Diversity of Lepidoptera associated with macadamia nut damage in South Africa and development of molecular tools to monitor pest populations

Ashleigh K. Smith[1], Bernard Slippers[1], Brett P. Hurley[2], Gerda Fourie[1]

[1] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa

[2] Department of Zoology and Entomology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa

Corresponding author:

Gerda Fourie

Tel.: +2712 420 3854

E-mail: gerda1.fourie@fabi.up.ac.za

Address: FABI University of Pretoria, Lunnon Rd, Hillcrest 0083, Pretoria

Abstract

1. Lepidopteran pests are of major economic importance to macadamia growers worldwide. In South Africa, four species have been associated with macadamia nut damage. Detailed information regarding species composition of the borer larvae in the nuts and the genetic diversity of these species is not however available.
2. Lepidoptera obtained from nuts from the KwaZulu-Natal, Mpumalanga and Limpopo provinces were identified based on COI sequencing.
3. *Thaumatotibia batrachopa* represented 95% of the larvae collected in damaged nuts across all growing regions.
4. The population genetic diversity and structure of the dominant species were determined using a 650 bp section of the mtDNA cytochrome oxidase I gene to construct a parsimony network.
5. Rapid diagnostic tools were developed, which included a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assay for the identification of all four species and species-specific primers for *T. batrachopa* and *T. leucotreta*.
6. Population studies of *T. batrachopa* showed a dominant COI haplotype present in all growing regions, although several unique haplotypes were also present in each region. Overall, the high haplotype diversity observed in this study is in agreement with that of a native population. The populations between the growing regions are, however, not clearly separated and might reflect recent gene flow.

Key words: Lepidoptera, macadamia, *Thaumatotibia batrachopa*, population genetics, phylogenetics

INTRODUCTION

The South African macadamia industry has grown rapidly in the past two decades, with kernel production increasing from an estimated 8789 to 56,550 tonnes between 2000 and 2018 (SAMAC, 2019). The Mpumalanga province is the main production region in South Africa, generating approximately half of the total annual kernel, followed by KwaZulu-Natal and Limpopo, respectively. Other macadamia growing provinces include the Eastern Cape, Western Cape and Gauteng (SAMAC, 2019). At an average price of 71.63 ZAR (4.94 USD) per kg for macadamias in shell in 2018 (SAMAC, 2019), macadamia is considered a high value crop with growing global demand. The growth of macadamia production in South Africa is expected to continue, with significant increases in annual plantings. However, this growing industry faces various obstacles in terms of pests and diseases.

Lepidopteran pests are of particular concern in macadamia orchards because the larval stage feeds on developing nuts. Approximately, 27 Lepidoptera from 11 different families have been associated with macadamia worldwide (Smith, 2020). Tortricidae and Pyralidae are the most dominant families, with eight tortricid and seven pyralid species associated with macadamia (Smith, 2020). Of these species, only four have been reported as pests of macadamia in South Africa, namely, *Thaumatotibia batrachopa* (Meyrick, 1908) (Tortricidae), *T. leucotreta* (Meyrick, 1913) (Tortricidae), *Ectomyelois ceratoniae* (Zeller, 1839) (Pyralidae) and *Cryptophlebia peltastica* (Meyrick, 1921) (Tortricidae) (Schoeman, 2009; Timm et al., 2007). Species composition of these nut borers within and between growing regions in South Africa remains unresolved.

Accurate pest identification is critical for the implementation of responsible and effective control approaches. These control approaches include the use of biological control and pheromones, which have become increasingly attractive due to the need to comply with international import regulations and restrictions regarding chemical usage and the associated maximum residue limits (MRLs) (Ambrus & Yang, 2016). Biological control offers a more environmentally friendly control option by making use of natural enemies to reduce pest populations, as opposed to spraying pesticides. However, the successful implementation of biological control requires detailed knowledge of both the target organism and the biological control agent. Similarly, pheromones, which are an invaluable and environmentally friendly tool for pest control and population monitoring (Reddy & Tangtrakulwanich, 2014), are highly species-specific, and therefore, accurate pest identification is crucial prior to their use.

Morphology is most commonly used for species identification of the four nut borers in South African macadamia orchards. However, identification guides are primarily based on adult morphology and often require expert knowledge (Marsberg et al., 2015; Rentel, 2013; Timm et al., 2007). Although monitoring adults provides insight into the species present within an orchard, it cannot confirm whether macadamia served as a food source for the larval stage. Surveys focussing on lepidopteran nut damage therefore require the collection of larvae from nuts and subsequent rearing to adults, which is time consuming and oftentimes unsuccessful (Marsberg et al., 2015). Another drawback is the difficulty to identify damaged or degraded specimens, especially when the key morphological features used to differentiate closely related species are compromised.

Molecular tools can be used to overcome some of the limitations of morphological taxonomy (Hebert & Gregory, 2005; Hebert, Ratnasingham, et al., 2003; Pentinsaari et al., 2016). The mitochondrial cytochrome c oxidase subunit 1 (COI) gene is one of the most popular genes

used in eukaryotic barcoding and systematics due to its central role in metabolism and appropriate levels of variation for species delineation (Hebert & Gregory, 2005; Hebert, Ratnasingham, et al., 2003; Pentinsaari et al., 2016; Rodrigues et al., 2017; Strüder-Kypke & Lynn, 2010). Sequence data generated from COI can also be used in the development of species-specific PCRs and PCR-restriction fragment length polymorphisms (PCR-RFLPs), which provide rapid and simple diagnostic techniques for routine identification (Bogale et al., 2007; Edwards, 1998; Zhao et al., 2016). By using a molecular approach to identify species, insight into population genetic structure can also be gained to guide management of lepidopteran pests that cause nut damage.

The primary aims of this study were to determine the species composition of Lepidoptera found within macadamia nuts in South Africa and to explore genetic variation within identified species to understand population structure and develop rapid diagnostic tools. Larvae were collected from nuts from farms in the Mpumalanga, KwaZulu-Natal and Limpopo provinces for two consecutive growing seasons and identified using mtDNA COI sequence data. The sequence data were also used to characterize the phylogenetic relationships between these species and other known lepidopteran pests of macadamia, as well as to determine the population structure of the most dominant species. Finally, effective identification systems for the species of concern using species-specific PCRs and PCR-RFLPs were generated.

METHODS

Sampling

Sampling was conducted in the main growing regions in South Africa during the 2017/2018 and 2018/2019 macadamia nut growing seasons, which range from late September to March. Macadamia nuts with visible borer holes were collected throughout each season from farms near Ramsgate in the KwaZulu-Natal province, White River and Kiepersol in the Mpumalanga province and Levubu in the Limpopo province (Figure 1). The nuts were transported to a laboratory at the University of Pretoria, cracked open and inspected for larvae. Over 1000 lepidopteran larvae were extracted and grouped According to location and collection date. The larvae were further grouped based on generic morphological similarities, specifically body form, colour and size. The larvae were stored in absolute ethanol at −20°C.

