Identification of aphid (Hemiptera: Aphididae) species of economic importance in Kenya using DNA barcodes and PCR-RFLP-based approach

G. Kinyanjui[1,2], F.M. Khamis[1,2]*, S. Mohamed[1], L.O. Ombura[1], M. Warigia[2] and S. Ekesi[1]

[1]International Centre of Insect Physiology and Ecology, PO Box 30772-00100 GPO, Nairobi, Kenya: [2]Kenyatta University, PO Box 43844-00100 GPO, Nairobi, Kenya

Abstract

Aphids are among pests of economic importance throughout the world. Together with transmitting plant viruses, aphids are capable of inflicting severe crop production losses. They also excrete honeydew that favours the growth of sooty mold which reduces the quality of vegetables and fruits and hence their market values. Rapid and accurate identification of aphids to the species level is a critical component in effective pest management and plant quarantine systems. Even though morphological taxonomy has made a tremendous impact on species-level identifications, polymorphism, morphological plasticity and immature stages are among the many challenges to accurate identification. In addition, their small size, presence of cryptic species and damaged specimens dictate the need for a strategy that will ensure timely and accurate identification. In this study, polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)-based on mitochondrial cytochrome *c* oxidase subunit I gene and DNA barcoding were applied to identify different aphid species collected from different agro-ecological zones of Kenya. Three restriction enzymes *Rsa*I, *Alu*I and *Hinf1* produced patterns that allowed unambiguous identification of the species except *Aphis craccivora* and *Aphis fabae*. Analyses of the barcode region indicated intraspecific and interspecific sequence divergences of 0.08 and 6.63%, respectively. DNA barcoding identified all species, including the morphologically indistinguishable *A. craccivora* and *A. fabae* and separated two subspecies of *A. fabae*. Based on these results, both PCR-RFLPs and DNA barcoding could provide quick and accurate tools for identification of aphid species within Aphididae subsequently aiding in effective pest management programmes and enhance plant quarantine systems.

Keywords: Aphids, Aphididae, species identification, mitochondrial COI gene, PCR-RFLPs, DNA barcoding

Introduction

Globally, aphids (Hemiptera: Aphididae) are among the most economically important insect pests of crops (Emden &

Harrington, 2007). Aphids also rank high as invasive pests due to their ease of transport and parthenogenetic mode of reproduction (Foottit *et al.*, 2008). They are known to cause 70–80% of yield losses on different crops worldwide (Aslam *et al.*, 2007). These losses are due to direct feeding damage on plant sap that results in stunted growth, distortion, wilting, yellowing of plants consequently leading to severe crop production losses (Aslam *et al.*, 2007). Indirect damages results from transmission of plant viruses and their related diseases,

*Author for correspondence
Fax: +254 (20) 8632001
E-mail: fkhamis@icipe.org

plant deformation arising from toxic salivary secretions and excretion of honey dew that favours the growth of sooty mold fungus (Blackman & Eastop, 2000). Other than causing a reduction in photosynthetic activity, honey dew and sooty mold contaminate the quality of crops which in turn reduce both the aesthetic appeal and marketability of these crops (Worf *et al.*, 1995).

Considering their economic importance, timely and accurate identification of aphid species is crucial for effective pest management strategies and phytosanitary management (Miller & Foottit, 2009; Lee *et al.*, 2011). Traditionally, aphid species have been identified based on their morphological characters (Emden & Harrington, 2007). However, their small size and reduction or loss of key morphological characteristics poses a serious problem in morphological identification (Miller & Foottit, 2009). Aphids have complex lifecycles involving parthenogeneticity and high polymorphisms, and within a single species, there are different morphs with distinct morphological characters, which may colonize different host plants, further complicating species identification (Foottit *et al.*, 2008). Additionally, aphids are prone to morphological plasticity due to environmental and host plant effects making identification very difficult (Miller & Foottit, 2009). Moreover, identification of immature stages, cryptic species and damaged specimens by morphological means is problematic due to the absence of key morphological characteristics (Armstrong & Ball, 2005; Lee *et al.*, 2011). Usually, closely related species with similar morphological characteristics have been identified on the basis of their host plants association (Coeur d'acier *et al.*, 2014), which is again complicated by the polyphagous species. The recognition of these difficulties has signalled the need for novel detection methods to supplement morphology.

Several molecular techniques have been explored to assist in aphid species identification and overcome the limitations of morphological identifications. Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) is based on digestion of PCR amplicons with appropriate restriction enzymes to produce distinct polymorphic fragments, visualized as markers for species identification. Several studies have utilized PCR-RFLP for species identification of various pests of economic importance (Armstrong *et al.*, 1997; Brunner *et al.*, 2002; Masahiro *et al.*, 2008). In the family Aphididae, RFLPs have also been used to distinguish several aphid species (Raboudi *et al.*, 2002; Valenzuela *et al.*, 2007, 2009).

Increasingly, DNA barcoding is being widely employed due to its accuracy in species identification and delineation. DNA barcoding is based on the principle that a short standardized DNA sequence can characterize species in a myriad of taxonomic groups in the animal kingdom (Hebert & Gregory, 2005). The 5′ end of mitochondrial cytochrome *c* oxidase subunit I (COI) gene has been adopted as the global barcoding gene (Hebert *et al.*, 2003*a*), and has demonstrated effectiveness in identification of most pest species of economic importance (Armstrong & Ball, 2005; Ball & Armstrong, 2006; Khamis *et al.*, 2012). In the family Aphididae, not only has DNA barcoding facilitated species identification (Foottit *et al.*, 2008; Kim *et al.*, 2010; Lee *et al.*, 2011; Wang *et al.*, 2011; Coeur d'acier *et al.*, 2014), but has also proved as a potential tool in pest management by identifying aphids at any stage of their lifecycle; adult, nymphal, larval, and the various morphs within a species; colour morphs, winged and wingless morphs (Foottit *et al.*, 2009). Furthermore, DNA barcoding has enabled identification of cryptic species and discovery of new aphid species (Foottit *et al.*, 2009; Rebijith *et al.*, 2012).

In tropical Africa, the melon/cotton aphid, *Aphis gossypii* Glover (reported from 40 countries in Africa) is one of the most important pests of vegetables in the continent. For example in Kenya, yield losses due to insect pest problems on okra in Nguruman and Muhaka (for which *A. gossypii* is regarded as among the key pests) was estimated at 24–40 and 15–24%, respectively (Sithanantham *et al.*, 1998). *A. gossypii* can transmit more than 50 plant viruses causing symptoms that impair vegetable quality and yield but the true impact on crop losses have not been quantified. Two less polyphagous aphids, the cabbage aphid *Brevycorine brassicae* (L.) – restricted to mid- and high-altitude agro-ecologies, and the turnip aphid *Lipaphis pseudobrassicae* Davis – restricted to lowland agro-ecologies, are found largely on cruciferous plants of which cabbage and kale are the most affected (Nyambo & Löhr, 2005; Sæthre *et al.*, 2011). The cabbage aphid is most widely distributed (22 countries in Africa) and is known to transmit 23 plant viruses, of which *Cauliflower mosaic virus* (CaMV) and *Turnip mosaic virus* (TuMV) are known to occur in tropical Africa and can cause substantial reduction in cabbage production (Spence *et al.*, 2007).

One of the most important pests of cowpea is *Aphis craccivora* Koch. Infestation by this aphid species can result in yield losses of up to 35% (Singh & Allen, 1980) and in extreme cases in complete crop failure (Ansari *et al.*, 1992). Apart from direct feeding damage, *A. craccivora* also vectors the cowpea aphid-borne mosaic virus (Bock & Conti, 1974). The black bean aphid, *Aphis fabae* Scopoli is a significant insect pest of common beans in tropical Africa and particularly in the higher altitudes regions (Karel & Autrique, 1989). *A. fabae* is very polyphagous, causing direct physical damages on a wide range of agricultural crops and an important vector of over 30 plant viruses including the bean common mosaic virus (Emden & Harrington, 2007).

