Austral Entomology

Austral Entomology (2021) **60**, 598–609

DNA barcode identification of *Conopomorpha cramerella* (Snellen, 1904) (Lepidoptera: Gracillariidae) and other moths affecting cacao in Papua New Guinea

David Gopurenko,[1,2]* Peter S Gillespie,[3] Rodney Minana[4] and Olivia L Reynolds[2,5]†

[1]*NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650, Australia.*
[2]*Graham Centre for Agricultural Innovation, Wagga Wagga, NSW 2650, Australia.*
[3]*NSW Department of Primary Industries, Orange Agricultural Institute, Orange, NSW 2800, Australia.*
[4]*Cocoa Board of Papua New Guinea, Tavilo, East New Britain Province Papua New Guinea.*
[5]*Cesar Australia, Parkville, VIC 3052, Australia.*

Abstract

Economically important cacao (*Theobroma cacao* Linnaeus 1753) plantations in South East Asia and Papua New Guinea (PNG) are significantly affected by the cocoa pod borer (CPB) moth. Species identity of the pest is attributed to *Conopomorpha cramerella* (Snellen 1904), a gracillariid moth endemic to Australasian and Oriental tropic regions that has evolved a host preference for introduced cacao. Suspected presence of cryptic CPB biotypes is largely unsupported by earlier genetic work but remains a concern to organisations developing species-specific lure and/or control tools for managing this pest. We report the use of DNA barcoding to investigate population and species genetic diversity of CPB infecting cacao plantations in PNG at the eastern periphery of the pest's distribution. DNA barcodes from 94.4% of 179 moths from three disjunct PNG provinces (East Sepik, Bougainville and East New Britain) matched to reported *C. cramerella* sequence accessions. Genetic diversity among *C. cramerella* in PNG was limited to four closely related haplotypes, two of which were common in PNG and have previously been reported in the Malay Archipelago at different frequencies. We found evidence of significant population genetic structure between mainland and eastern offshore PNG provinces marked by a reduction of genetic diversity at the offshore provinces. We suggest *C. cramerella* populations in these offshore island provinces likely arose recently from a genetically depauperate cohort of founders from western sources. Ten moths were genetically unmatched to *C. cramerella*. Of these, six incidentally captured and degraded adults had novel unmatched DNA barcodes. Four larvae infesting cacao pods were genetically matched to *Thaumatotibia zophophanes* (Turner 1946) (Lepidoptera: Tortricidae) and an undefined *Conopomorpha* species distantly related to *C. cramerella*. The extent to which these two additional moth species affect the cacao industry in PNG (and potentially elsewhere) remains unknown. However, their low representation here (<2.3% of samples) indicates they may be of lesser prominence than the widespread *C. cramerella*. Nevertheless, organisations designing and/or implementing CPB controls tailored specifically for *C. cramerella* will need to ensure empirical evidence of control efficiency is monitored with respect to species identity of the captured pests.

Key words cocoa pod borer (CPB), cytochrome *c* oxidase subunit I (COI) gene, Papua New Guinea, *Theobroma cacao*.

INTRODUCTION

Plantation production of cacao (*Theobroma cacao* Linnaeus (1753) (Malvaceae)) is of major economic importance to Papua New Guinea (PNG) and other countries of the Malay Archipelago, which collectively contribute approximately 14% of annual global cacao production (FAO 2004). Damage to cacao fruit caused by oviposited larvae of the 'cocoa pod borer' (CPB) pest moth has significantly affected cacao production across much of the Malay Archipelago. Losses of cacao yield as high as 80% due to CPB damage have been reported (Loke *et al*. 2006), and in some instances, CPB infestations have forced temporary abandonment of orchards and plantations (Day 1989).

Taxonomic (Bradley 1986) and genetic (Shapiro *et al*. 2008) authorities have attributed identity of the CPB pest to the gracillariid moth *Conopomorpha cramerella* (Snellen 1904) present at various tropic locations in the Orient and Australasia. Apart from introduced cacao, *C. cramerella* is associated with at least 10 native host plant species, mainly within Sapindaceae (De Prins & De Prins 2019) including economically important rambutan (*Nephelium lappaceum* Linnaeus 1767). *Conopomorpha cramerella* show a strong preference for oviposition on introduced cacao over its native plant hosts (Gende 2012; Niogret *et al*. 2020). This is surprising given that cacao was introduced during the 16th century to South-East Asia from the new world (Day 1985) where *Conopomorpha* is absent (De Prins & De Prins 2019). Pest Lepidoptera damage to cacao pods was first

*david.gopurenko@dpi.nsw.gov.au
Declaration of conflict of interest: The authors declare there are no conflicts of interest.
†Present address: Susentom, Melbourne, VIC 3081, Australia.

© 2021 Australian Entomological Society
doi: 10.1111/aen.12559

noted during the mid-19th century among the early Indonesian cacao plantations in central Java and Sulawesi (Mumford 1986). Substantial damage to pods reported later at plantations in the Philippines from 1936 (Uichanco 1936), in East Malaysia from 1980 (Day 1989) and in PNG from 2006 (Loke *et al*. 2006) indicated the broad geographical scale of the problem. A scant record of *C. cramerella* in PNG prior to 2006 indicated presence of its larvae found in fruit of some Fabaceae (Froggatt 1940). It was argued that it had not been recorded in the country in over a decade since 1940 and therefore was of little concern to the local cacao industry of the day (Dunn 1955). The infestations in PNG during 2006 occurred initially in the offshore East New Britain Province (Loke *et al*. 2006), and within a year, CPB had affected multiple mainland and offshore plantations separated by up to 1300 km (Gende 2012).