FIGURE 1. Map of South Africa showing nut borer sampling sites in Limpopo, Mpumalanga and Kwazulu-Natal in relation to the University of Pretoria. The dark grey represents the macadamia growing regions in South Africa. The underlined provinces represent the three main macadamia production regions in South Africa

DNA isolation, PCR amplification and sequencing

DNA was extracted from 190 larvae for downstream analysis. The selected larvae included representatives from all collection dates and morphological groups. Approximately 60 larvae were sequenced from each of the three provinces. DNA was extracted from abdominal tissue using prepGem® Insect (ZyGEM, Hamilton, New Zealand) following the manufacturer's instructions. A ~650 bp section of the mitochondrial gene, cytochrome c oxidase subunit 1 (COI), was amplified using a standard protocol, with primers HCO-2198 (50-TAA ACT TCA GGG TGA CCA AAA AAT CA-30) and LCO-1490 (50-GGT CAA CAA ATC ATA AAG ATA TTG G-30) (Folmer et al., 1994). Reactions were performed in a total volume of 25 μl, using 1 μl DNA (50 ng), 0.5 μl MyTaq™ DNA Polymerase, 5 μl MyTaq™ buffer (BioLine, South Africa), 0.5 μl of each primer and 17.5 μl Sabax water (Adcock Ingram, Bryanston, S.A). An Eppendorf AG Mastercycler (Hamburg, Germany) was used for amplification under the following conditions: 2 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 52°C, 30 s at 72°C and final extension of 1 min at 72°C. The PCR products were analysed on a 1% agarose gel, using a 100 bp DNA Ladder (Promega®, WI) to estimate fragment lengths. PCR products were purified using ExoSap (Applied Biosystems, Foster City, CA) following the manufacturer's instructions.

Sequencing reactions were carried out for both the forward and reverse primers using Big Dye Terminator 1.1 kit (Applied Biosystems, Port Elizabeth) following the manufacturer's instructions. The sequencing PCR products were sequenced using an ABI Prism™ 3100 Genetic Analyser (Applied BioSystems, USA). Forward and reverse sequence reads were checked for base calling accuracy, manually edited and consensus sequences were generated using Biological Sequence Alignment Editor (BioEdit) (Hall, 1999) version 7.0.9. Sequences were also translated into amino acids in order to screen for the presence of premature stop codons. Thereafter, sequence identities were determined by Basic Local Alignment Search Tool (BLAST) against the NCBI database (www.ncbi.nlm.nih.gov). In the event that the most significant alignment indicated an identity score of under 98%, COI divergence between the query sequence and found sequence was calculated using the Kimura 2-parameter model in Molecular Evolutionary Genetics Analysis (MEGA) version 7.0.26 (Tamura et al., 2011).

Phylogenetic analysis

COI sequences generated in this study (Table 1), sequences of closely related species from BLAST, as well as sequences of other known lepidopteran pests of macadamia, were downloaded from GenBank and included in the analysis (Table 2). The majority of the downloaded sequences were generated and published by Marsberg et al. (2015) where the identities of the specimens were also confirmed morphologically by experts. Sequences were aligned using multiple alignment using fast Fourier transform (MAFFT) version 7 (https://mafft.cbrc.jp/alignment/software/) (Katoh & Standley, 2013) set on the L-INS-i option (Katoh et al., 2002, 2005).

TABLE 1. Collection details of specimens sequenced to determine species presence and composition of larval infestations inside macadamia nuts in South Africa

Family	Species	Province	GPS coordinates		No. of	Lab ID
			Latitude	Longitude	larvae	
Tortricidae	*Thaumatotibia batrachopa*	KwaZulu-Natal	−30.8881	30.3056	64	MLK1-MLK64
		Limpopo	−23.0464	30.2685	65	MLL1-MLL65
		Mpumalanga	−25.0806	31.0194	26	MLM1-MLM26
			−25.3299	31.0163	26	MLM27-MLM52
	Thaumatotibia leucotreta	KwaZulu-Natal	30.3056	30.3056	1	MLK65
		Limpopo	30.2685	30.2685	1	MLL66
		Mpumalanga	31.0194	31.0194	3	MLM53-MLM55
Pyralidae	Unknown	KwaZulu-Natal	30.3056	30.3056	3	MLK66-MLK68
		Mpumalanga	31.0194	31.0194	1	MLM56

TABLE 2. COI sequences downloaded from NCBI GenBank and included in phylogenetic analysis

Family	Species	Isolate/voucher	Sequence length	Year	GenBank accession
Papilionidae	*Papilio demodocus*	Voucher L33	636	2014	KP083445
		Voucher L34	571	2014	KP083446
Pyralidae	*Amyelois transitella*	Voucher BIOUG<CAN>:MDOK-3714	658	2019	HM388100
	Assara seminivale	Voucher 11ANIC-03443	658	2013	KF398833
	Batrachedra amydraula	Voucher GB316	676	2016	KT827248
	Cadra cautella	Isolate BRG8	477	2018	KX399347
	Cadra figulilella	Isolate WH20	477	2018	KX399381
	Cryptoblabes gnidiella	Isolate nar_kad-mrkz	474	2014	KM035797
	Cryptoblabes hemigypsa	Voucher 11ANIC-01888	658	2013	KF405266
	Ectomyelois ceratoniae	Voucher 11ANIC-02067	658	2013	KF397550
		Voucher 11ANIC-03182	658	2013	KF405701
		Voucher L28	663	2014	KP083440
		Isolate Ke1	604	2016	KU896479
		Isolate 11	661	2018	MG489943
		Voucher KLM Lep 03053	658	2018	MH417697
Tortricidae	*Choristoneura rosaceana*	Voucher BIOUG<CAN>:10BBLEP-00555	658	2019	HQ985901
	Cryptophlebia illepida	Voucher JWB-08-0109-1	658	2013	KF491660
	Cryptophlebia peltastica	Voucher KLM Lep 03097	658	2018	MH415950
		Isolate KS7B	677	2017	KX150514
		Isolate KS6A	677	2017	KX150511
		Voucher USNM:ENT:00676545	658	2014	KJ592246
	Cryptophlebia rutilescens	Voucher 11ANIC-12831	658	2013	KF405947
		Voucher 11ANIC-12832	658	2013	KF397939
	Epiphyas postvittana	Isolate LBAM_090326-5	639	2016	GU827562
	Gymnandrosoma aurantianum	BOLD:AAA4028	658	2012	HM429740
	Thaumatotibia batrachopa	Voucher L24	660	2014	KP083436
		Voucher L25	663	2014	KP083437
		Voucher USNM:ENT:00676531	658	2014	KJ592085
	Thaumatotibia leucotreta	Voucher USNM:ENT:00676526	658	2014	KJ592079
		Voucher USNM:ENT:00718973	658	2014	KJ592153
		Voucher USNM:ENT:00808074	658	2014	KJ592321
	Thaumatotibia maculata	Voucher 11ANIC-12868	658	2014	KF399439
		Voucher YAWCATCR0540	636	2019	MK019945

Phylogenetic analyses were performed using maximum likelihood (ML) and Bayesian inference. For ML analysis, the GTR model with gamma correction (G) to account for among site rate heterogeneity (Tavaré, 1986) was the best-fit substitution model according to jModeltest (Posada, 2008). Phylogenetic analysis was conducted using Randomized Axelerated Maximum Likelihood (RAxML) version 8 (Stamatakis, 2014), and clade support

was estimated with 1000 bootstrap replicates. *Papilio demodocus* (Esper, 1799) (Papilionidae) was used as an outgroup. Bayesian inference was conducted using a Bayesian method implemented in BEAST v1.10.4 (Suchard et al., 2018). A Markov chain Monte Carlo was run for 10 million generations and sampled every 1000 generations. GTR + G was used for the substitution model (estimated for the dataset).