No molecular identification tools have been developed to separate these aphid species in Kenya and Africa at large. The current study was undertaken to identify a rapid tool for separation of aphid species collected from different agro-ecological zones of Kenya. Emphasis was placed on the aphid species listed above but additional aphid species including *Myzus persicae* Sulzer and *Acyrthosiphon pisum* Harris which are equally damaging to various crops were also added to the investigation. The overall goal is to avail tools that will contribute to timely and accurate identification of these pests, which should in turn facilitate quicker and effective implementation of pest management and strengthening of the quarantine systems in countries affected by the aphid species.

Materials and methods

Sample collection and processing

Adult aphids and late instar nymphs were collected from various agro-ecological zones of Kenya (fig. 1), preserved in 95% ethanol and brought to the laboratory for processing. Aphids were collected from the leaves of their host plants using a camel brush. Twenty samples from each site were randomly selected, identified and photographed dorsally, laterally and ventrally at x25 with a Leica LAS EZ4D stereo microscope with an integral digital camera (Leica Microsystems Limited, Switzerland) before DNA extraction. Voucher specimens from each collection are deposited at Arthropod pathology unit molecular laboratory in the International Centre of Insect Physiology and Ecology (*icipe*).
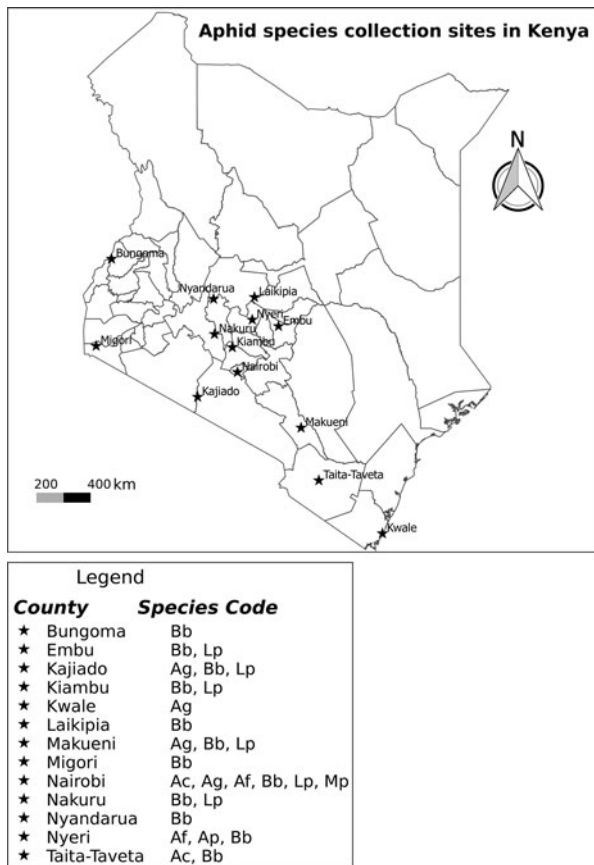
Fig. 1. Map of Kenya showing the sampling sites for aphid species between February and September 2013. Ac, *Aphis craccivora*; Af, *Aphis fabae*; Ag, *Aphis gossypii*, Ap, *Acyrthosiphon pisum*; Bb, *Brevicoryne brassicae*; Lp, *Lipaphis pseudobrassicae* and Mp, *Myzus persicae*.

DNA extraction

Each individual photographed sample was surface-sterilized using 3% bleach and rinsed three times with distilled water. Genomic DNA was extracted using proteinase K buffer extraction protocol. Each aphid sample was put into a sterile 1.5 ml Eppendorf tube and 100 µl 1 × proteinase K buffer (10 × proteinase K buffer composed of 25 mM KCl, 10 Mm Tris-HCl at pH 9.0 at 25°C and 10 mM Triton X-100) and 0.5 µl of 20 mg ml^{-1} proteinase K (Thermo Scientific, USA) were added. The sample was homogenized using a sterile pestle and incubated overnight at 37°C in a water bath. The homogenate was then heated at 95°C in a water bath for 5 min to inactivate proteinase K. The sample was centrifuged at 15,000 *g* for 5 min and 40 µl of supernatant was aliquoted into a sterile Eppendorf tube. The extracted DNA was then stored at −20°C for downstream processes.

PCR-RFLP

A 1540 bp region of mitochondrial COI gene was amplified by PCR using the primers A-pissum Fwd TCAACTAATCAT AAAGATATTGGAA and A-pissum Rv TATAAATGAATT TTAAGTTC. The primers were generated from a sequence that was retrieved from a complete genome of mitochondrion of *Acyrthosiphon pisum* Harris in the National Center for Biotechnology Information (NCBI) database (FJ411411.1), using sequence manipulation suite software and were tested *in silico* for their ability to amplify the different species under study. The primers were synthesized at a commercial facility (Macrogen Inc., Europe). PCR was carried out in a 20 µl volume containing 5 × My *Taq* Reaction Buffer (Bioline, London, UK) composed of 5 mM dNTPs, 15 mM MgCl$_2$ stabilizers and enhancers, 10 µM of each primer, 1.25 mM MgCl$_2$, 6.25 units My *Taq* DNA polymerase (Bioline, London, UK) and 15 ng µl^{-1} of DNA template. PCR thermo cycling conditions of 1 min at 95°C, followed by 35 cycles of 15 s at 95°C, 1 min at annealing temperature of 49.1°C and 1 min at 72°C, then a final elongation step of 10 min at 72°C were used. The reaction was set up in Arktik thermal cycler (Thermo Fisher Scientific Inc., USA). Amplified products were analyzed by electrophoresis on 1% agarose gel with ethidium bromide and O' Gene Ruler 1 kb Plus DNA ladder (Thermo Fisher Scientific Inc., USA) incorporated. The program NEB cutter V2.0 (New England BioLabs Inc., MA, USA) (http://tools.neb.com/ NEBcutter2/index.php) was used to predict the potential restriction sites. Three restriction enzymes; *Rsa*I, *Alu*I and *Hinf*I were selected based on the size of the resulting fragments and the ability to distinguish the species studied. Restriction digest was done in 18 µl volume containing 10 µl nuclease-free water, 2 µl fast digest buffer (Thermo Fisher Scientific Inc., USA), 1 µl restriction enzyme and 5 µl PCR product. PCR products were digested separately and incubation conditions of 37°C for 14 h, enzyme inactivation at 65°C for 5 min and a hold temperature of 10°C were used. Restriction products were resolved through a 2% agarose gel with O' Gene Ruler 1 kb Plus and 100 bp DNA ladders (Thermo Fisher Scientific Inc., USA). Electrophoresis was set at 70 V for 90 min, followed by visualization of DNA under ultraviolet (UV) transilluminator in a KETA gel documentation imaging system (Wealtec Corp., Nevada, USA).

DNA barcode region amplification and sequencing

PCR was carried out using universal primers, LCO 1490 5′-GGTCAACAAATCATAAAGATATTGG-3′ and HCO 2198 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′ (Folmer *et al.*, 1994) to amplify an approximately 680 bp fragment of the COI gene. Amplification was carried out in a 20 µl volume containing 5 × My *Taq* Reaction Buffer (Bioline, London, UK) composed of 5 mM dNTPs, 15 mM MgCl$_2$ stabilizers and enhancers, 10 µM of each primer, 1.25 mM MgCl$_2$, 6.25 units My *Taq* DNA polymerase (Bioline, London, UK) and 15 ng µl^{-1} of DNA template. PCR thermo cycling conditions of 1 min at 95°C, then 35 cycles of 15 s at 95°C, 1 min at annealing temperature of 48.3°C and 1 min at 72°C, followed by a final elongation step of 10 min at 72°C were used. The reaction was set up in Arktik thermal cycler (Thermo Fisher Scientific Inc., USA). Amplified PCR products were resolved through a 1% agarose gel. Electrophoresis was set at 100 volts for 1 h, followed by visualization of DNA under UV-illumination in a KETA gel documentation imaging system (Wealtec Corp., Nevada, USA). PCR products were purified using Isolate II PCR and Gel Kit (Bioline, London, UK) according to the manufacturer's instructions. Purified DNA samples were bi-directionally sequenced using ABI 3730xl DNA sequencer (Applied Biosystems, Foster City, California) at a commercial sequencing facility (Macrogen Inc., Europe).