It is uncertain how this pest had rapidly emerged across distant and disjunct cacao plantations across PNG and elsewhere. Adult CPB in a host orchard fly up to 5 m at a time when disturbed (Niogret *et al*. 2019), but their capacity for migratory flight or wind assisted dispersal between host plant stands over far greater distances is unreported. The natal distribution of *C. cramerella* prior to the advent of the cacao industry is uncertain. Subsequently, it is unclear if the species was naturally endemic across affected areas of the Malay Archipelago, or if expansion of the cacao industry (and/or other cash crop hosts) facilitated movement of the pest into areas where it was absent. Alternatively, inter-plantation transfers of cacao planting materials contaminated by CPB eggs, larvae and/or pupae may have provided a recent pathway for assisted dispersal and introduction of this pest among distant cacao plantations (Singh & Rethinam 2006; Yen *et al*. 2010; Gende 2012). This was indirectly supported by population genetic analysis of *C. cramerella* sampled across the Malay Archipelago (Shapiro *et al*. 2008). Evidence of diminished multi-locus genetic diversity across much of the pest's distribution resembled that seen in populations of introduced insects that have experienced bottleneck events associated with their movement into new territory (Shapiro *et al*. 2008).

Shapiro *et al*. (2008) also genetically tested for cryptic biotype and/or species presence in CPB, as this was an issue initially raised in reports of allozyme variation among putative host-specific biotypes of *C. cramerella* in Malaysia (Rusnah *et al*. 1985; Muhamad & Tan 1987). The possibility of cryptic diversity within CPB remains a point of concern to the cacao industry (Singh & Rethinam 2006; Yen *et al*. 2010) and relevant to organisations using and/or developing species-specific lure and control tools and practices for management of the pest. Undetected taxonomic diversity present in the distribution of the pest CPB may potentially confound generically applied trapping and controls of the pest (Beevor *et al*. 1993; Zhang *et al*. 2008). The earlier suggestions of cryptic species diversity among CPB were unsubstantiated by Shapiro *et al*. (2008) who reported evidence only of intraspecific level variations among samples with no clear associations between CPB genetic lineages and their host use. Notwithstanding the broad geographic scale of their survey across the Malay Archipelago (~5000 km), the per-location sampling employed by Shapiro *et al*. (2008) was low (average 6.1 samples over 15 locations) and in some cases obtained from historical collections of adult specimens found on various hosts. It is possible that a more intensive genetic sample of Lepidoptera directly affecting cacao pods may reveal presence of previously undetected species diversity.

DNA barcode (Hebert *et al*. 2003) analysis of mitochondrial encoded 5′ cytochrome *c* oxidase subunit I (COI) gene sequence variation within and between faunal species provides a standardised comparative means to test species hypotheses and to aid in discovery of novel species diversity (Gopurenko *et al*. 2015). The method allows genetic identification of query specimens to species when matched to representative DNA barcodes of taxonomically vouchered taxa (Mitchell & Gopurenko 2016). This is particularly useful in cases when query specimens are morphologically ambiguous due to their physical condition and/or developmental life stage (Gopurenko *et al*. 2013). Where levels of intraspecific polymorphism are sufficient, DNA barcodes can also be informative of maternal population genetic structure in a broadly sampled target species (Bellis *et al*. 2015). Here, we report use of DNA barcoding to investigate species and population genetic diversity of CPB infesting cacao plantations in three geographically separated PNG provinces at the eastern periphery of the pest's distribution. For comparative genetic analyses, we included COI DNA barcode sequences available at repositories for *Conopomorpha* species, and six *C. cramerella* COI haplotypes reported in earlier genetic analysis of CPB (Shapiro *et al*. 2008).

MATERIALS AND METHODS

Sampling locations

Lepidoptera were sampled during February to June 2019 from cacao plantations in three PNG provinces: Autonomous Region of Bougainville (Bougainville), East New Britain (first outbreak of CPB in PNG recorded 2006) and East Sepik (Fig. 1; Supporting Information S1). The three provinces are primary centres of cacao production in PNG and are each sea separated by distances exceeding 200 km, with greatest straight-line distance of 1350 km between East Sepik (west mainland PNG) and Bougainville to the east. Sampled locations within provinces were separated by no more than 150 km and collectively analysed per province.

Four methods of sampling were used at each province (summarised from detailed report in Chapter 5, Reynolds *et al*. 2019):

1 Sampling infested cacao pods
C. cramerella infested cacao pods (displaying typical uneven ripening) were harvested from trees, split open at field site laboratories and inspected for the presence of an insect or insect damage. Specimens were collected as larvae in pods, pupae on pods and adults raised from larvae infected pods.

2 Leaf inspection
Leaf litter and leaves on host and non-host plants in cacao plantations were visually inspected for the presence of *C. cramerella* pupae.