Morphological species identification

To validate species identification based on COI BLAST searches against the NCBI database and phylogenetic analysis, *T. batrachopa* adults (Figure 2) and sent to a taxonomist at the Agricultural Research Council (ARC), Biosystematics division, in Roodeplaat, South Africa, for identification. Larvae were reared to adults using a modified version of an artificial diet described previously for *T. leucotreta* (Moore et al., 2014). Modifications included increasing the water content to prevent desiccation of the diet, and the addition of sodium propionate and methyl paraben to inhibit fungal growth (Smith, 2020).

FIGURE 2. Adult *Thaumatotibia batrachopa*

Population genetic diversity analysis

All generated *T. batrachopa* sequences were used for haplotype diversity analysis. Genetic diversity parameters such as the number of haplotypes (h), haplotype diversity (Hd), number of segregating sites (S), average number of nucleotide differences (k), nucleotide diversity (Pi), average number of nucleotide substitutions per site (Dxy) and net nucleotide substitutions per site (Da) were calculated using DNA Sequence Polymorphism (DnaSP)

version 5.10.01 (Librado & Rozas, 2009). Pairwise mean genetic distance between populations was computed using the Kimura 2-parameter model in MEGA version 7.0.26 (Tamura et al., 2011). A haplotype network was constructed using Population Analysis with Reticulate Trees (PopART) version 1.7 (Leigh & Bryant, 2015).

Species-specific PCRs and PCR-RLFPs

To develop diagnostic PCRs for *T. batrachopa* and *T. leucotreta*, sequences of a ~ 650 bp region of the COI gene were manually screened for polymorphic regions. Primers were designed to span polymorphic regions that contained at least three single nucleotide polymorphisms unique to the target species using CLC Main Workbench 8.1.3 (QIAGEN Bioinformatics). Primers were designed to be between 18 and 22 bp in length, with CG content between 40% and 60% and melting temperatures above 45°C. Forward and reverse primers were generated for both species (Table 3).

TABLE 3. Species-specific primers for *Thaumatotibia batrachopa* and *T. leucotreta*

Species	Primer set	Primer	Sequence	Length (bp)	Tm (°C)	Fragment length (bp)
Thaumatotibia batrachopa	MNB	MNBF	CCCACAGCTCAAACAAATA	19	46.77	290 bp
		MNBR	TGGCTTTCCCCCGAATAA	18	45.28	
Thaumatotibia leucotreta	FCM	FCMF	CTGATCTTCCGCTATGTG	18	48.04	120 bp
		FCMR	CTTCTACCCCCCTCTATTT	19	48.93	

To determine specificity of each diagnostic primer pair, DNA of 10 *T. batrachopa* larvae and 10 *T. leucotreta* larvae were used in PCR reactions. Since only five *T. leucotreta* larvae were collected during the survey, additional larvae were purchased from River BioScience (Pty) Ltd. (Port Elizabeth, South Africa). The PCR conditions were kept similar to previously described, with the exception of the annealing temperatures of 61°C for the MNB primer pair and 48°C for the FCM primer pair. To further test the diagnostic power of the PCRs, 40 additional larvae, collected during the survey, were randomly selected for amplification.

A PCR-RFLP assay was developed to distinguish between all four Lepidoptera species reported from macadamia in South Africa, namely, *T. batrachopa*, *T. leucotreta*, *E. ceratoniae* and *C. peltastica*. For each species, COI sequences were either generated or downloaded from NCBI GenBank, aligned and trimmed to 656 bp in BioEdit (Hall, 1999) version 7.0.9. The restriction site analysis tool in CLC Main Workbench 8.1.3 (QIAGEN Bioinformatics) was used to digest the sequence dataset in silico with various restriction enzymes. The resulting COI-RFLP profiles were screened and restriction digests that resulted in unique COI-RFLP profiles for each of the four species were identified. Endonucleases *Vsp*I and *Acc*I were selected for a double digest, which produced unambiguous identification of the four species. To test the specificity of the assay, COI sequences of closely related Tortricidae and Pyralidae species were downloaded from NCBI GenBank, included in the sequence dataset and digested in silico with *Vsp*I and *Acc*I. The species chosen consists of those that have previously been associated with macadamia for which COI sequences are available, including *Epiphyas postvittana* (Walker, 1863), *Gymnandrosoma aurantianum* (Lima, 1927), *Amyelois transitella* (Walker, 1863), *Assara seminivale* (Turner, 1904), *Batrachedra amydraula* (Meyrick, 1916), *Cadra cautella* (Walker, 1863), *Cadra figulilella* (Gregson, 1871), *Cryptoblabes gnidiella* (Millière, 1867) and *C. hemigypsa* (Turner, 1913). Of these species, Ca. *cautella* and *Cr. gnidiella* have been reported in South Africa (Smith, 2020).

RESULTS

Phylogenetic analysis

Based on percentage identity values of 98% or higher, BLAST searches against the NCBI database identified 181 of the 190 individuals as *T. batrachopa* and five as *T. leucotreta* (GenBank accession numbers MZ420787 - MZ420976). Species identity of *T. batrachopa* was also confirmed based on adult morphology by a taxonomist. Of the *T. leucotreta* larvae, one was collected from Limpopo, one from KwaZulu-Natal and three from Mpumalanga. Four larvae could not be identified by BLAST analysis. The closest sequence producing significant alignment for the four unknown individuals was *E. ceratoniae* (KF405701), with max scores of 712 or lower and percentage identity of around 90%. The COI divergences between known *E. ceratoniae* sequences and the unknown sequences all exceeded 12%. Mean COI divergence between congeneric species pairs ranges from 6.6% to 11.5% (Hebert, Cywinska, et al., 2003), and thus the unknown sequences could not be identified as *E. ceratoniae*. No *E. ceratoniae* or *C. peltastica* larvae were identified from macadamia nuts during the survey.