G. Kinyanjui *et al.*

Table 1. Collection data of aphid samples collected from different agro-ecological zones of Kenya with species name, host plant, global positioning system (GPS) coordinates of the sampling locations and GenBank accession numbers.

Species name	Sample name	Host plant	Coordinates	GenBank accession numbers
Aphis gossypii	Ag1, Ag2, Ag3, Ag4, Ag5	Okra	S4°22′32.94″ E39°31′57.29″	KR085019, KR085018, KR085017, KR085016, KR085015
Aphis gossypii	Ag6, Ag7, Ag8, Ag9, Ag10	Okra	S2°23′33.1″ E37°59′49.1″	KR085020, KR085021, KR085022, KR085023, KR085024
Aphis gossypii	Ag11, Ag12, Ag13, Ag14, Ag15	Okra	S1°48′22.1″ E36°03′41.2″	KR085025, KR085026, KR085027, KR085028, KR085029
Aphis gossypii	Ag16, Ag17, Ag18, Ag19	Okra	S1°13′14.59″ E36°53′43.73″	KR085030, KR085031, KR085032, KR085033
Brevicoryne brassicae	Bb1, Bb2, Bb3, Bb4, Bb5	Kales	N0°02′15.98″ E36°21′40.59″	KR085076, KR085075, KR085034, KR085074, KR085073
Brevicoryne brassicae	Bb6, Bb7, Bb8, Bb9, Bb10	Kales	S0°30′42.09″ E36°19′21.30″	KR085072, KR085071, KR085070, KR085069, KR085068
Brevicoryne brassicae	Bb11, Bb12, Bb13, Bb14, Bb15	Cabbage	S0°37′27.19″ E36°22′42.60″	KR085067, KR085066, KR085065, KR085064, KR085063
Brevicoryne brassicae	Bb16, Bb17, Bb18, Bb19, Bb20	Kales	S0°17′06.11″ E36°04′52.08″	KR085062, KR085061, KR085060, KR085059, KR085058
Brevicoryne brassicae	Bb21, Bb22, Bb23, Bb24, Bb25	Kales	N0°00′01.78″ E36°13′51.88″	KR085057, KR085056, KR085055, KR085054, KR085053
Brevicoryne brassicae	Bb26, Bb27, Bb28, Bb29, Bb30	Kales	S0°50′46.05″ E34°09′20.29″	KR085052, KR085051, KR085050, KR085049, KR085048
Brevicoryne brassicae	Bb31, Bb32, Bb33, Bb34, Bb35	Kales	S1°09′03.16″ E36°57′33.75″	KR085047, KR085046, KR085045, KR085044, KR085043
Brevicoryne brassicae	Bb36, Bb37, Bb38, Bb39, Bb40	Cabbage	S1°04′37.98″ E36°37′12.39″	KR085042, KR085041, KR085040, KR085039, KR085038
Brevicoryne brassicae	Bb41, Bb42, Bb43, Bb44, Bb45	Cabbage	S3°22′41.20″ E38°20′14.86″	KR085037, KR085036, KR085035, KR085122, KR085121
Brevicoryne brassicae	Bb46, Bb47, Bb48, Bb49	Kales	S1°48′22.1″ E36°03′41.2″	KR085120, KR085119, KR085118, KR085117
Brevicoryne brassicae	Bb50, Bb80, Bb81, Bb86, Bb87	Kales	S1°26′13.30″ E36°41′06.06″	KR085116, KR085086, KR085085, KR085080, KR085079
Brevicoryne brassicae	Bb51, Bb52, Bb53, Bb54, Bb66	Kales	S0°04′20.29″ E37°07′42.99″	KR085115, KR085114, KR085113, KR085112, KR085100
Brevicoryne brassicae	Bb61, Bb62, Bb63, Bb64, Bb65	Kales	N0°47′30.91″ E34°26′40.42″	KR085105, KR085104, KR085103, KR085102, KR085101
Brevicoryne brassicae	Bb67, Bb68, Bb69, Bb70, Bb71	Kales	S0°28′36.10″ E37°34′59.84″	KR085099, KR085098, KR085097, KR085096, KR085095
Brevicoryne brassicae	Bb55, Bb72, Bb73, Bb74, Bb82	Kales	S1°13′14.59″ E36°53′43.73″	KR085111, KR085094, KR085093, KR085092, KR085084
Brevicoryne brassicae	Bb56, Bb57, Bb75, Bb76, Bb77	Kales	S2°23′11.20″ E38°00′14.20″	KR085110, KR085109, KR085091, KR085090, KR085089
Brevicoryne brassicae	Bb58, Bb59, Bb60, Bb78, Bb79	Kales	S0°21′10.69″ E37°5′14.35″	KR085108, KR085107, KR085106, KR085088, KR085087
Brevicoryne brassicae	Bb83, Bb84, Bb85, Bb88, Bb89	Kales	S1°12′56.82″ E36°53′47.73″	KR085083, KR085082, KR085081, KR085078, KR085077
Lipaphis pseudobrassicae	Lp1, Lp2, Lp3, Lp4, Lp5	Kales	S1°10′12.85″ E36°54′09.88″	KR085144, KR085143, KR085142, KR085141, KR085140
Lipaphis pseudobrassicae	Lp6, Lp7, Lp8, Lp9, Lp10	Kales	S1°39′41.32″ E37°26′56.78″	KR085139, KR085138, KR085137, KR085136, KR085135
Lipaphis pseudobrassicae	Lp11, Lp12, Lp13, Lp14, Lp15	Kales	S0°37′27.19″ E36°22′42.60″	KR085134, KR085133, KR085132, KR085131, KR085130
Lipaphis pseudobrassicae	Lp16, Lp17, Lp18,	Kales	S1°09′03.16″ E36°57′33.75″	KR085129, KR085128, KR085127
Lipaphis pseudobrassicae	Lp19, Lp20, Lp21, Lp22, Lp23	Kales	S0°28′27.95″ E37°34′49.23″	KR085126, KR085125, KR085124, KR085152, KR085153
Lipaphis pseudobrassicae	Lp24, Lp25, Lp26, Lp27, Lp28	Kales	S1°13′14.59″ E36°53′43.73″	KR085154, KR085155, KR085156, KR085157, KR085158
Lipaphis pseudobrassicae	Lp29, Lp30, Lp31, Lp32, Lp33	Kales	S1°12′56.82″ E36°53′47.73″	KR085159, KR085148, KR085149, KR085150, KR085151
Lipaphis pseudobrassicae	Lp34, Lp35, Lp36, Lp37,	Kales	S1°48′22.1″ E36°03′41.2″	KR085123, KR085145, KR085146, KR085147
Aphis craccivora	Ac1, Ac2, Ac3, Ac4, Ac5	Cowpeas	S3°16′8.3″ E37°44′17.7″	KR084997, KR085001, KR085002, KR085003, KR085004
Aphis craccivora	Ac6, Ac7, Ac8, Ac9, Ac10	Cowpeas	S1°13′25.5″ E36°53′50.5″	KR084999, KR084998, KR085000, KR084996, KR084995
Aphis fabae	Af1, Af2, Af3, Af4, Af5	Rose coco beans	S0°21′10.69″ E37°5′14.35″	KR085014, KR085013, KR085012, KR085011, KR085010
Aphis fabae	Af6, Af7, Af8, Af9, Af10	Black night shade	S1°13′14.59″ E36°53′43.73″	KR085009, KR085008, KR085007, KR085006, KR085005
Myzus persicae	Mp1, Mp2, Mp3, Mp4, Mp5	Cabbage	S1°13′14.59″ E36°53′43.73″	KR085160, KR085161, KR085162, KR085164, KR085163
Acyrthosiphon pisum	Ap1, Ap2, Ap3, Ap4, Ap5	Garden peas	S0°21′11.30″ E37°5′20.18″	KR084994, KR084993, KR084992, KR084991, KR084990

