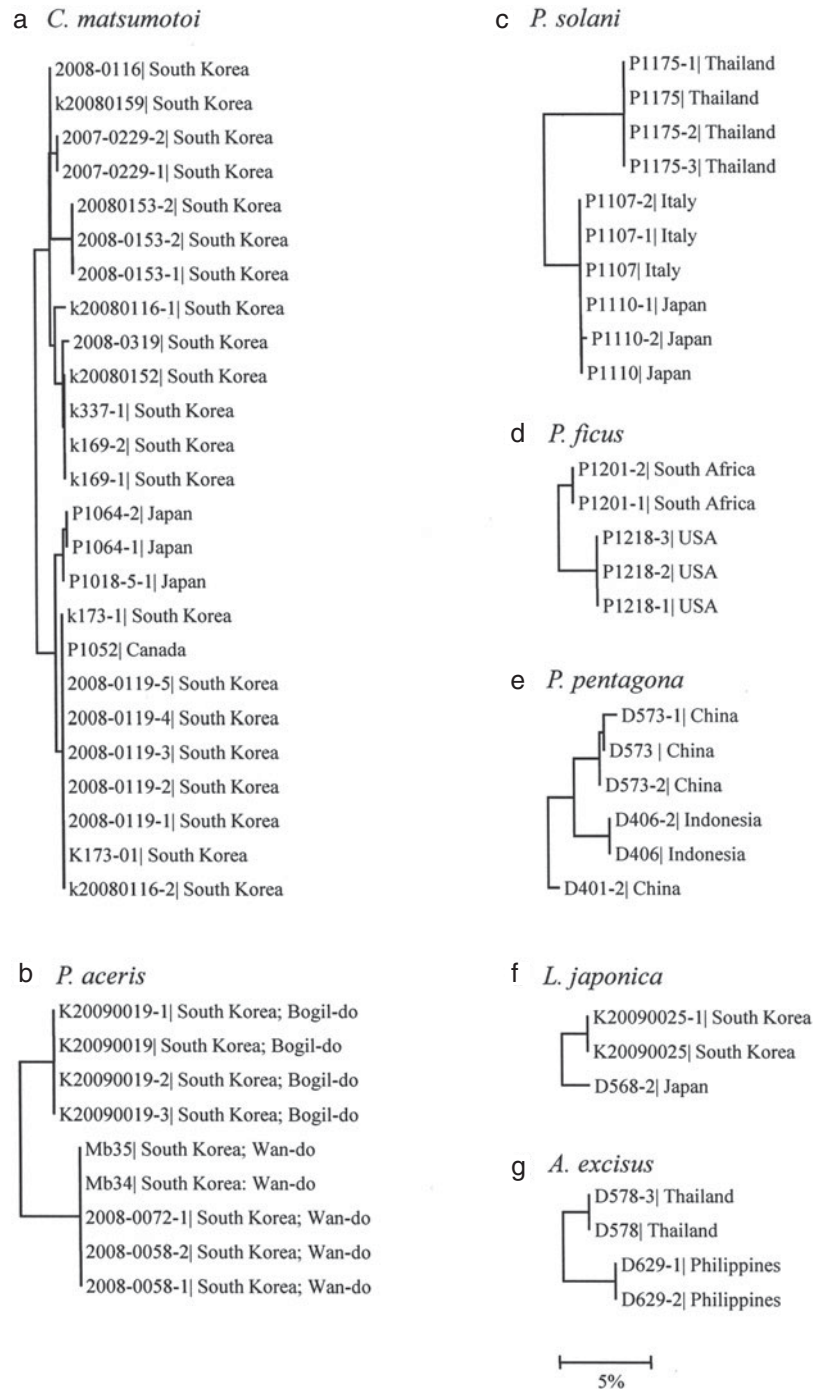
5%

Fig. 3. NJ analysis (K2P) showing intra-specific variation in the barcode region for seven species of scale insects.

cluster were collected on *Eriobotra* trees on the island of Bogil-do, while specimens in the second cluster were collected from three other tree genera (*Liquidamber, Zelkova* and *Sorbus*) on another island, Wan-do. A second case of intra-specific variation involved *C. matsumotoi*, an important pest of pears in South Korea. Specimens of this species were obtained from three countries (South Korea, Japan and Canada). The sole Canadian specimen was closely similar to some Korean specimens, while the three Japanese specimens had slightly divergent sequences (fig. 3a). However, this species included two clusters with an intra-specific K2P distance of 2.5–4.0%. We also found two *Pseudococcus* species (*comstocki* and *logispinus*) that possessed two very genetically divergent lineages that could not be separated morphologically (fig. 2).

D.-S. Park *et al.*

To clarify these potential cases of cryptic species, more extensive sampling followed by detailed phenotypic and molecular studies will be required.