ML analysis resolved the sequences into two distinct family clades, namely Tortricidae and Pyralidae (Figure 3). The family clades were supported with bootstrap values of 97% and 84%, respectively. Within the Tortricidae clade, eight of the nine tortricid species were separated into individual clades. These clades represented *Choristoneura rosaceana* (Harris, 1841), *C. peltastica*, *Cr. rutilescens, Ep. postvittana*, *G. aurantianum*, *T. batrachopa*, *T. leucotreta* and *T. maculata*. The remaining tortricid species, *C. illepida* (Butler, 1882) (KF491660), was grouped within the *C. peltastica* clade with strong bootstrap support (Figure 3).
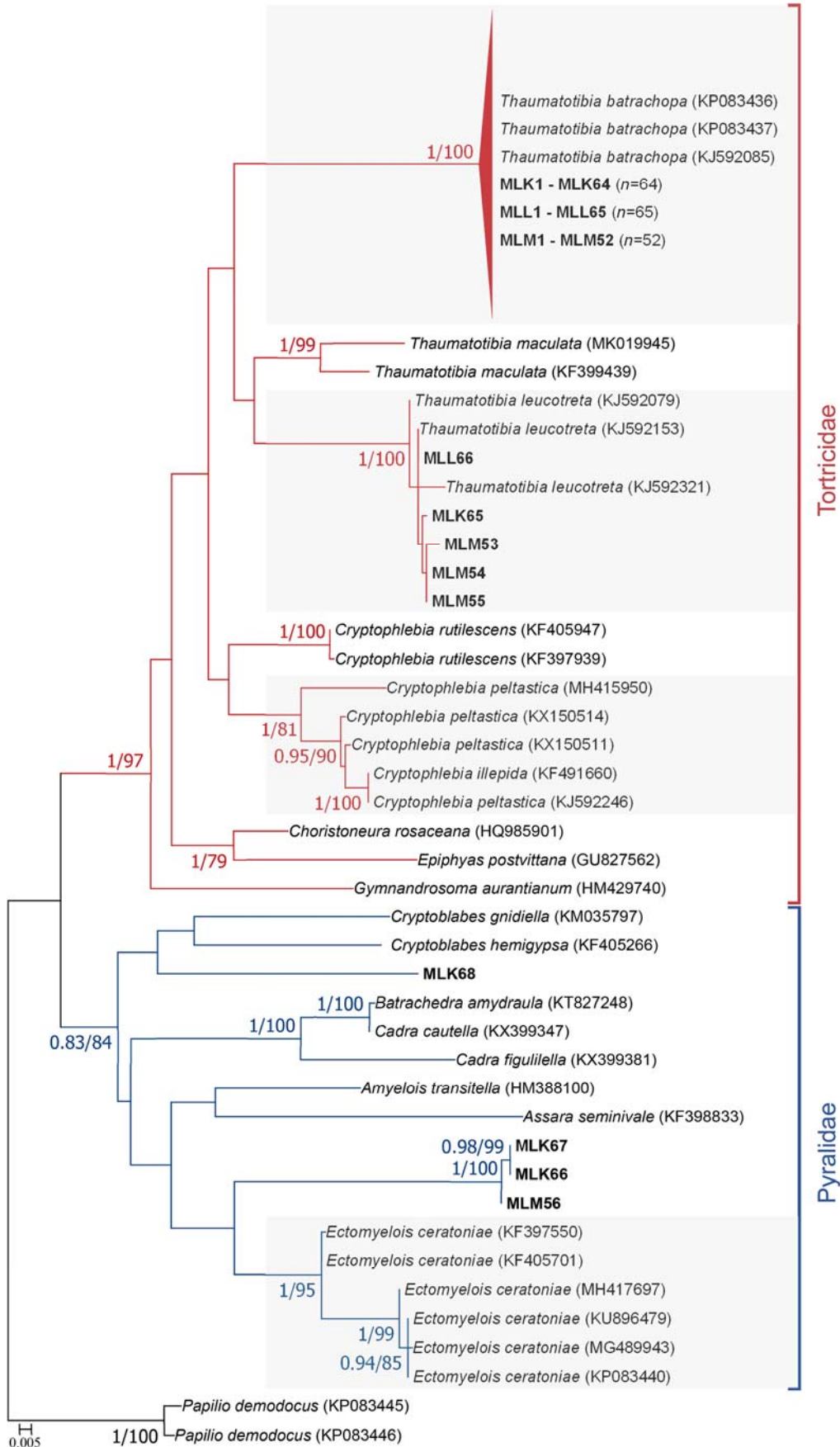
FIGURE 3. Phylogeny representing mtDNA COI barcoding sequences of Lepidoptera associated with macadamia and closely related lepidopteran species. Node numbers indicate Bayesian posterior probabilities (BEAST) expressed as decimals and maximum likelihood bootstrap support values (RaxML) expressed as percentages. Bold isolates represent sequences generated during this study from lepidopteran larvae collected from macadamia nuts during the 2017/2018 and 2018/2019 growing season in Mpumalanga (M), Limpopo (L) and KwaZulu-Natal (K) (Table 1). Grey blocks highlight clades representative of species reported as pests of macadamia in South Africa, namely *Thaumatotibia batrachopa, T. leucotreta, Cryptophlebia peltastica* and *Ectomyelois ceratoniae*

All eight of the pyralid species included in the analysis grouped within the Pyralidae clade, namely *Am. transitella*, *As. seminivale*, *B. amydraula*, Ca. *cautella*, Ca. *figulilella*, *Cr. gnidiella*, *Cr. hemigypsa* and *E. ceratoniae*. Four unknown individuals from our study also grouped within this clade. However, the resolution of the species relationships within the family clades remains low, with many polytomies and low bootstrap support for branches. As a result, the closest relative of the unknown individuals could not be resolved (Figure 3).

Population genetic diversity analysis

Analysis of the mtDNA COI gene of *T. batrachopa* yielded 29 haplotypes from 181 sequences (Table 4). The haplotype network indicated a dominant haplotype (haplotype 1) shared between the three regions (Mpumalanga, KwaZulu-Natal and Limpopo) (Figure 4). Haplotype 1 represented 36% of the total number of individuals, while the second largest haplotype (haplotype 2) represented 14%. Both haplotypes were shared between all three regions. A further nine haplotypes were shared between either all or two regions; however, there were various haplotypes that were only found in one of the growing regions (Figure 5). KwaZulu-Natal had 10 unique haplotypes, Limpopo had five, while Mpumalanga had two (Table 5).