© 2021 Australian Entomological Society

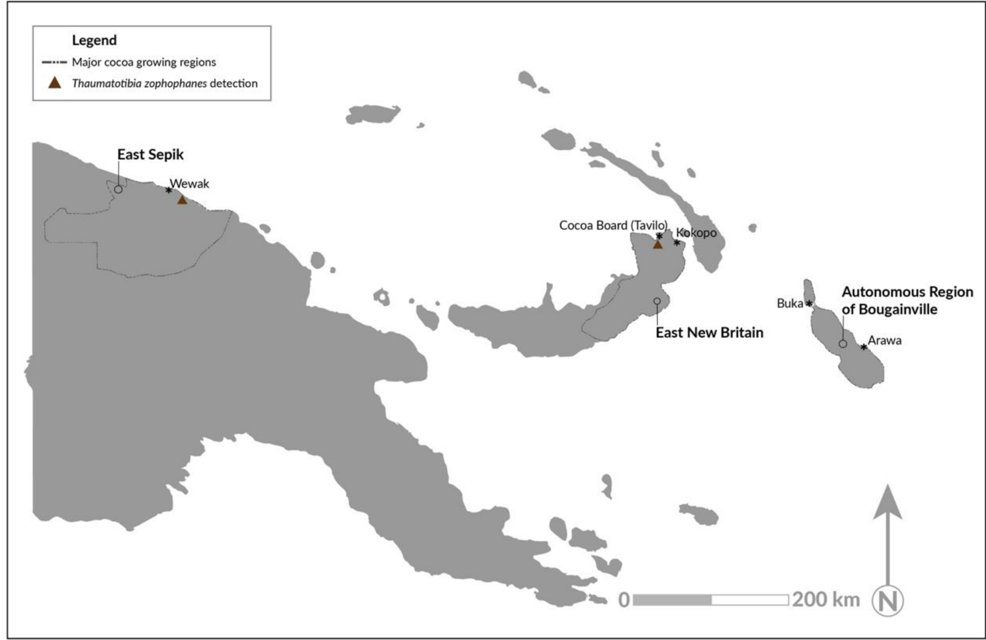

Fig 1. Map of Papua New Guinea showing three major cocoa growing provinces, East New Britain, East Sepik and the Autonomous Region of Bougainville, that were sampled for *Conopomorpha cramerella* and other cocoa pod borer pests, including the first records of *Thaumatotibia zophophanes* sampled from cocoa.

3 Pheromone trapping
Two types of traps were used, a modified Biotrap that had four 2-cm-diameter holes cut out of the clear plastic lids (BioTrap Australia Pty Ltd, Victoria, Australia) and a UNI-Trap (AlphaTrap, Oregon, USA). These were baited with *C. cramerella* lures (PCI Pest Control Private Limited, Mumbai, Maharashtra) and contained a 1 cm^2 of dichlorvos-impregnated strip. Traps were placed randomly (singly) in fruiting cacao trees (all provinces), fruiting rambutan trees (East Sepik only), native forest (ENB only) and native rainforest (East Sepik only), for at least 24 h and up to 14 days. Traps were checked daily or weekly/fortnightly for adult moths.

4 Sweep netting
Regular sweeping of leaves around and within the cacao canopy yielded some adult moths. This activity was typically 5 min in duration in places where traps were deployed, or where fruit was collected but was limited in some instances by the degree to which the cacao canopy was closed over thus restricting movement.

Curation of samples

Initial specimen handling and rearing was done at PNG field site laboratories. Immature stages were preserved in 99% ethanol, while adults were either micro-pinned and dried or, if degraded, placed into 99% ethanol. Procedures for taxonomic identification of specimens were detailed in Reynolds *et al*. 2019 (Chapter 5). Preserved and pinned specimens were received and curated at the Biological Collections unit in NSW DPI (Orange Agricultural Institute) and allocated with unique alpha-numeric sample tags before subsequent DNA barcoding at Wagga Wagga Agricultural Institute (Supporting Information S1).

DNA barcoding

Specimens were cleared of ethanol by overnight evaporation prior to digestion. Samples were non-destructively digested overnight at 55°C in 330 μL of aliquots containing 310 μL of buffer ATL (with 100% ethanol additive) and 20 μL of Proteinase K (600 mAU/mL) as provided in DNeasy®, Blood & Tissue Kit (Cat ID: 69506; QIAGEN). DNA was extracted from 250 μL of each digestion and eluted to 170 μL, using volume adjusted buffers and protocols provided in GenElute™ 96 well tissue genomic DNA extraction kit (G1N9604; Sigma-Aldrich).

For PCR amplification, we targeted the 5′ mitochondrial cytochrome *c* oxidase I (COI) gene fully encompassing the standard faunal DNA barcode region (Hebert *et al*. 2004) and all contiguous downstream sequence examined by Shapiro *et al*. (2008). For this, we used primers *LepF1* (Hebert *et al*. 2004) and *C1-N-2410D* (5′ AATTTTAATwCCwGTwGG) designed here as a modified and reverse complement of primer C1-J-2441 (Simon *et al*. 1994), in amplification of an 895 base-pair (bp) target. In a few instances, this primer combination failed, and we repeated PCR but replaced *C1-N-2410D* either with *C1-N-2329* (Simon *et al*. 1994) to amplify a 814 bp target or with *LepR1* (Hebert *et al*. 2004) to amplify just the DNA barcode region (658 bp). All primers were 5′ tailed with directional M13 vector sequences (17 mer) to simplify bi-directional Sanger sequencing at the Australian Genome Research Facility (AGRF, Brisbane node). PCR reactions (15 μL of total volume) contained 2 μL of DNA extract, 1 μM each of forward and reverse primers (Sigma-Aldrich), 200 μM of dNTP mix (Invitrogen) and Invitrogen PCR reagents (0.4 U Platinum® *Taq* DNA Polymerase, 2.9 mM of $MgCl_2$ and 1 X PCR buffer). Thermal cycling using an Eppendorf Mastercycler EP Gradient

© 2021 Australian Entomological Society

S Thermal Cycler consisted of a 94°C denature step (2 min), followed by 40 cycles of step profile (94°C denature 30 s, 50°C anneal 30 s, 72°C extension 45 s) and ending with a 72°C extension (5 min) and 4°C storage. PCR products stained with Invitrogen SYBR® Safe DNA Gel Stain were visualised under UV light in a BioRad Universal Hood II following electrophoresis through a 1.5% agarose gel including of size markers and negative controls. PCR products were Sanger sequenced at the Brisbane node of the Australian Genome Research Facility (AGRF). Semi-automated liquid handling protocols for PCR preparation and sequence picking followed Gopurenko *et al*. (2013).

Bi-directional sequence chromatograms were consensus assembled to sample ID and quality checked using Lasergene SeqMan Pro ver. 8.1.0 (DNASTAR Inc., Madison, WI, USA). Assembly was orientated for 5′ directional alignment against reference *C. cramerella* COI sequence (895 nucleotide combination of overlapping GenBank accessions HQ824804 and EU644524). Aligned FASTA formatted sequences were primer truncated using BioEdit ver. 7.0.5.3 (Hall 1999) and examined for evidence of frame shift indels and heterozygous sites indicative of nuclear encoded mitochondrial pseudogenes. Specimen records and their associated DNA barcode sequences were publicly released to the Barcode of Life Data Systems [BOLD] (Ratnasingham & Hebert 2007) repository under project 'Cocoa pod borer in Papua New Guinea' (project code: CPBNG) and accessioned at GenBank (MW378131–MW378310). For general specimen summary, refer Supporting Information S1.