Samples were collected from February to September 2013. Sample name corresponds to the name of the species as used during the experiments and analysis.
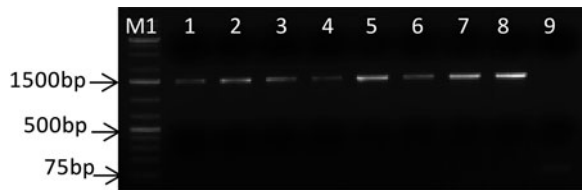
Fig. 2. PCR products of 1500 bp fragment amplified from genomic DNA of aphid species on a 1% agarose gel. M1: O' Gene Ruler 1 kb Plus DNA ladder, Lane 1: *Aphis gossypii*, Lane 2: *Myzus persicae*, Lane 3: *Lipaphis pseudobrassicae*, Lane 4: *Brevicoryne brassicae*, Lane 5: *Acyrthosiphon pisum*, Lane 6: *Aphis craccivora*, Lane 7: *Aphis fabae*, Lane 8: *Aphis fabae* and Lane 9: Negative control.
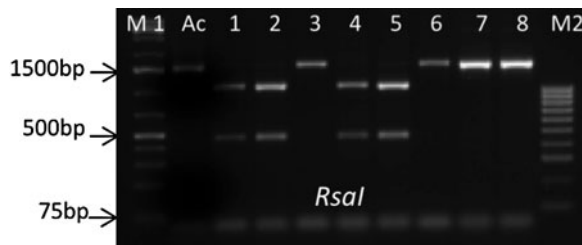

Fig. 3. PCR-RFLP analysis of aphid species on a 2% agarose gel. DNA was digested with *Rsa*I. M1: O' Gene Ruler 1 kb Plus DNA ladder, Lane 1: *Aphis gossypii*, Lane 2: *Myzus persicae*, Lane 3: *Lipaphis pseudobrassicae*, Lane 4: *Brevicoryne brassicae*, Lane 5: *Acyrthosiphon pisum*, Lane 6: *Aphis craccivora*, Lane 7: *Aphis fabae*, Lane 8: *Aphis fabae* and M2: O' Gene Ruler 100 bp DNA ladder.

Fig. 4. PCR-RFLP analysis of aphid species on a 2% agarose gel. DNA was digested with *Alu*I. M1: O' Gene Ruler 1 kb Plus DNA ladder, Lane 1: *Aphis gossypii*, Lane 2: *Myzus persicae*, Lane 3: *Lipaphis pseudobrassicae* Lane 4: *Brevicoryne brassicae*, Lane 5: *Acyrthosiphon pisum*, Lane 6: *Aphis craccivora*, Lane 7: *Aphis fabae*, Lane 8: *Aphis fabae* and M2: O' Gene Ruler 100 bp DNA ladder.

Fig. 5. PCR-RFLP analysis of aphid species on a 2% agarose gel. DNA was digested with *Hinf*I. M1: O' Gene Ruler 1 kb Plus DNA ladder, Lane 1: *Aphis gossypii*, Lane 2: *Myzus persicae*, Lane 3: *Lipaphis pseudobrassicae* Lane 4: *Brevicoryne brassicae*, Lane 5: *Acyrthosiphon pisum*, Lane 6: *Aphis craccivora*, Lane 7: *Aphis fabae*, Lane 8: *Aphis fabae* and M2: O' Gene Ruler 100 bp DNA ladder.

Data analysis

Sequences were assembled and edited using Chromas v2.1.1 (Technelysium Pty Ltd, Queensland, Australia). For conclusive identifications, sequences were queried via basic local alignment search tool (BLAST) at the GenBank database hosted by NCBI (http://www.ncbi.nlm.nih.gov/). Megablast (for highly similar sequences) program was used for all the sequences. Alignment was done using MUSCLE v3.8.31 (Edgar, 2004). Unaligned sequence ends were trimmed and gaps removed in Jalview v2.8.2 (Waterhouse *et al*., 2009). The program jModeltest v2.1.7 (Darriba *et al*., 2012) was used to determine the appropriate substitution model for phylogenetic analyses. The best-fit model TIM1 + G was selected by the four different criteria (AIC, AICc, BIC and DT). Maximum likelihood (ML) estimates were obtained using the TIM1 + G model under a general time reversible GTRGAMMA substitution model with 1000 bootstrap replicates in RAxML v8.2.0 (Stamatakis, 2014). Generated trees were viewed and edited in Fig Tree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree). Genetic divergences were determined at the species, genus and family levels using pairwise distance model as generated by the distance summary tool available in Barcode of Life Data systems (BOLD) (http://www.boldsystems.org/). Evolutionary divergence over sequence pairs between groups were estimated using the p-distance model in MEGA 6.0 (Tamura *et al*., 2013). To further infer relationships among the aphid species, principal component analysis (PCA) was conducted. A table of genetic distances generated by MEGA 6.0 (Tamura *et al*.,2013) was used to create principal component plots using GenAlEx 6.41 (Peakall & Smouse, 2006). The program DnaSP 5.0 (Librado &
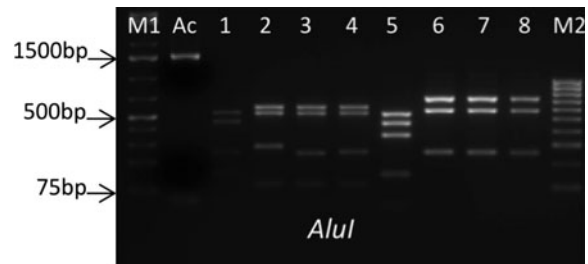
Rozas, 2009) was used to analyze DNA polymorphism in the nucleotide sequences within and between populations. In this program, DNA sequence variation generated a haplotype file which was used to construct a phylogenetic network using the haplotype median-joining algorithm in Network 4.6.1.1 (Fluxus Technology Ltd, Suffolk, England). Finally, COI sequences were submitted to the BOLD database and deposited in GenBank (Accession numbers are found in table 1).

Results

PCR-RFLP

In silico restrictions predicted a specific profile for each species except *A craccivora* and *A. fabae*. These predictions were checked on several aphid species collected from different agro-ecological zones of Kenya. Amplified 1540 bp fragment (fig. 2) was subjected to restriction digests and the products ranged from 1000 to less than 75 bp. Fragments less than 100 bp were not considered as diagnostic. Restriction profiles exhibited by these enzymes corresponded to *in silico* predictions. *Rsa*I yielded fragments of 500 and 1000 bp in *A. gossypii*, *M. persicae*, *B. brassicae* and *A. pisum* and therefore, failed to differentiate the four species. *Rsa*I also ruled out *L. pseudobrassicae*, *A. craccivora* and *A. fabae* (fig. 3).