TABLE 4. The number and names of sequences representing each of the 29 haplotypes identified for *Thaumatotibia batrachopa* using mtDNA COI sequence data

Haplotype	No. of sequences	Sequence name
1	65	MLK1, MLK2, MLK3, MLK10, MLK11, MLK13, MLK15, MLK16, MLK18, MLK20, MLK25, MLK26, MLK28, MLK31, MLK32, MLK41, MLK44, MLK45, MLK51, MLK52, MLK54, MLK61, MLL2, MLL7, MLL9, MLL17, MLL19, MLL23, MLL36, MLL40, MLL44, MLL48, MLL51, MLL55, MLL64, MLL65, MLM3, MLM4, MLM6, MLM7, MLM8, MLM11, MLM12, MLM15, MLM17, MLM18, MLM19, MLM21, MLM22, MLM24, MLM25, MLM26, MLM30, MLM31, MLM40, MLM41, MLM42, MLM43, MLM44, MLM47, MLM48, MLM49, MLM50, MLM51, MLM52
2	26	MLK6, MLK17, MLK43, MLK50, MLK53, MLK55, MLK56, MLK62, MLK64, MLL1, MLL18, MLL21, MLL25, MLL33, MLL46, MLL52, MLL53, MLL54, MLL60, MLL61, MLM23, MLM35, MLM36, MLM37, MLM38, MLM39
3	11	MLK12, MLK27, MLK29, MLK33, MLL8, MLL42, MLL43, MLM9, MLM10, MLM14, MLM29
4	10	MLK8, MLL5, MLL6, MLL10, MLL26, MLL28, MLL39, MLL47, MLL56, MLM32
5	8	MLL3, MLL4, MLL11, MLL20, MLL59, MLL63, MLM45, MLM46
6	5	MLK19, MLK24, MLM16, MLM33, MLM34
7	5	MLL34, MLL49, MLL50, MLM2, MLM13
8	5	MLL16, MLL24, MLL27, MLL38, MLM1
9	3	MLK57, MLK58, MLL14
10	3	MLL29, MLL41, MLM20
11	2	MLK14, MLM5
12	2	MLK49, MLL15
13	8	MLK5, MLK7, MLK9, MLK21, MLK22, MLK23, MLK59, MLK60
14	3	MLK30, MLK40, MLK63
15	3	MLK35, MLK36, MLK37
16	2	MLK34, MLK39
17	1	MLK4
18	1	MLK48
19	1	MLK38
20	1	MLK47
21	1	MLK42
22	1	MLK46
23	6	MLL22, MLL31, MLL45, MLL57, MLL58, MLL62
24	3	MLL30, MLL32, MLL35
25	1	MLL37
26	1	MLL12
27	1	MLL13
28	1	MLM27
29	1	MLM28

Note: The collection location of the specimen is indicated by the last letter in the sequence name, namely KwaZulu-Natal (K), Limpopo (L) and Mpumalanga (M).
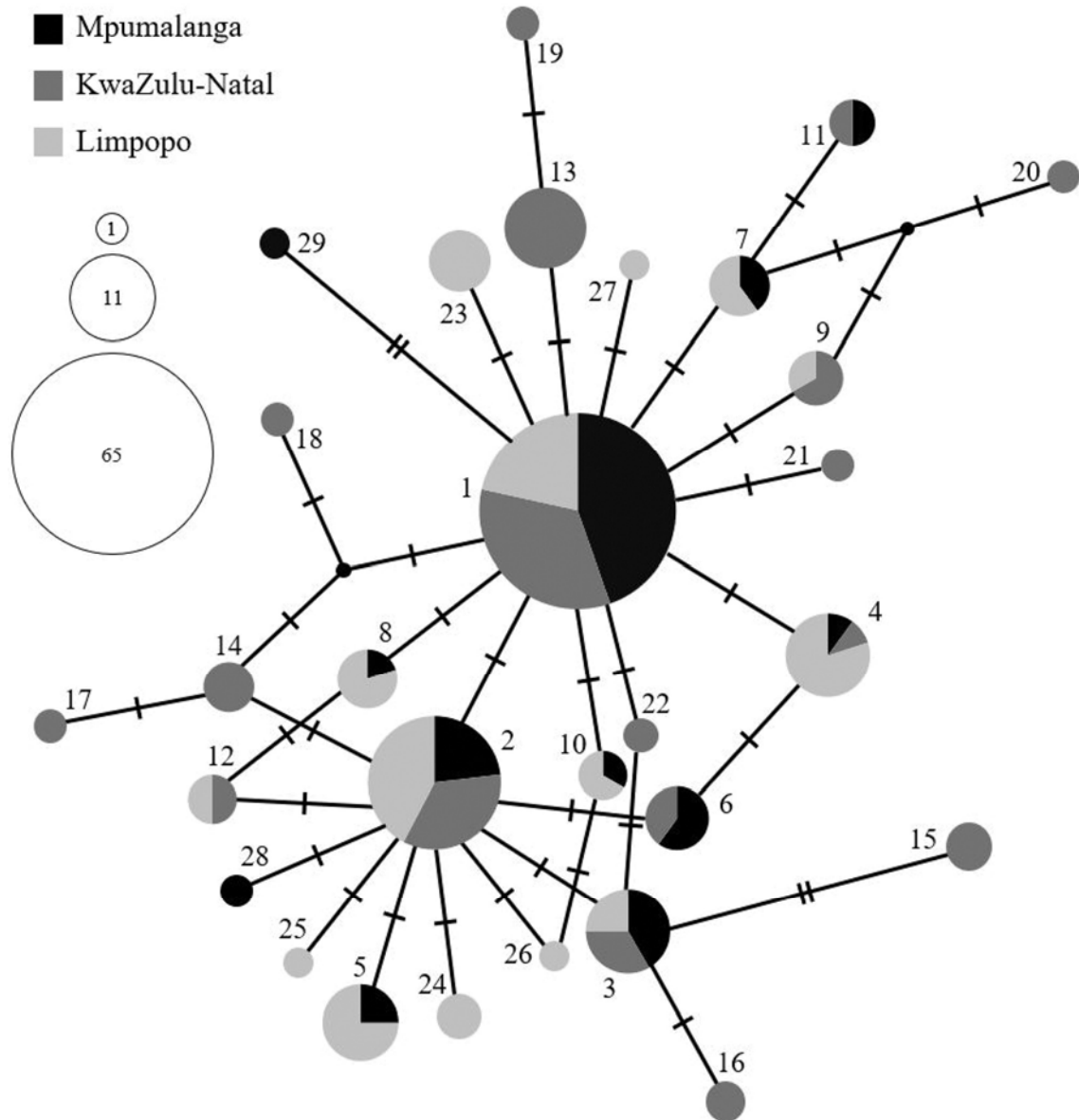
FIGURE 4. Cytochrome c oxidase subunit I (COI) haplotype network of *Thaumatotibia batrachopa*. Each circle represents a single haplotype and is proportional to the number of individuals represented. The number of individuals is shown in the empty circles adjacent to the haplotype network. Different colours indicate the geographical localities in which the haplotype occurs. Hatch marks on lines connecting circles represents mutational steps
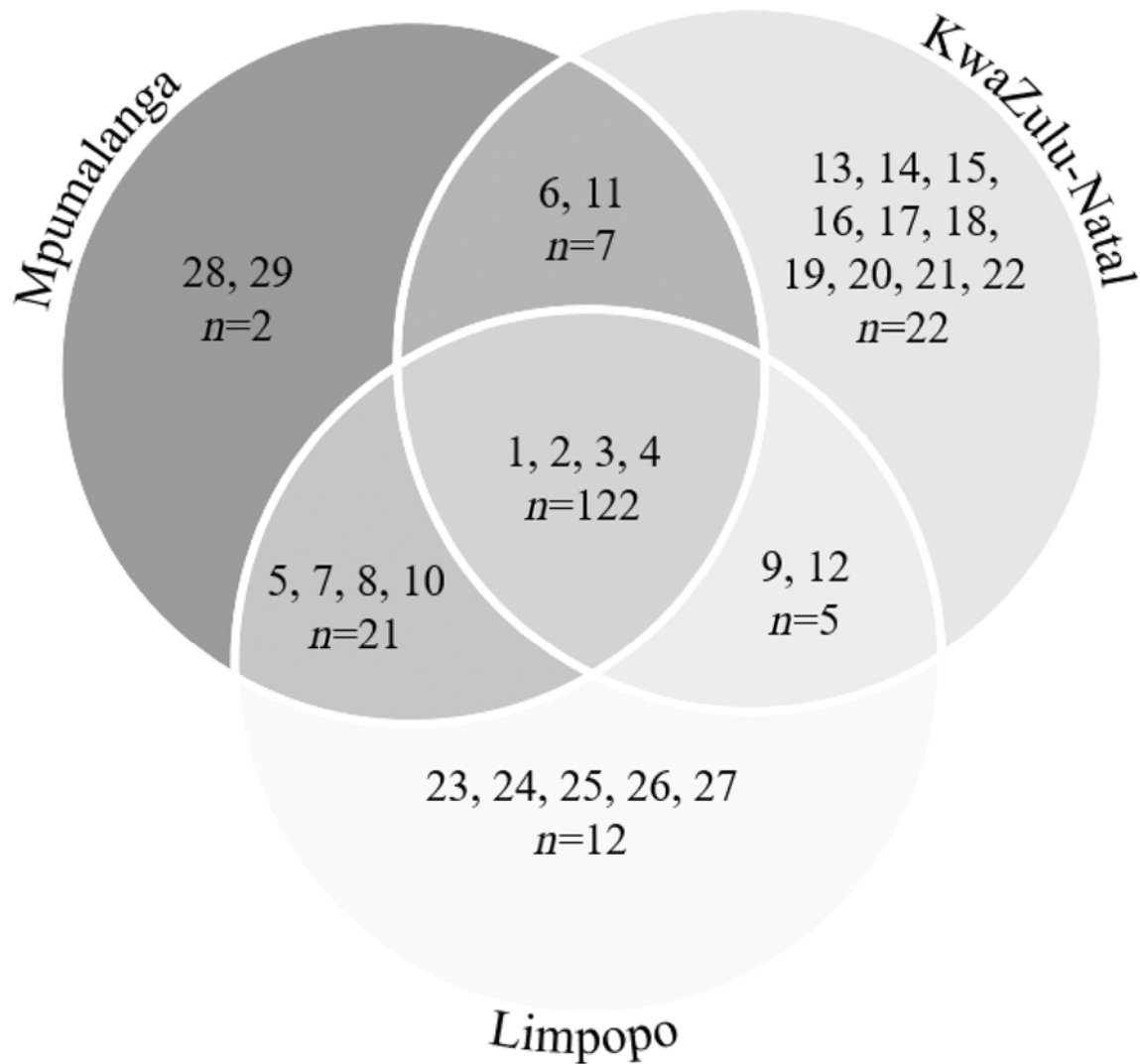
FIGURE 5. Venn diagram indicating geographic occurrences of the 29 haplotypes identified for *Thaumatotibia batrachopa*. The sum of the individuals (*n*) represented by the haplotypes present is indicated for each section