Exemplar sequences in the alignment were listed using FABOX ver. 1.35 (Villesen 2007) and queried (05/May/2021) for closest genetic match to publicly available COI accessions using the online BOLD identification engine (http://v4.boldsystems.org/index.php/IDS_OpenIdEngine) (Ratnasingham & Hebert 2007). Sequences matched to a reported accession (>99% sequence similarity criterion) were treated as putatively conspecific and identified to the accession's taxonomic record. Sequences unmatched to accessions in a genus (>3% sequence difference) were treated as unidentified; in those cases, the taxonomic description of the nearest matched accession was cautiously noted as the closest comparative reference to the query. As an alternative to this genetic distance approach, we also reported the Barcode Index Numbers (BINs) and associated taxonomic descriptions auto assigned to our specimen sequences deposited at BOLD. BINs provide alpha numeric indexing of genetically similar COI barcode sequences grouped by coalescence algorithms, to allow ongoing delimitation of molecular operational taxonomic units in the BOLD database (Ratnasingham & Hebert 2013). BINs reported in Lepidoptera studies are frequently inclusive of specimens with a shared single species description (Wilson *et al*. 2011; Hausmann *et al*. 2013; Lees *et al*. 2013; Zahiri *et al*. 2014). We also used assigned BIN identifiers as interim species hypotheses for indexing query sequences either unmatched to existing accessions in the BOLD database or only matched to accessions lacking a complete taxonomic description.

Pairwise genetic distance relationships among sequences minimally identified to genus, and their matched exemplar accessions at BOLD, were constructed by Neighbour-Joining (NJ) distance analysis implemented in MEGA 6.06 (Tamura *et al*. 2013). The NJ analysis included all publicly available sequence exemplars of *Conopomorpha*, to compare genetic distances among taxa within the genus (Supporting Information S2). Pairwise sequence distances were calculated using a Kimura two-parameter substitution model to accommodate transition–transversion nucleotide rate variation. Missing nucleotides in the alignment were excluded from paired sample distances. Significance of nodes in the optimal NJ tree were determined by bootstrap replication (N = 10 000 replicates).

Sequences identified to *C. cramerella* were truncated to 624 bp to allow equal length comparisons against six available *C. cramerella* haplotypes labelled CO-A to CO-F (Shapiro *et al*. 2008; GenBank accessions EU644563, EU644578, EU644565, EU644539, EU644577 and EU644589). Novel haplotypes in PNG not matched at 100% similarity to the six haplotypes were labelled following on from earlier nomenclature (Shapiro *et al*. 2008) and deposited at GenBank (accessions MW403484 and MW403485). Genealogical relationships among *C. cramerella* haplotypes were inferred as a parsimony network using TCS ver. 1.21 (Clement *et al*. 2000) set to a 95% limit imposed on haplotype connections.

Summary population genetic statistics reporting haplotype (h) and nucleotide (π) diversity among provincial *C. cramerella* populations were estimated from the 624 bp haplotype alignment using DnaSP ver. 6.0 (Rozas *et al*. 2017). DnaSP was used to ascertain Tajima's D statistic and associated significance tests for evidence of deviation from nucleotide neutrality among identified *C. cramerella* sequences. Pairwise F_{st} estimates of population genetic structure between provinces based on nucleotide data were determined using DnaSP, with significance of estimates obtained by permutation (10 000 replicates).

RESULTS

Sampling of cacao plantations in PNG provided 189 moths for DNA barcode analyses and included 32 specimens from East Sepik, 99 from East New Britain and 58 from Bougainville. Most specimens were obtained as larvae in cacao pods (N = 149), and of these many (N = 89) were raised to adult stage. The remainder of specimens consisted of pupae sampled from the cacao pod husk (N = 7) and cacao leaf (N = 9), or adults captured using lure-baited traps (N = 21). Sweep netting was ineffective as a means for capture of adult CPB moths, although other incidental species were captured by this means and retained for molecular comparison against sampled larvae (see later).

COI sequences were obtained from 179 of 189 sampled specimens (Supporting Information S1); 10 specimens failed to PCR amplify despite repeats using various primer combinations. Aligned sequences were free of frameshift indels and/or degenerate nucleotide sites found in pseudogenes. Over 94% of primer truncated sequences were of 814 to 895 nucleotides in length, and all sequences fully overlapped the DNA barcode region (658 bp) except at specimen ww26793 from which only 455

© 2021 Australian Entomological Society

bp of DNA barcode region sequence was obtained. Nucleotide polymorphisms were confirmed by bi-directional sequence reads and by additional selective resequencing effort.