*Alu*I digestion profile gave informative results for *A. gossypii* and *A. pisum* and allowed accurate discrimination of the two species (fig. 4). *Alu*I yielded fragments of 150, 250, 450 and 550 bp in *A. gossypii* whereas fragments in *A. pisum*

Table 2. Genetic sequence divergence of mitochondrial COI gene region between different taxonomic levels of Aphididae with minimum (min), maximum (max), mean and standard error (SE) distance values as determined using pairwise distance model.

	n	Taxa	Comparisons	Min distance (%)	Mean distance (%)	Max distance (%)	SE distance (%)
Within species	175	7	4863	0	0.08	1.04	0
Within genus	39	1	480	5.67	6.63	7.91	0
Within family	175	1	9882	5.01	6.90	9.89	0

Table 3. Estimates of evolutionary divergence of mitochondrial COI gene region over sequence pairs between groups as determined using p-distance model in Mega 6.0 (Tamura *et al.*, 2013).

	Ag	Ac	Af	Bb	Lp	Ap	Mp
Ag	0						
Ac	0.077	0					
Af	0.061	0.062	0				
Bb	0.075	0.086	0.077	0			
Lp	0.074	0.088	0.085	**0.052**	0		
Ap	0.089	0.078	0.085	0.060	0.071	0	
Mp	0.089	**0.098**	0.088	0.083	0.085	0.089	0

Ag, *Aphis gossypii*; Ac, *Aphis craccivora*; Af, *Aphis fabae*; Bb, *Brevicoryne brassicae*; Lp, *Lipaphis pseudobrassicae*; Ap, *Acyrthosiphon pisum* and Mp, *Myzus persicae*. Numbers in bold indicate the lowest (Lp and Bb) and the highest (Mp and Ac) values between the species.

were 150, 350, 400 and 500 bp. *Alu*I also provided a clear distinction between *M. persicae* and *L. pseudobrassicae* and *B. brassicae*. *Alu*I yielded fragments of 150, 300, 500 and 600 bp in *M. persicae* and fragments of 150, 250, 500 and 600 bp in *L. pseudobrassicae* and *B. brassicae*. Finally, *Alu*I digestion yielded fragments of 250, 500 and 700 bp in *A. craccivora* and *A. fabae* and could not therefore discriminate between the two species.

Digestion with *Hinf*I allowed effective differentiation of five of the species studied (fig. 5). *Hinf*I digestion of *A. gossypii* resulted in fragments of 300, 450 and 600 bp, while those of *M. persicae* were 100, 200, 300 and 1000 bp. *Hinf*I fragments of *L. pseudobrassicae* were 250, 400 and 500 bp, whereas those of *B. brassicae* were 300, 400 and 1250 bp. *Hinf*I digestion of *A. pisum* yielded fragments of 200, 500 and 700 bp. However, it was quite a challenge to distinguish between the cowpea specialist *A. craccivora* and the bean specialist *A. fabae* because they yielded fragments of equal sizes (250, 400 and 1000 bp) upon digestion with *Hinf*I enzyme.

Barcode data analyses

COI sequences were generated for 175 aphid samples collected across Kenya. BLAST hits between our sequence dataset and those from the NCBI Genbank database ranged from 95 to 100%. Genetic sequence divergences were analyzed at the species, genus and family level, and as expected genetic divergences increased with higher taxonomic levels. Sequence divergence within species ranged from 0 to 1.04% with an average of 0.08% (table 2) and all species displayed intraspecific genetic distances of less than 2%. Mean sequence divergence between species of the same genus was 6.63%, with a range of 5.67–7.91% and comparison of all species pair revealed that all species showed genetic distances greater than 2%. The sequence divergence between different genera of the same family ranged from 5.01 to 9.89% with a mean of 6.90%. Less

interspecific divergences were found between *L. pseudobrassicae* and *B. brassicae* (0.052) whereas *M. persicae* and *A. craccivora* were the most genetically divergent species (0.098) (table 3). Mean nucleotide frequencies were $T = 0.4377$, $A = 0.3440$, $C = 0.1149$ and $G = 0.1035$, with a bias towards thymine and adenine composition and no stop codons were observed within the sequences. Negative log likelihood was 1751.1144 and ML estimate of the gamma shape parameter was 0.0820.

Phylogenetic analyses

The phylogenetic tree was generated from 175 samples belonging to different species and populations. The tree separated into two major groups that were supported by high bootstrap values of 92% (fig. 6). The first group had the clustering of *Aphis* species, in which *A. gossypii* formed a sister group with strong support of 92% bootstrap value with the cluster containing the two clades of *A. craccivora* and *A. fabae*. The second group had the clustering of four species with robust support of 100% bootstrap values. *L. pseudobrassicae* and *B. brassicae* branched from the same node, whereas *M. persicae* formed a clade with *A. pisum*. Aphid samples of *B. brassicae* were collected from 18 different regions, some from cabbages and others from kales (Brassicaceae) (table 1). It was noted that these populations clustered together irrespective of their host plants and geographic origins. This clustering was also evident in *L. pseudobrassicae*, *A. gossypii* and *A. craccivora* and could suggest the existence of the same populations within a species. However, samples of *A. fabae* were clustered into two clades based on their host plants.

The PCA clearly separated the seven species into distinct clusters (fig. 7). The cluster belonging to *B. brassicae* was shown to be much closer to the cluster consisting of *L. pseudobrassicae* than it was to the cluster consisting of *A. pisum*. Consistently, the PCA results revealed a close relationship between *A. fabae*, *A. craccivora* and *A. gossypii*. Haplotype network patterns further confirmed the previous results by positioning the species into distinct nodes (fig. 8). *Aphis craccivora* and *A. fabae* diverged from the same median vector, suggesting a close relationship between these species. Moreover, *A. craccivora*, *A. fabae* and *A. gossypii* were found to be related as they all split from a common median vector. The two sub groups of *A. fabae* were distinctly clustered, confirming the results obtained from the phylogenetic tree. Also, a close relationship was confirmed between *B. brassicae* and *L. pseudobrassicae*, whose median vectors split from a common median vector. Finally, it was observed that *A. pisum* and *M. persicae* formed separate clusters which were derived from the central median vectors independently.

Discussion

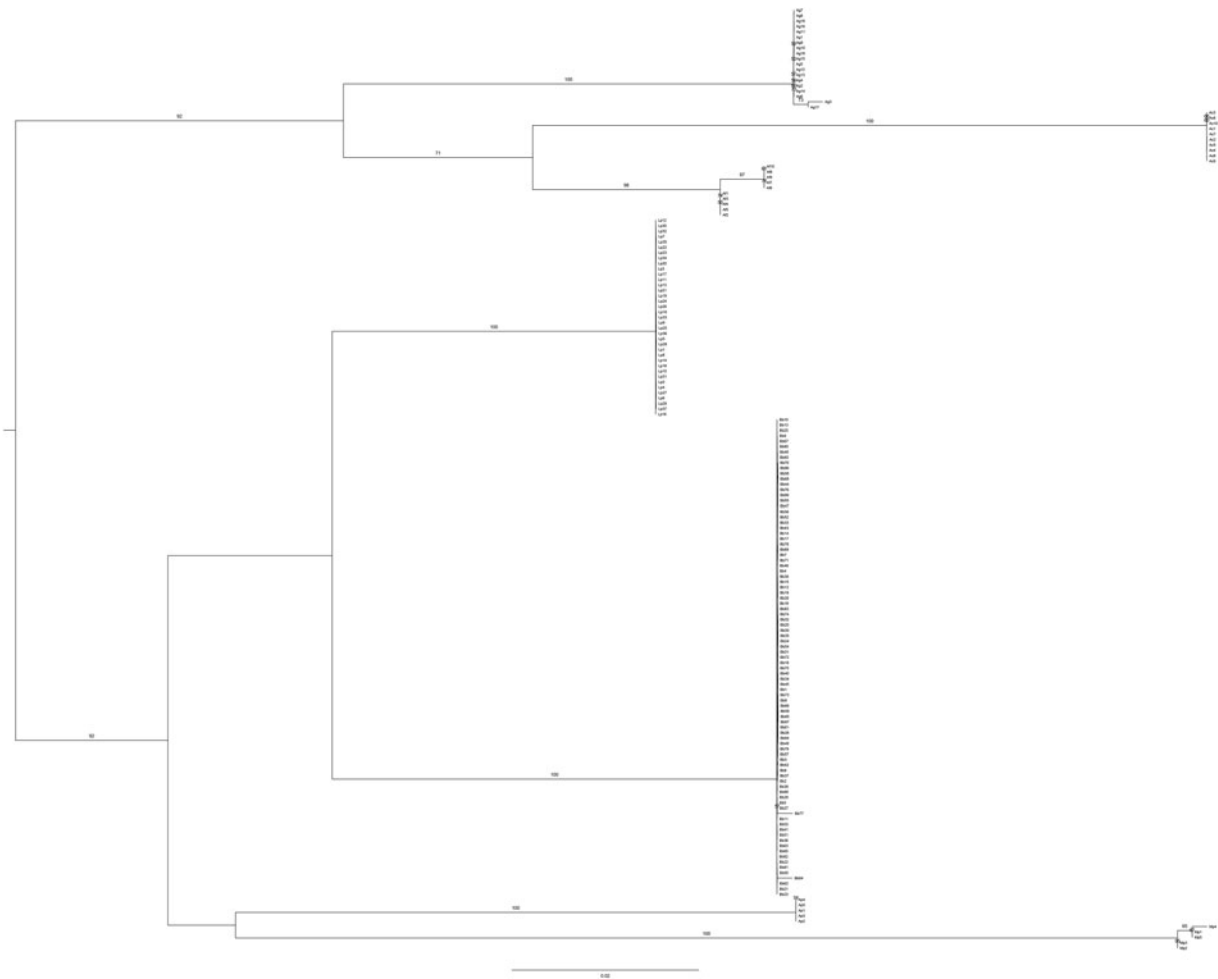In this study, RFLPs were first developed from the COI region to detect sequence variation among seven aphid species