TABLE 5. Summary statistics of genetic diversity of *Thaumatotibia batrachopa* per sampling region based on the mtDNA cytochrome oxidase I gene using DnaSP. 5.10

Population	#	h	u	Hd	S	k	Pi
All	181	29		0.805	13	1.54187	0.00280
Mpumalanga	52	12	2	0.674	11	1.14103	0.00207
KwaZulu-Natal	64	18	10	0.846	16	1.84325	0.00335
Limpopo	65	15	5	0.894	12	1.64135	0.00298

Abbreviations: #, number of sequences; h, number of haplotypes; u, number of unique haplotypes; Hd, haplotype diversity; S, number of variable sites; k, average number of nucleotide differences; Pi, nucleotide diversity.

An estimated overall nucleotide diversity (Pi) of 0.0028 (S = 13, h = 29, Hd = 0.805, K = 1.54187) was obtained across all sampling regions, with the KwaZulu-Natal population showing a Pi of 0.00335 (S = 16, h = 18, Hd = 0.846, K = 1.84325), Limpopo of 0.00298, (S = 12, h = 15, Hd = 0.894, K = 1.64135) and Mpumalanga of 0.00207 (S = 11, h = 12, Hd = 0.674, K = 1.14103) (Table 5). Pairwise mean genetic distance between populations

14

was 0.003 for all compared regions. The average number of nucleotide substitutions per site (Dxy) and the net nucleotide substitutions per site (Da) between populations ranged from 0.00252 to 0.00295 and −0.00003 to 0.00001, respectively.

Species-specific PCRs and PCR-RFLPs

Initial species-specific PCRs conducted using DNA from 20 previously identified larvae confirmed that the primer sets designed for *T. batrachopa* (MNB) and *T. leucotreta* (FCM) (Table 3) only amplified DNA of the species for which they were designed. Primer set MNB produced a single band of approximately 290 bp for *T. batrachopa*, while primer set FCM produced a single band of approximately 120 bp for *T. leucotreta*. Both primer sets were tested on 40 randomly selected larvae, all of which were identified as *T. batrachopa*.