Sequences of 169 PNG specimens matched closest (>99.84% similarity) to existing DNA barcodes of *Conopomorpha cramerella* at BOLD (N = 106 public records, allocated to BIN BOLD:AAA4000). Two existing and two novel *C. cramerella* haplotypes were identified among 163 of the sequences that could be truncated to 624 bp for equal length comparisons against the six haplotypes 'CO-A' to 'CO-F' reported by Shapiro *et al*. (2008). A 95% parsimony network among all equal length *C. cramerella* haplotypes evidenced a genetically shallow multifurcating phylogeny with maximum sequence differences not exceeding 0.32% (Fig. 2a). Haplotype CO-A was central to all haplotypes in the network and identified as the most recent common ancestral lineage in the species. Each of the remaining seven haplotypes differed from CO-A by a single nucleotide mutation. CO-A was present in each of the three sampled PNG provinces but represented less than 8.6% of the current *C. cramerella* sample (Fig. 2b). Most *C. cramerella* in PNG (>88.3%) had haplotype CO-B, and this was dominant in each sampled province. The relative frequencies of haplotypes CO-A and CO-B in the current PNG sample were a reversal of that reported collectively among Malay Archipelago locations to the west of PNG (Fig. 2b) where previously haplotypes CO-A and CO-B were found in 80.2% and 11.1% of moths, respectively (Shapiro *et al*. 2008). Four rare haplotypes (CO-C, CO-D, CO-E and CO-F) reported to the west of PNG (Shapiro *et al*. 2008) were absent from our PNG sample. Two novel and rare haplotypes (CO-G and CO-H; GenBank: MW403484 and MW403485, respectively) in our PNG sample were only observed at the mainland East Sepik Province. Haplotype and nucleotide diversity measures at East Sepik were more than twice as high as that at Bougainville and East New Britain, with the least genetic diversity observed at East New Britain (Table 1). Significantly high population genetic structure was observed between East Sepik and Bougainville (F_{st} = 0.144, $P < 0.01$) and between East Sepik and East New Britain (F_{st} = 0.234, $P < 0.001$). In contrast, population genetic structure between Bougainville and East New Britain was minimal (F_{st} = 0.022) and non-significant ($P > 0.05$) and reflected the almost identical haplotype frequency profiles present at these two offshore provinces (Fig. 2b). Tajima's D measure of deviation from a neutral model of nucleotide evolution at COI was negatively deviated ($D = -0.817$) for the collective PNG sample and individually at Bougainville ($D = -0.178$) and East New Britain ($D = -0.783$). In contrast, the East Sepik was positively deviated ($D = 0.153$) from neutrality expectations. Deviations of the observed D statistic in each of these four tests were not significantly different ($P > 0.10$) from neutral model expectations.

Ten remaining moths were genetically unrelated to *C. cramerella*, differing by more than 10% and far exceeding the maximum 0.41% sequence difference evidenced among vouchered accessions of this species (Table 2). These unrelated moths included four larvae infesting cacao fruit in East New Britain and East Sepik, and six adult moths captured mainly by sweep netting in Bougainville and East New Britain plantations. The DNA barcode of larval specimen ww27790 found in a cacao pod at East New Britain differed from *C. cramerella* by >11.1% but was matched at 99.1% similarity to accessions of an unidentified *Conopomorpha* species (indexed as BOLD:AAI5875;

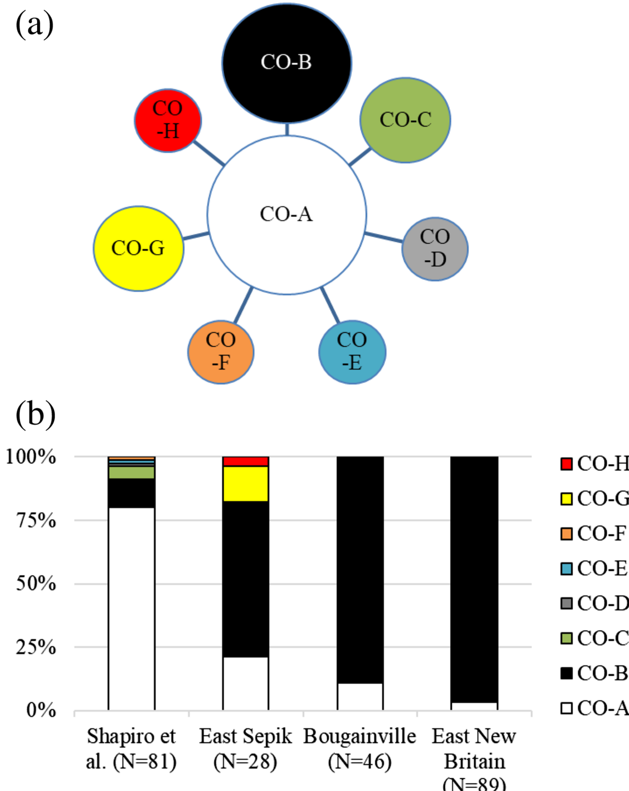

Fig 2. *Conopomorpha cramerella* COI haplotype relatedness and distribution. (a) Genetic relationships among eight haplotypes (CO-A to CO-H) determined as a 95% parsimony network. Connection between haplotypes represents one nucleotide difference in 624 base sequence alignment. Decreasing circle size represents frequency of haplotype as high, moderate, low and singleton. (b) Relative frequency of haplotypes (coloured as in (a)) at geographic regions including Malay Archipelago west of Papua New Guinea (* data from Shapiro *et al*. 2008), and three provinces (East Sepik, Bougainville and East New Britain) in Papua New Guinea sampled in this study.

Table 1 COI haplotypes (624 bp) identified to *Conopomorpha cramerella* sampled from three provinces in Papua New Guinea (PNG)

	Province			Total PNG
	East Sepik	Bougainville	East New Britain	
$N_{sequences}$	28	46	89	163
$N_{haplotypes}$	4	2	2	4
H	0.585	0.198	0.066	0.213
π	0.00119	0.00028	0.00009	0.00038
Taj. D	0.153^{NS}	-0.178^{NS}	-0.783^{NS}	-0.817^{NS}

The number (N) of sequences and haplotypes, and measures of genetic diversity reported, including haplotype (H) and nucleotide (π) diversity. Tajima's D (Taj. D) statistic and its probability of (±) deviation from expectations under a neutral model of evolution here shown as not significant (^{NS}) at $P > 0.10$.