Fig. 6. Maximum likelihood tree inferred from COI sequences of 175 samples using TIM1 + G model of molecular evolution (-lnL = 1751.1144 and gamma shape = 0.0820) under GTRGAMMA model for seven aphid species. Support values designated at the nodes (next to the branches) represent percentage bootstrap values after 1000 replications. Bootstrap values less than 50% are not indicated. The tree was drawn to scale, and the branch lengths are denoted by the rate of substitution per nucleotide position.
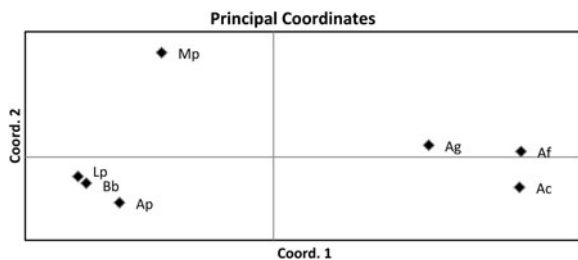

Fig. 7. Plots of the principal component analysis (PCA) for seven aphid species calculated using GenAlEx. The first and second principal components were 25.91 and 21.28, respectively.

and based on the results *B. brassicae, L. pseudobrassicae, A. pisum, M. persicae* and *A. gossypii* were distinguished successfully. A single restriction enzyme was not sufficient to distinguish the species and therefore, three restriction enzymes were combined to produce a profile that could serve as a molecular diagnostic key for species identification. A similar study by Valenzuela *et al.* (2007) demonstrated the utility of restriction analysis of PCR-amplified COI region for the identification of immature stages of 25 aphid species. Similarly, the technique has been used to separate four *Rhopalosiphum* Koch species (Yeh *et al.*, 2005) and differentiate between aphid haplotypes (Raboudi *et al.*, 2002; Shufran, 2003). PCR-RFLP provides a simple, fast and cost effective diagnostic tool for aphid species identification. It is a relatively sound technique that could be beneficial in low budget situations. It is also applicable in situations where there may be no access to sequence analysis software and skills to analyze sequence data. However, RFLPs depends on a few informative DNA sequence positions and thus, only a fraction of sequence variations is detected (Brunner *et al.*, 2002). For example, *A. craccivora* and *A. fabae* could not be distinguished and the two subgroups of *A. fabae* could not be separated. In addition, RFLPs depends on the selection of specific restriction enzymes that require previous knowledge of the samples to be analyzed.

Legend
H_1: *Aphis gossypii*
H_2: *Aphis craccivora*
H_3: *Aphis fabae*
H_4: *Aphis fabae*
H_5: *Brevicoryne brassicae*
H_6: *Lipaphis pseudobrassicae*
H_7: *Acyrthosiphon pisum*
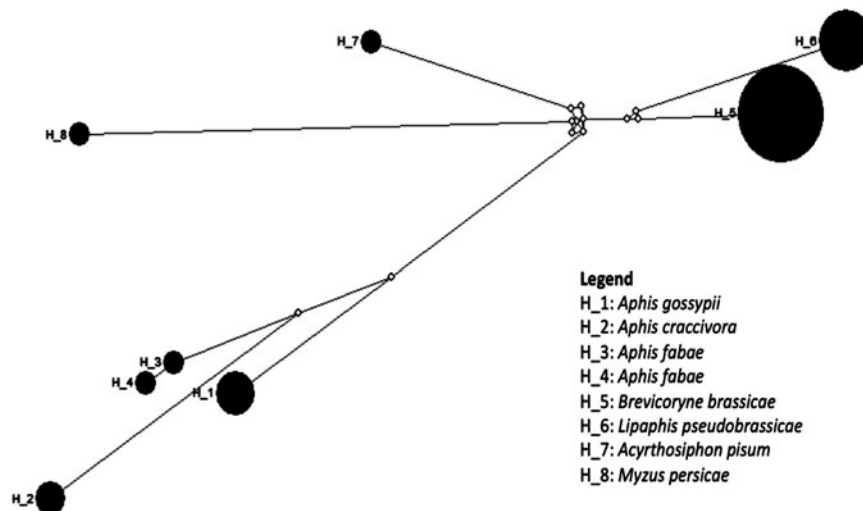H_8: *Myzus persicae*

Fig. 8. Phylogenetic network showing evolutionary relationships among seven aphid species. Eight distinct haplotypes were detected with a haplotype diversity, Hd = 0.6822. All sites containing gaps and missing data were excluded and the number of variable sites was 122. There were a total of 650 sites in the final dataset. The black full nodes represent the sampled sequences and the size of each node is proportional to the corresponding haplotype (sequence) frequency. The small empty nodes represent the median vectors. The size of the branch is proportional to the number of mutations that occurred.

In DNA barcoding, sequences were retrieved from 175 aphid samples collected from different agro-ecological zones of Kenya. COI sequences gave a high resolution of all species and two subgroups of *A. fabae*. These results are consistent with previous studies carried out in Europe, North America, Korea and India that demonstrated DNA barcoding based on COI gene as a potentially useful tool for identification of species within Aphididae (Foottit *et al.*, 2008; Lee *et al.*, 2011; Rebijith *et al.*, 2013; Coeur d'acier *et al.*, 2014). Also, by providing quantitative data with sequence divergence and bootstrap values as a measure of reliability (Armstrong & Ball, 2005), DNA barcoding presented a more accurate and robust approach towards separation of the aphid species and defining their identity as compared with PCR-RFLPs. Virgilio *et al.* (2010) reported that intraspecific sequence divergences of DNA barcodes belonging to 1995 insect species ranged from 0.0 to 7.64% and aphids were reported to lie at the lower ranges. In our study, the mean intraspecific sequence divergence obtained was as low as 0.08% and is therefore within the ranges reported by previous authors. COI sequence divergences among species can vary among different groups of animals ranging from 0.0 to 53.7% among 13,320 species pairs (Hebert *et al.*, 2003*b*). In our study, the mean interspecific sequence divergence was 6.63%, and significantly greater than the mean intraspecific divergence. This created a substantial barcode gap that enabled accurate discrimination of all the species. Our results are within the expected limits and fall within the range of variation exhibited in previous studies within Aphididae (Foottit *et al.*, 2008, 2009; Lee *et al.*, 2011; Wang *et al.*, 2011) and demonstrate that COI DNA barcode region is a reliable marker for separating aphid species. Clustering of species on the phylogenetic tree was in accordance with the results obtained from the PCA plot and phylogenetic network. Moreover, the close relationship between *B. brassicae* and *L. pseudobrassicae* and between *A. gossypii*, *A. craccivora* and *A. fabae* was confirmed.