RFLP patterns for 656 bp COI sequences for *T. batrachopa*, *T. leucotreta*, *E. ceratoniae* and *C. peltastica* were distinguishable after a double digest with restriction enzymes *Vsp*I and *Acc*I (Figure 6). *Vsp*I was used to differentiate between *T. batrachopa*, *T. leucotreta* and *E. ceratoniae*, while *Acc*I was used to differentiate *C. peltastica*. *T. batrachopa* was distinguished from *T. leucotreta* and *E. ceratoniae* in that it harboured a single *Vsp*I restriction site, resulting in a COI-RFLP profile consisting of two fragments of 599 and 59 bp. Both *T. leucotreta* and *E. ceratoniae* had two *Vsp*I restriction sites; however, they were easily distinguished based on differing sizes of their largest fragments, with *T. leucotreta* producing fragments of 199, 174 and 287 bp, and *E. ceratoniae* producing fragments of 440, 23 and 197 bp. *C. peltastica* was distinguished from the other species in that it did not contain a *Vsp*I restriction site but had a unique *Acc*I restriction site, resulting in a COI-RFLP profile consisting of two fragments of 338 and 320 bp. In silico digests of numerous tortricid and pyralid species confirmed unique COI-RFLP profiles for the four species in question.
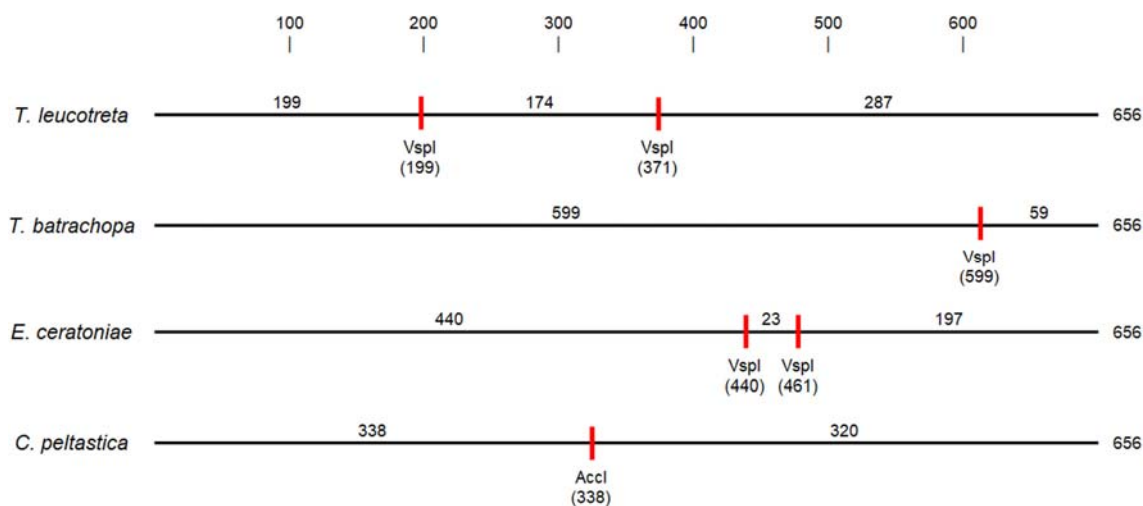

FIGURE 6. Restriction enzyme map showing restriction sites for endonucleases *Vsp*I and *Acc*I in a 656 bp fragment of the mtDNA COI gene for *Thaumatotibia batrachopa*, *T. leucotreta*, *Ectomyelois ceratoniae* and *Cryptophlebia peltastica*

DISCUSSION

In this study, the dominant lepidopteran pest of macadamia was identified as *T. batrachopa*. *T. batrachopa* has a limited geographic distribution, reported only in Malawi and South

Africa (Marsberg et al., 2015; Rentel, 2013), and is believed to feeds almost exclusively on macadamia (Timm et al., 2006). A single report of *T. batrachopa* feeding on citrus has been recorded from a packhouse in South Africa (Bedford et al., 1998), but due to their morphologically similar larval stages, *T. leucotreta* may be mistaken for *T. batrachopa*, particularly in areas where citrus and macadamia orchards are in close proximity (Marsberg et al., 2015; Newton, 1988).

T. batrachopa comprised over 95% of the larvae obtained from macadamia nuts throughout the 2017/2018 and 2018/2019 growing seasons in Mpumalanga, Limpopo and KwaZulu-Natal. During the survey, only five *T. leucotreta* larvae were isolated, while no *E. ceratoniae* or *C. peltastica* larvae were collected from nuts. However, due to the limited geographic spread of the sampling sites, it is possible that the species composition may be different in unsampled areas. Our results are in contrast to a survey conducted in 2002 in Levubu (Limpopo) that showed *T. leucotreta* comprised 60.7% of the collected lepidopteran larvae, while *T. batrachopa* comprised 34.0% (Mlanjeni et al., 2002). The diminution of *T. leucotreta* within macadamia orchards since 2002 may be a result of control regimes. South Africa has a range of products registered for the control of lepidopteran pests in macadamia orchards (Subtrop, 2018); however, these chemicals offer broad-spectrum action (Aktar et al., 2009). It is therefore unlikely that these products have largely eradicated *T. leucotreta* from macadamia orchards, while *T. batrachopa* populations remain prevalent. The contrast in results may be better explained by natural species fluctuations caused by factors such as climate change, food availability and quality, parasitoids, pathogens and predators (Myers & Cory, 2013). Due to the dynamic nature of populations, pest populations within orchards need to be monitored continually to tailor control strategies to specific pests present.

Cryptophlebia peltastica and *E. ceratoniae* have previously been described as pests of macadamia in South Africa (Manrakhan et al., 2008; Moore, 2012), but no larvae of these species were collected during this survey. This may suggest that these species are not pests of concern to South African macadamia growers but instead exist as sporadic and/or opportunistic feeders. In South Africa, *E. ceratoniae* is known to mainly feed on injured fruits and stored goods (Moore, 2012) and thus may be more problematic in macadamia drying bins or storage facilities. The four larvae that could not be identified are also unlikely to be cause for concern in macadamia orchards. They occurred in low numbers and were only collected from Mpumalanga and KwaZulu-Natal. It is possible these individuals occurred as opportunistic pests. Nevertheless, further monitoring over a wider range of farms and growing seasons is required to confirm this, and the identification techniques developed in this study would prove useful for that purpose.

Phylogenetic analyses resolved all of the species except for *C. illepida* and *C. peltastica*. Unfortunately, only a single COI sequence was available for *C. illepida* from NCBI GenBank. *C. illepida* is endemic to the islands of Kauai, Oahu, Molokai, Maui, Lānai and Hawaii (Jones, 1994; Jones et al., 1997). In contrast, *C. peltastica* occurs throughout Africa, Madagascar, Seychelles, Mauritius, Asia and the Bahamas (Timm et al., 2006). The grouping of *C. illepida* within the *C. peltastica* clade may be a result of sequence labelling or identification error, or the same species has been described in separate regions as two taxa. This would require further collections from the different regions and examination of type specimens to confirm.

The population genetic analysis of *T. batrachopa* based on a region of the mtDNA COI gene revealed a high haplotype diversity, which is to be expected from an insect population in its

native range, although clear genetic separation between populations from different regions was absent. Results from this study suggest that there is gene flow between *T. batrachopa* populations from different regions and that geographically distinct populations that influence biological traits of relevance to management are not expected. In contrast, Timm et al. (2006) showed significant genetic differentiation among populations of *T. batrachopa* sampled from Mpumalanga and Limpopo using amplified fragment length polymorphism (AFLP) analysis, indicating low gene flow between populations. During the decade between this study and our study, the South African macadamia industry has rapidly expanded. It is possible that this expansion may have attributed to the increased gene flow between regions through the movement of host material. However, population genetic studies using more diverse nuclear markers will be needed to gain further insight into population genetic structuring.

Considering the global increase in annual macadamia plantings, consistent monitoring of pest populations is necessary for the growth and long-term sustainability of the industry. The development of species-specific primers and PCR-RFLP assays has promising application for rapid identification of species responsible for causing damage within South African macadamia orchards, information that is paramount for the implementation of effective control regimes. Our research provides insight into the current species composition of lepidopteran pests of macadamia in South Africa and provides a cost-effective means to monitor pest populations in future.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors of this manuscript certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership or other equity interest and expert testimony or patent-licensing arrangements) or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this article.