© 2021 Australian Entomological Society

Table 2 DNA barcodes of 10 non-*Conopomorpha cramerella* specimens sampled from Papua New Guinea cacao plantations

ID	Stage	BIN	BOLD nearest matched record*		%
			Family	Species	
ww26780	A	ADZ3683	Riodinidae	*Semomesia*	90.1
ww26787	A	AEB9394	Cosmopterigidae	*Labdia*	91.3
ww26791	A	ADZ4996	Cosmopterigidae	*Limnaecia*	93.0
ww26793	A	N/C	Cosmopterigidae		95.0
ww27760	A	ADZ5450	Gracillariidae	*Stomphastis*	96.8
ww27790	L	AAI5875	Gracillariidae	*Conopomorpha* sp.	99.2
ww27741	A	ADZ5878	Noctuidae	N/A	91.9
ww27757	L	AAD6549	Tortricidae	*Thaumatotibia zophophanes*	100
ww27752	L	AAD6549	Tortricidae	*Thaumatotibia zophophanes*	99.2
ww26734	L	AAD6549	Tortricidae	*Thaumatotibia zophophanes*	100

DNA barcodes queried at BOLD; taxonomic information and sequence similarity (%) to nearest genetic match as indicated. Specimen sample ID (refer Supporting Information S1), developmental stage (A = adult; L = larva), and Barcode Index Number (BIN, except where non-compliant (N/C)) as registered at BOLD. Per cent (%) sequence similarity of specimen DNA barcode to its nearest matched public record at the BOLD identification engine. Identification to species* is only indicated where matched sequence similarity > 99%, otherwise closest matched family and genus indicated except where not available (N/A).

Fig. 3; Table 2). This unidentified *Conopomorpha* species was marginally more similar (92.3% sequence similarity) to *C. litchiella* Bradley 1986 than to *C. cramerella* (88.9%). Three other specimens sampled from cacao fruit at East New Britain and East Sepik comprised haplotypes matched (99.5–100% similarity) to accessions of the tortricid moth *Thaumatotibia zophophanes* (Turner 1946) (Lepidoptera: Tortricidae) indexed to BOLD:AAD6549 (Fig. 3; Table 2). Six unrelated adult moths captured by sweep netting ($N = 4$) or in a pheromone trap ($N = 2$) were each uniquely indexed at BOLD and were ambiguously related (<97% sequence similarity) to various Lepidoptera (Table 2). Morphological examination of these specimens was not possible owing to their damaged and/or degraded condition Reynolds *et al*. 2019). These six adult moths were captured in the cacao plantations, but there was no evidence linking them to crop damage.

DISCUSSION

DNA barcoding confirmed *Conopomorpha cramerella* was by far the most prevalent (94.4%) species in our sample of 179 sequenced moths from cacao plantations in three PNG provinces. Additional captures represented less than 6% of the sample and included at least two other pest taxa infesting cacao pods and several incidental captures. Presence of the additional pest taxa has relevance for CPB pest management.

Genetic paucity of *Conopomorpha cramerella*, the main pod-boring pest in Papua New Guinea

Our DNA barcode analysis indicated low levels of mitochondrial genetic diversity among 169 *C. cramerella* in PNG at the eastern periphery of the species distribution. The maximum pairwise sequence difference among PNG haplotypes did not exceed 0.32%. This was consistent with earlier evidence of diminished mitochondrial genetic diversity identified among a lesser sample of *C. cramerella* (N = 92) collected across >5000 km of the Malay Archipelago (Shapiro *et al*. 2008). Four of the eight comparable mitochondrial COI haplotypes currently identified to *C. cramerella* were present in PNG (Fig. 2a). Two haplotypes (CO-A and CO-B) common to all three provinces in our PNG sample were reported as widespread in the Malay Archipelago (Shapiro *et al*. 2008), albeit at contrasting relative frequencies to that found in the current sample (Fig. 2b). Two novel haplotypes (CO-G and CO-H) were present at low frequencies (collectively 3.1% of the current sample) and unique to the East Sepik Province of mainland PNG. Four other rare haplotypes reported west of PNG (Shapiro *et al*. 2008) were not observed in the current sample.

The 95% parsimony network among all *C. cramerella* haplotypes was characteristic of a low complexity intraspecific 'star' phylogeny often observed in species that have undergone historically recent population expansion from a common ancestor (Slatkin & Hudson 1991). Haplotype CO-A was genetically equidistant to all other haplotypes in the *C. cramerella* network and identified as the most recent common ancestral mitochondrial lineage for the species (Fig. 2a). The common presence of this ancestor haplotype across the species distribution in the Malay Archipelago was consistent with intraspecific coalescent theory predictions concerning the distribution of the progenitor lineage (Posada & Crandall 2001). Although our sample of *C. cramerella* in PNG was almost twice the quantity sampled across the Malay Archipelago by Shapiro *et al*. (2008), we did not detect additional genetic depth in the haplotype network of mitochondrial lineages reported earlier. Based on this shallow level of coalescence, we suggest the eastern peripheral *C. cramerella* populations in PNG were unlikely natal or ancestral sources of diversity within the species distribution.

Our analysis indicated presence of significantly high population genetic structure between the mainland population of *C. cramerella* at East Sepik and each of the two offshore provinces ($F_{st} > 0.144$, $P < 0.01$). In contrast, the measure of population structure between the two offshore provinces (Bougainville and East New Britain) was lower and