Clades belonging to *B. brassicae*, *L. pseudobrassicae*, *A. craccivora* and *A. gossypii* consisted of populations collected from various agro-ecological zones of the country. COI

sequences revealed that no genetic variation existed between populations belonging to the same species, despite different geographical locations as they were all grouped together within a cluster. Genetic uniformity was also observed among populations of *B. brassicae* collected from kales and cabbages. These results concur with previous studies on aphid species in the family Lamiaceae and *Brachycaudus helichrysi* Kaltenbach, where the clustering of populations was not related to their geographic locations (Piffaretti *et al.*, 2012; Cocuzza & Cavalieri, 2014). Genetic uniformity has also been observed in the genus *Hyalopterus* Koch, where populations of each species clustered together irrespective of the widespread geographical sampling (Lozier *et al.*, 2008). Similarly, COI sequences belonging to *A. gossypii* and *M. persicae*, which are both cosmopolitan and highly polyphagous species, revealed that there was no significant genetic variation among populations collected from different geographical locations and different host plants (Rebijith *et al.*, 2012).

The great morphological similarity between *A. craccivora* and *A. fabae* makes it very difficult to separate them based on morphological characters. However, previous studies have presented DNA barcoding as a very useful tool in separation of the two species (Coeur d'acier *et al.*, 2007; Lee *et al.*, 2011). In the current study, not only were the species identified, but DNA barcoding was able to reveal that both aphids are sister species in a clade that also formed a sister group relationship with *A. gossypii*. This relationship is also evident in the PCA plot and the median joining haplotype network, where the three species formed neighbouring clusters. The study also distinctly separated two subgroups of *A. fabae* collected from black night shade (*Solanum nigrum* L.) and rose coco beans (*Phaseolus vulgaris* L.). All sequences of *A. fabae* displayed very high BLAST hits between 98 and 100% with a Genbank sequence assigned to the same species. However, these results could not separate the two subgroups and only tended to be diagnostic at the species level. It has been reported that *A. fabae* species constitutes a complex of six morphologically inseparable subspecies, whose identification is based on host plants affiliations

(Zhang *et al.*, 2010). Our results are in agreement with previous studies on *Aphis* species that distinctly clustered *A. fabae* into four highly supported subspecies (Coeur d'acier *et al.*, 2007). The results were further confirmed by the phylogenetic network, which positioned the two subgroups as distinct haplotypes with a very close relationship.

Complex life cycles of aphids associated with parthenogeneticity and polymorphism including colour morphs, winged, unwinged, sexual and asexual morphs is a common trait among aphids (Foottit *et al.*, 2008). In our study, no genetic variation was observed in the DNA sequences obtained from samples of the yellow and black morphs of *A. gossypii*. Based on these results, we conclude that morphological variations that occur among species populations in response to environmental factors and host plant effects are not reflected in their genetic makeup. The immature stages of the various aphid species must be reared to adult stages to allow for morphological identification – a process that is laborious and time consuming. The current study have shown that both RFLPs and DNA barcoding could be used reliably for the identification of immature stages, different morphs and various life stages within a species (Valenzuela *et al.*, 2007; Foottit *et al.*, 2009; Shufran & Puterka, 2011).

RFLP markers were developed for the identification of five aphid species, which are among the most damaging pest of vegetables in Kenya. However, the restriction enzymes used in the study could not separate *A. craccivora* and *A. fabae.* Our DNA barcoding results contribute to the growing database of DNA barcodes of aphid species of the world. Nevertheless, extensive sampling is needed in future studies to generate a more comprehensive barcode database for aphids attacking crops in Kenya. Additional work is also necessary to investigate the occurrence of *A. fabae* subspecies on the basis of their host plants. Overall, the study has presented both PCR-RFLP and DNA barcoding as diagnostic tools to distinguish among Aphididae species. With the availability of quick and accurate identification tools, monitoring and detection of potentially invasive species could be heightened, facilitating successful pest management strategies and contributing to effective phytosanitary management systems in Kenya and beyond.

Acknowledgements

References

Ansari, A.K., van Emden, H.F. & Singh, S.R. (1992) Varietal resistance of cowpea to cowpea aphid, *Aphis craccivora* Koch. *International Journal of Tropical Insect Science* **13**, 199–203, doi: 10.1017/S1742758400014351.

Armstrong, K.F. & Ball, S.L. (2005) DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **360**, 1813–1823.

Armstrong, K.F., Cameron, C.M. & Frampton, E.R. (1997) Fruit fly (Diptera: Tephritidae) species identification: a rapid molecular diagnostic technique for quarantine application. *Bulletin of Entomological Research* **87**, 111–118.

Aslam, M., Razaq, M., Ahmad, F. & Mirza, Y.H. (2007) Population abundance of aphids (*Brevicoryne brassicae* L. and *Lipaphis erysimi* (Kalt.) on Indian mustard (*Brassica juncea* L.). *African Crop Science Conference Proceedings* **8**, 935–938.

Ball, S.L. & Armstrong, K.F. (2006) DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera: Lymantriidae). *Canadian Journal of Forest Research* **36**, 337–350.

Blackman, R.L. & Eastop, V.F. (2000) *Aphids on the World's Crops: an Identification and Information Guide*. 2nd edn.. England, John Wiley & Sons Ltd.

Bock, K.R. & Conti, M. (1974) Cowpea aphid-borne mosaic virus. *CMI/AAB Descriptions of plant viruses*. No. 134.

Brunner, P.C., Fleming, C. & Frey, J.E. (2002) A molecular identification key for economically important thrips species (Thysanoptera: Thripidae) using direct sequencing and a PCR–RFLP-based approach. *Agricultural and Forestry Entomology* **4**, 127–136.

Cocuzza, G.E.M. & Cavalieri, V. (2014) Identification of aphids of *Aphis frangulae*-group living on Lamiaceae species through DNA barcode. *Molecular Ecology Resources* **14**, 447–457.

Coeur d'acier, A., Jousselin, E., Martin, J.F. & Rasplus, J.Y. (2007) Phylogeny of the Genus *Aphis* Linnaeus, 1758 (Homoptera: Aphididae) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **42**, 598–611.

Coeur d'acier, A., Cruaud, A., Artige, E., Genson, G., Clamens, A. L., Pierre, E., Hudaverdian, S., Simon, J.C., Jousselin, E. & Rasplus, J.Y. (2014) DNA barcoding and the associated PhylAphidB@se website for the identification of European aphids (Hemiptera: Aphididae). *PLoS ONE* **9**, e97620. doi: 10.1371/journal.pone.0097620.

Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**, 772.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797.

Emden, H.F.V. & Harrington, R. (2007) *Aphids as Crop Pests*. Oxford, UK, CAB International.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial *cytochrome c oxidase* subunit 1 from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–297.

Foottit, R.G., Maw, H.E.L., Von Dohlen, C.D. & Hebert, P.D.N. (2008) Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. *Molecular Ecology Resources* **8**, 1189–1201.

Foottit, R.G., Maw, H.E.L. & Pike, K.S. (2009) DNA barcodes to explore diversity in aphids (Hemiptera: Aphididae and Adelgidae). *Redia* **92**, 87–91.