AUTHOR CONTRIBUTIONS

All authors conceived research. Ashleigh K. Smith conducted experiments. All authors contributed material. Ashleigh K. Smith and Gerda Fourie analysed data and conducted statistical analyses. Ashleigh K. Smith wrote the manuscript. Gerda Fourie secured funding. All authors read and approved the manuscript.

REFERENCES

Aktar, M.W., Sengupta, D. & Chowdhury, A. (2009) Impact of pesticides use in agriculture: their benefits and hazards. *Interdisciplinary Toxicology*, 2, 1– 12.

Ambrus, Á. & Yang, Y.Z. (2016) Global harmonization of maximum residue limits for pesticides. *Journal of Agricultural and Food Chemistry*, 64, 30– 35.

Bedford, E.C.G., Van den Berg, M.A. & De Villiers, E. (1998) Citrus pests in the Republic of South Africa. Nelspruit, South Africa: Institute for Tropical and Subtropical Crops.

Bogale, M., Wingfield, B.D., Wingfield, M.J. & Steenkamp, E.T. (2007) Species-specific primers for *fusarium redolens* and a PCR-RFLP technique to distinguish among three clades of *Fusarium oxysporum*. *FEMS Microbiology Letters*, 271, 27– 32.

Edwards, K.J. (1998) Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLPs). In: A. Karp, P.G. Isaac & D.S. Ingram (Eds.) Molecular tools for screening biodiversity: plants and animals. The Netherlands: Springer, pp. 131– 134.

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.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95– 98.

Hebert, P.D., Ratnasingham, S. & De Waard, J.R. (2003) 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.

Hebert, P.D.N., Cywinska, A., Ball, S.L. & Dewaard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B. Biological Sciences*, 270, 313– 321.

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

Jones, V.P. (1994) Feeding by *Cryptophlebia illepida* and *C. ombrodelta* (Lepidoptera: Tortricidae) on macadamia nut abortion. *Journal of Economic Entomology*, 87, 781– 786.

Jones, V.P., Tome, C.H.M. & Caprio, L.C. (1997) Life tables for the koa seedworm (Lepidoptera: Tortricidae) based on degree-day demography. *Environmental Entomology*, 26, 1291– 1298.

Katoh, K., Kuma, K., Toh, H. & Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33, 511– 518.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30, 3059– 3066.

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772– 780.

Leigh, J.W. & Bryant, D. (2015) PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6, 1110– 1116.

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

Manrakhan, A., Abeeluck, D. & Gokool, A. (2008) Assessment of damage by *Cryptophlebia peltastica* (Meyrick) (Lepidoptera: Tortricidae) in litchi orchards in Mauritius. *African Entomology*, 16, 203– 208.

Marsberg, T., Moore, S.D., Hill, M.P. & Timm, A.E. (2015) DNA-based identification of Lepidoptera associated with citrus in South Africa. *African Entomology*, 23, 165– 171.

Mlanjeni, N., Moore, S. & Richards, G. (2002) Investigation of the macadamia nutborer species complex and the natural enemies of their larvae in South Africa. *Yearbook of the Southern African Macadamia Growers Association*, 10, 62– 71.

Moore, S., Richards, G.I., Chambers, C. & Hendry, D. (2014) An improved larval diet for commercial mass rearing of the false codling moth, *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera: Tortricidae). *African Entomology*, 22, 216– 219.

Moore, S. D. (2012) *Moths and butterflies: carob moth*, Citrus Research International, viewed 8 February 2019, <https://www.citrusres.com/system/files/documents/production-guidelines/Ch%203-9-2%20Carob%20moth%20-%20Jan%202012.pdf>

Myers, J.H. & Cory, J.S. (2013) Population cycles in forest Lepidoptera revisited. *Annual Review of Ecology, Evolution, and Systematics*, 44, 565– 592.

Newton, P.J. (1988) Inversely density-dependent egg parasitism in patchy distributions of the citrus pest *Cryptophlebia leucotreta* (Lepidoptera: Tortricidae) and its agricultural efficiency. *Journal of Applied Ecology*, 25, 145– 162.

Pentinsaari, M., Salmela, H., Mutanen, M. & Roslin, T. (2016) Molecular evolution of a widely adopted taxonomic marker (COI) across the animal tree of life. *Scientific Reports*, 6, 35275.

Posada, D. (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, 25, 1253– 1256.

Reddy, G.V.P. & Tangtrakulwanich, K. (2014) Potential application of pheromones in monitoring, mating disruption, and control of click beetles (Coleoptera: Elateridae). *ISRN Entomology*, 2014, 531061.

Rentel, M. (2013) *Morphology and taxonomy of tortricid moth pests attacking fruit crops in South* Africa. Unpublished doctoral dissertation, Western Cape Provence, South Africa, University of Stellenbosch.

Rodrigues, M.S., Morelli, K.A. & Jansen, A.M. (2017) Cytochrome c oxidase subunit 1 gene as a DNA barcode for discriminating *Trypanosoma cruzi* DTUs and closely related species. *Parasites & Vectors*, 10, 488.

SAMAC. (2019) *Statistics on the Southern African macadamia industry*, viewed 29 June 2020, <https://www.samac.org.za/industry-statistics/>

Schoeman, P.S. (2009) *Key biotic components of the indigenous Tortricidae and Heteroptera complexes occurring on macadamia in South Africa*. Unpublished doctoral dissertation, North-West Provence, South Africa, University of Potchefstroom.

Smith, A. (2020) *Characterisation of Lepidoptera species associated with macadamia in South Africa*. Unpublished masters dissertation. Gauteng Provence, South Africa, University of Pretoria.

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

Strüder-Kypke, M.C. & Lynn, D.H. (2010) Comparative analysis of the mitochondrial cytochrome c oxidase subunit I (COI) gene in ciliates (Alveolata, Ciliophora) and evaluation of its suitability as a biodiversity marker. *Systematics and Biodiversity*, 8, 131– 148.

Subtrop. (2018) *Products registered for insect control on macadamias in South Africa*, viewed 23 June 2020, <https://www.samac.org.za/old/wp-content/uploads/2019/01/Macadamias___list_of_insecticides_and_acaricides.pdf>

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J. & Rambaut, A. (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, 4, vey016.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731– 2739.

Tavaré, S. (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences*, 17, 57– 86.

Timm, A.E., Geertsema, H. & Warnich, L. (2006) Analysis of population genetic structure of two closely related tortricid species of economic importance on macadamias and litchis in South Africa. *Agricultural and Forest Entomology*, 8, 113– 119.

Timm, A.E., Warnich, L. & Geertsema, H. (2007) Morphological and molecular identification of economically important Tortricidae (Lepidoptera) on tropical and subtropical fruit in South Africa. *African Entomology*, 15, 269– 286.

Zhao, Z.-H., Cui, B.-Y., Li, Z.-H., Jiang, F., Yang, Q.-Q., Kučerová, Z. et al. (2016) The establishment of species-specific primers for the molecular identification of ten stored-product psocids based on ITS2 rDNA. *Scientific Reports*, 6, 21022.