© 2021 Australian Entomological Society

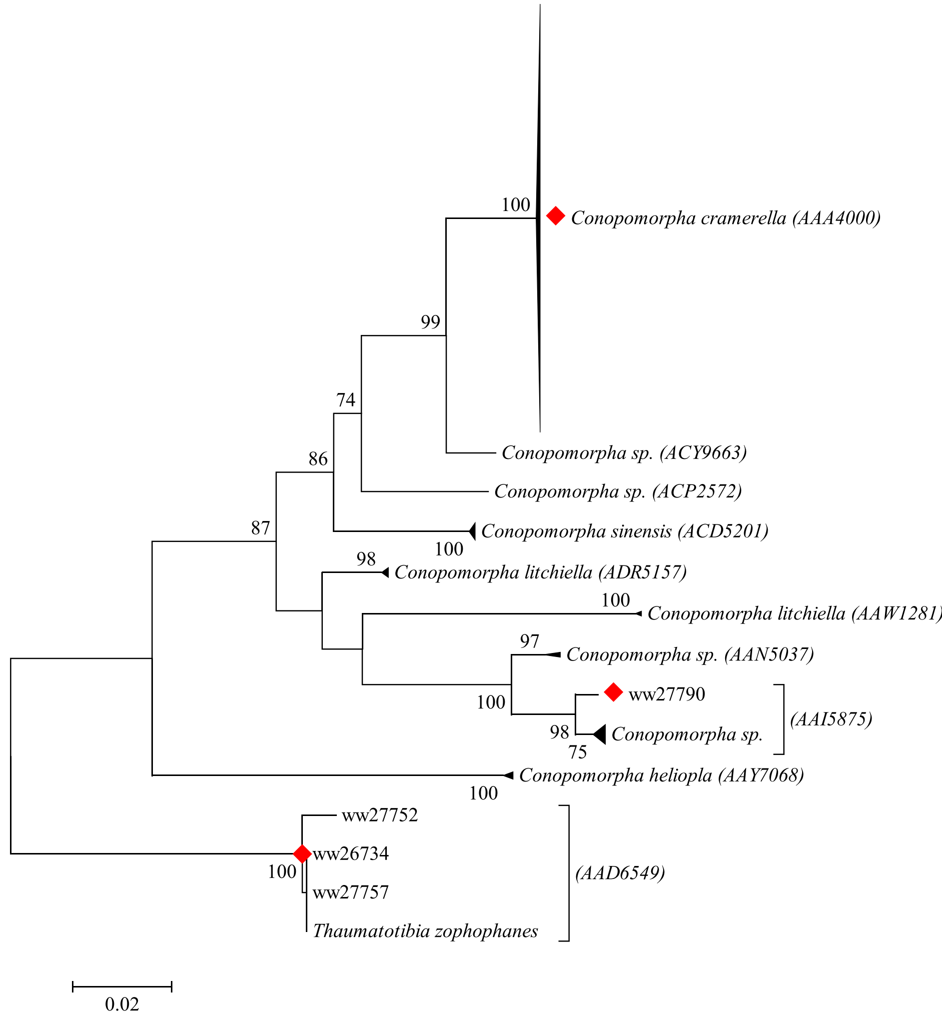

Fig 3. Neighbour-Joining distance tree among DNA barcodes of moths sampled from cocoa pods in Papua New Guinea, and all available *Conopomorpha* accessions at BOLD. Current samples denoted by closed red diamond and specimen labels as listed in Supporting Information S2. Barcode Index Numbers in parentheses as assigned by BOLD to terminal DNA barcode clusters (collapsed to species). DNA barcodes matched at >99% sequence similarity to accessions using the BOLD identification engine, unmatched DNA barcodes not included here. Bootstrapped (10 000 replicates) node supports (>70%) as shown. Scale bar equals 2% Kimura two-parameter adjusted sequence distance.

non-significant ($F_{st} = 0.022$, $P > 0.05$), suggesting the two offshore provinces were demographically connected or recently sourced from a shared founder cohort. Comparative statistical tests of population structure across the Malay Archipelago were unreported in the earlier analysis of *C. cramerella* (Shapiro *et al*. 2008) possibly due to the low per-site sample replication in that study ($N_{average} = 6.1$, sampled over 15 countries). Regardless, Shapiro *et al*. (2008) indicated CPB haplotypes among sea-separated regions across a 5000 km zone were unsorted by any obvious pattern or discernible population structure. In contrast, our evidence indicated an interesting phylogeographic shift in the frequencies of the two common haplotypes at PNG relative to the remainder of the Malay Archipelago. The most recent common ancestor haplotype (CO-A) of *C. cramerella* reported as abundant and ubiquitous among sampled locations west of PNG (collectively 80.2% of 81 specimens, Shapiro *et al*. 2008) occurred at much lower frequency in our PNG sample (collectively 8.6% of 163 PNG specimens). The abundance of haplotype CO-A at each province in PNG was greatly surpassed by the derived haplotype CO-B (88.3% of all PNG *C. cramerella* specimens) reported earlier at lower frequencies (11% of 81 specimens) among disjunct locations west of PNG. Although not commented on directly by Shapiro *et al*. (2008), their total sample of 11 *C. cramerella* from PNG (all from East New Britain) was fixed for haplotype CO-B (refer Table 1, Shapiro *et al*. 2008) and consistent with our evidence of an upwards shift in the frequency of this haplotype in PNG. Shapiro *et al*. (2008) sampled too few specimens from Irian Jaya ($N = 6$), to the west of PNG, to permit an analysis of the broader regional scale of this pattern across the New Guinea landscape. Future intensive sampling is required to determine if the pattern is ubiquitous across *C. cramerella* populations elsewhere in New Guinea, or if the relative abundance of these two common haplotypes transition as a clinal or abrupt frequency shift in the landscape.

© 2021 Australian Entomological Society

We speculate that stochastic sampling of mitochondrial lineages during *C. cramerella* founder events led to fortuitous prevalence of the derived haplotype (CO-B) over the ancestral haplotype CO-A in PNG. Founder and invasive events involving reduction(s) of effective population size can accentuate stochastic losses and sorting of alleles in newly established populations (Puillandre *et al*. 2008). Our evidence of low haplotype diversity ($H = 0.066$ and 0.198) and lack of population structure between the two eastern offshore provinces is consistent with our contention that these eastern provinces were foundered by a limited and shared cohort of *C. cramerella*. Presence of *C. cramerella* in PNG prior to entry of the cacao industry is unknown; however, expansion of the industry onwards from 2006 may have inadvertently aided both broad dispersal and establishment of a common cohort of *C. cramerella* among plantations in the country. Cacao planting materials are commonly transported among PNG plantations, and survey evidence suggests relaxed quarantine and bio-sanitary handling procedures were prevalent at the onset of the industry in PNG and potentially allowed movement of CPB infected materials among disjunct plantations and distant grower provinces (Yen *et al*. 2010; Gende 2012). Stochastic sampling of maternal genetic lineages during successive translocation events can allow rare haplotypes to rise up at high frequencies in new areas (Puillandre *et al*. 2008) and may account for the observed shifts in relative frequencies of haplotypes CO-A and CO-B in PNG. Similar shifts in neutral allelic diversity incurred by genetic drift during population bottleneck and/or founder events have been evidenced at isolated peripheral populations (Gopurenko & Hughes 2002), at areas of recent natural range expansion (Bellis *et al*. 2015) and at the populations of introduced species (Zhu *et al*. 2019).