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

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B: Biological Sciences* **270**, 313–322.

Hebert, P.D.N., Ratnasingham, S. & deWaard, J.R. (2003b) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Series B: Biological Sciences* **270**, S96–S99.

Karel, A.K. & Autrique, A. (1989) Insects and other pests in Africa. pp. 455–504 *in* Schwartz, H.F. & Pastor-Corrales, M.A. (*Eds*) *Bean Production Problems in the Tropics*. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

Khamis, F.M., Masiga, D.K., Mohamed, S.A., Salifu, D., de Meyer, M. & Ekesi, S. (2012) Taxonomic identity of the invasive fruit fly pest, *Bactrocera invadens*: concordance in morphometry and DNA barcoding. *PLoS ONE* **7**, e44862. doi: 10.1371/journal.pone.0044862.

Kim, H., Hoelmer, K.A., Lee, W., Kwon, Y. & Lee, S. (2010) Molecular and morphological identification of the soybean aphid and other *Aphis* species on the primary host *Rhamnus davurica* in Asia. *Annals of the Entomological Society of America* **103**, 532–543.

Lee, W., Kim, H., Lim, J., Choi, H-R., Kim, Y., Kim, Y-S., Ji, J-Y., Foottit, R.G. & Lee, S. (2011) Barcoding aphids (Hemiptera: Aphididae) of the Korean Peninsula: updating the global data set. *Molecular Ecology Resources* **11**, 32–37.

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

Lozier, J.D., Foottit, R.G., Miller, G.L., Mills, N.J. & Roderick, G.K. (2008) Molecular and morphological evaluation of the aphid genus *Hyalopterus* Koch (Insecta: Hemiptera: Aphididae), with a description of a new species. *Zootaxa* **1688**, 1–19.

Masahiro, M.H.O., Kotsubo, Y., Tajima, R. & Hinomoto, N. (2008). Restriction fragment length polymorphism catalog for molecular identification of Japanese *Tetranychus* spider mites (Acari: Tetranychidae). *Journal of Economic Entomology* **101**, 1167–1175.

Miller, G.L. & Foottit, R.G. (2009) The taxonomy of crop pests: the aphids. pp. 463–473 *in* Foottit, R.G. & Adler, P.H. (*Eds*) *Insect Biodiversity: Science and Society*. Oxford, Wiley-Blackwell.

Nyambo, B. & Löhr, B. (2005) The role and significance of farmer participation in biocontrol-based IPM for brassica crops in East Africa. pp. 290–301 *in* Hoddle, M.S. (*Ed*.) *Proceedings of the Second International Symposium on Biological Control of Arthropods,* vol. I, 12–16 September 2005, Davos, Switzerland.

Peakall, R. & Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–295.

Piffaretti, J., Vanlerberghe-Masutti, F., Tayeh, A., Clamens, A.L., Cœur d'Acier, A. & Jousselin, E. (2012) Molecular phylogeny reveals the existence of two sibling species in the aphid pest *Brachycaudus helichrysi* (Hemiptera: Aphididae). *Zoologica Scripta* **41**, 266–280.

Raboudi, F., Marrakchi, M. & Makni, M. (2002) Polymerase chain reaction-restriction fragment length polymorphism of ribosomal internal transcribed spacer region analysis on analysis on polyacrylamide gel electrophoresis reveals two haplotypes coexisting in *Myzus persicae*. *Electrophoresis* **23**, 186–188.

Rebijith, K.B., Asokan, R., Krishna, V., Kumar, N.K.K. & Ramamurthy, V.V. (2012) Development of species-specific markers and molecular differences in mitochondrial and nuclear DNA sequences of Aphis gossypii and Myzus persicae (Hemiptera: Aphididae). *Florida Entomologist* **95**, 674–682.

Rebijith, K.B., Asokan, R., Kumar, N.K.K., Krishna, V., Chaitanya, B.N. & Ramamurthy, V.V. (2013) DNA barcoding and elucidation of cryptic aphid species (Hemiptera: Aphididae) in India. *Bulletin of Entomological Research* **103**, 601–610.

Sæthre, M-G., Godonou, I., Hofsvang, T., Tepa-Yotto, G.T. & James, B. (2011) Aphids and their natural enemies in vegetable agroecosystems in Benin. *International Journal of Tropical Insect Science* **31**, 103–117.

Shufran, K.A. (2003) Polymerase chain reaction-restriction fragment length polymorphisms identity mtDNA haplotypes of greenbug (Hemiptera: Aphidida). *Journal of the Kansas Entomological Society* **76**, 551–556.

Shufran, K.A. & Puterka, G.J. (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, doi: 10.1603/AN10129.

Singh, S.R. & Allen, D.J. (1980) Pests, diseases, resistance and protection of *Vigna unguiculata* (L.) Walp. pp. 419–433 *in* Summerfield, R.J. & Bunting, H.A. (*Eds*) *Advances in Legumes Science*. Royal Botanic Gardens, Kew, Ministry of Agriculture, Fisheries and Food, London, England.

Sithanantham, S., Nyarko, K.A., Ogutu, W., Chongoti, L., Kibata, G.N., Ouko, J.O., Mukindia, C., Agong, S.G., Seif, A.A. & Loehr, B. (1998) Stakeholder consultation towards initiatives for improved pest management in export vegetables in Kenya. pp. 258–267 *in Proceedings of the 2nd Biennial Crop Protection Conference*. Kenya: Crop Protection Research in Kenya, KARI-ODA.

Spence, N.J., Phiri, N.A., Hughes, S.L., Mwaniki, A., Simons, S., Oduor, G., Chacha, D., Kuria, A., Ndirangu, S., Kibata, G. N. & Marris, G.C. (2007) Economic impact of *turnip mosaic virus, cauliflower mosaic virus* and *beet mosaic virus* in three Kenyan vegetables. *Plant Pathology* **56**, 317–323.

Stamatakis, A. (2014) RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729.

Valenzuela, I., Hoffmann, A.A., Malipatil, M.B., Ridland, P.M. & Weeks, A.R. (2007) Identification of aphid species (Hemiptera: Aphididae: Aphidinae) using a rapid polymerase chain reaction restriction fragment length polymorphism method based on the *cytochrome oxidase* subunit I gene. *Australian Journal of Entomology* **46**, 305–312.

Valenzuela, I., Eastop, V.F., Ridland, P.M. & Weeks, A.R. (2009) Molecular and morphometric data indicate a new species of the aphid Genus *Rhopalosiphum* (Hemiptera: Aphididae). *Annals of the Entomological Society of America* **102**, 914–924.

Virgilio, M., Backeljau, T., Nevado, B. & Meyer, M.D. (2010) Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics* **11**, 206–216.

Wang, J-F., Jiang, L-Y. & Qiao, G-X. (2011) Use of a mitochondrial COI sequence to identify species of the subtribe Aphidina (Hemiptera, Aphididae). *Zookeys* **122**, 1–17.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. & Barton, G.J. (2009) Jalview version 2: a multiple sequence alignment and analysis workbench. *Bioinformatics* **25**, 1189–1191, doi: 10.1093/bioinformatics/btp033.

Worf, G.L., Heimann, M.F. & Pellitteri, P.J. (1995) *Sooty Mold*. pp. 2. University of Wisconsin Cooperative Extension Publication. Madison Wisconsin A2637.

Yeh, H.T., Ko, C.C., Hsu, T.C. & Shih, C.J. (2005) PCR-RFLP technique for identification of *Rhopalosiphum* (Aphididae). *Formosan Entomologist* **25**, 33–46.

Zhang, H-H., Huang, X-L., Jiang, L.Y. & Qiao, G-X. (2010) Subspecies differentiation of *Aphis fabae* Scopoli (Hemiptera: Aphididae) based on morphological and molecular data. *Acta Zootaxonomic Sinica* **35**, 537–547.