This study has established the feasibility of creating a comprehensive barcode library for scale insects. It has shown that this effort will reveal taxonomic situations worthy of deeper analysis and will create an effective system for identifying species in this group, leading to a major advance in our capacity to identify immature and male scale insects. Finally, this work has revealed that the barcode region of scale insects possesses both unusual nucleotide composition and a high incidence of indels, observations that suggest the value of studies to ascertain whole mitochondrial genomes in selected scale insects.

Acknowledgements

Supplementary material

The following supplementary material can be viewed online at http://journals.cambridge.org/ber.

S1. Information of samples used in this study.
S2. DNA sequences which have thymine residue deletions. This file was generated using MEGA software and hard copied to PowerPoint.
S3. The entire NJ tree of specimens generated using the 'Sequence Analysis tool' of BOLD.

References

Andersson, S.G. & Kurland, C.G. (1998) Reductive evolution of resident genomes. *Trends in Microbiology* **6**(7), 263–268.

Baumann, L. & Baumann, P. (2005) Cospeciation between the primary endosymbionts of mealybugs and their hosts. *Current Microbiology* **50**, 84–87.

Ben-Dov, Y., Miller, D.R. & Gibson, G.A.P. (2006) ScaleNet: a database of the scale insects of the world. United States Department of Agriculture (USDA), Beltsville, MD. Available online at http://www.sel.barc.usda.gov/scalenet/scalenet.htm (accessed 9 December 2010).

Beuning, L.L., Murphy, P., Wu, E., Batchelor, T.A. & Morris, B.A. (1999) Molecular-based approach to the differentiation of mealybug (Hemiptera: Pseudococcidae) species. *Journal of Economic Entomology* **92**, 463–472.

Downie, D.A. & Gullan, P.J. (2004) Phylogenetic analysis of mealybugs (Hemiptera: Coccoidea: Pseudococcidae) based on DNA sequences from three nuclear genes, and a review of the higher classification. *Systematic Entomology* **29**, 238–259.

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.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W. & Hebert, P.D. (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences USA* **103**, 968–971.

Hebert, P.D. & Gregory, T.R. (2005) The promise of DNA barcoding for taxonomy. *Systematic Biology* **54**, 852–859.

Hebert, P.D., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003a) Biological identifications through DNA barcodes. *Proceedings of Biological Sciences* **270**, 313–321.

Hebert, P.D., Ratnasingham, S. & deWaard, J.R. (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the National Academy of Sciences USA* **270** Supplement 1, S96–99.

Ivanova, N.V., deWaard, J.R. & Hebert, P.D.N. (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes* **6**, 998–1002.

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.

Kondo, T., Gullan, P.J. & Williams, D.J. (2008) Coccidology. The study of scale insects (Hemiptera: Sternorrhyncha: Coccoidea). *Revista Corpoica - Cienciay Technologia Agropecuaria* **9**, 55–61.

Malausa, T., Fenis, A., Warot, S., Germain, J.-F., Ris, N., Prado, E., Botton, M., Vanlerberghe-Masutti, F., Sforza, R., Cruaud, C., Couloux, A. & Kreiter, P. (2010) DNA markers to disentangle complexes of cryptic taxa in mealybugs (Hemiptera: Pseudococcidae). *Journal of Applied Entomology* (doi: 10.1111/j.1439-0418.2009.01495.x).

Min, X.J. & Hickey, D.A. (2007) DNA barcodes provide a quick preview of mitochondrial genome composition. *PLoS One* **2**, e325.

Moran, N.A. & Wernegreen, J.J. (2000) Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends in Ecology & Evolution* **15**, 321–326.

Moran, N.A., McCutcheon, J.P. & Nakabachi, A. (2008) Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics* **42**, 165–190.

Morse, G.E. & Normark, B.B. (2006) A molecular phylogenetic study of armoured scale insects (Hemiptera: Diaspididae). *Systematic Entomology* **31**, 338–349.

Park, D.-S., Suh, S.-J., Oh, H.-W. & Hebert, P.D.N. (2010a) Recovery of the mitochondrial COI barcode region in diverse Hexapoda through tRNA-based primers. *BMC Genomics* **11**, 423.

Park, D.S., Leem, Y.J., Hahn, K.W., Suh, S.J., Hong, K.J. & Oh, H.W. (2010b) Molecular identification of mealybugs (Hemiptera: Pseudococcidae) found on Korean pears. *Journal of Economic Entomology* **103**, 25–33.

Perez-Brocal, V., Gil, R., Ramos, S., Lamelas, A., Postigo, M., Michelena, J.M., Silva, F.J., Moya, A. & Latorre, A. (2006) A small microbial genome: the end of a long symbiotic relationship? *Science* **314**, 312–313.

Ratnasingham, S. & Hebert, P.D.N. (2007) BOLD: The Barcode of Life Data System? (http://www.barcodinglife.org). *Molecular Ecology Notes* **7**, 355–364.

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

Thao, M.L., Gullan, P.J. & Baumann, P. (2002) Secondary (gamma-Proteobacteria) endosymbionts infect the primary (beta-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. *Applied and Environmental Microbiology* **68**, 3190–3197.