Analyses of our genetic evidence under models of population equilibrium were equivocally supportive of demographic or bottleneck events affecting genetic diversity in PNG. Marginally negative neutrality statistics (Tajima's *D*; Table 1) observed at each of the two offshore provinces indicated low-frequency COI polymorphisms were potentially in excess of that expected if the provinces had sustained stable effective population sizes or were free from the effects of purifying selection linked to the examined locus. Shapiro *et al*. (2008) reported similar 'borderline' negative deviances from neutrality at unlinked mitochondrial and nuclear loci in their pooled sample from Malay Archipelago locations. They argued this apparent shared pattern at two independent genomes was symptomatic of a demographic bottleneck event affecting both loci, and unlikely to have occurred through independent selective sweeps for favoured alleles at each locus. We note however in all cases the observed marginal negative neutrality statistics reported by both the current and earlier studies were not significantly different from expectations under a stable population model. It is possible the low levels of available polymorphism at the examined loci affected statistical sensitivity of these neutrality tests (Pluzhnikov & Donnelly 1996). Future population genetic analysis of this species may benefit from analysis of multiple single nucleotide polymorphisms and/or hypervariable loci, where greater levels of assayable nucleotide diversity could allow more informative testing of this neutrality question. This approach has been reported for other pest Lepidoptera, where next-generation sequencing approaches have provided unparalleled population genetic resolution of taxonomic lineages defined previously by mitochondrial DNA barcoding (Perry *et al*. 2018, 2020; Picq *et al*. 2018).

CPB diversity: more than the one species

Reports of moth larvae affecting cacao pods in PNG prior to emergence of CPB infestations during 2006 identified the tortricids *Cydia* Hübner (1825) sp. and *Cryptophlebia encarpa* (Meyrick 1920) as pests of limited occurrence and minor economic concern to the industry (Froggatt 1938; Froggatt 1940; Baker 1976). *Conopomorpha cramerella* and several suspected pod borer moths were similarly reported to be of minor concern (Dunn 1955; Szent-Ivany 1961). First genetic evidence is presented here for two additional moth species directly damaging cacao pods in PNG cacao plantations, albeit at low levels of abundance, and may be significant for future moth pest research and management. This evidence heralds the need for vigilant attention to species diagnostics when interpreting empirical CPB data. Larvae of the tortricid, *T. zophophanes* and an unidentified *Conopomorpha* species found damaging cacao pods represented 2.3% of our DNA barcoded moths. *Thaumatotibia zophophanes* was sampled at mainland East Sepik and offshore East New Britain, indicating the pest may be broadly distributed across much of the PNG cacao plantations. The pest was previously recorded from North Queensland damaging fruit of native lemon Aspen *Acronychia acidula* von Mueller 1864 (Rutaceae) and introduced avocado *Persea americana* Miller 1768 (Lauraceae) and macadamia *Macadamia* sp. von Mueller 1857 (Proteaceae) (Horak & Komai 2016) but has not been recorded as causing damage to the sparse cacao plantations in that region. The other novel cacao pest species was sampled as a single specimen in East New Britain; it was genetically matched (BIN BOLD: AAI5875) to reported adult accessions of an unidentified *Conopomorpha* species sampled previously only from Madang Province (mainland PNG) on Sapindaceae host plant *Pometia pinnata* Forster & Forster 1776 (Sam *et al*. 2017). This unspecified *Conopomorpha* species is genetically distant from others in the genus and placed closer to *C. litchiella* than to *C. cramerella* (Fig. 3).

Uncertainty concerning species identity and diversity of pod borer moths has persisted for several decades. This uncertainty was driven initially by taxonomic confusion in the literature of the main pest *C. cramerella* with several other *Conopomorpha* species and to some extent by taxonomic inaccuracy of field diagnosed CPB. Bradley (1986) provided an interim taxonomic report clarifying descriptions of *C. cramerella* and three novel *Conopomorpha* species existing as a complex of four taxa indigenous to Oriental and Australasian locations that are '*remarkably similar in wing pattern and coloration*'. Bradley identified *C. cramerella* as the sole pest of cacao within the complex and distinguished it from *C. litchiella*, an economically important pest of litchi (*Litchi chinensis* Sonnerat 1782 [Sapindaceae]), that was previously misreported as *C. cramerella* following incorrect listing of the latter species in India by Fletcher (1916).

© 2021 Australian Entomological Society

Intraspecific variation in host specificity among phytophagous insects, driven by maternal responses to cues for ovipositional sites, may be associated with assortative mating behaviours involved in emergence of race and/or species diversity (Berlocher & Feder 2002). *Conopomorpha cramerella* is polyphagous for plant hosts across several families, and this led some authorities to suggest host-specific cryptic diversity was present in the species and/or possibly evolved in response to increased availability of novel crop hosts such as introduced cacao. Initial evidence of incomplete allele segregation at two allozyme loci between 'biotypes' of *C. cramerella* sampled from cacao and rambutan hosts in Malaysia (Rusnah *et al*. 1985) was prospectively reported as an indication of a need to investigate the taxonomic status of the biotypes (Muhamad & Tan 1987). In contrast, Yong *et al*. (1987) reported an absence of any geographic or host plant related trends in allozyme variation among *C. cramerella* biotypes sampled from Sri Lanka and Peninsula Malaysia. They conservatively concluded that biotypes of the pest, as defined by their use of host plants, were genetically conspecific. Additional allozyme analyses of eight polymorphic loci in *C. cramerella* sampled from cacao and rambutan hosts in Malaysia also failed to support earlier proposals of host-specific genetic variation in the pest; results were more consistent with outcomes of population genetic processes (Muhamad *et al*. 1989).

DNA sequence analyses of the pest by Shapiro *et al*. (2008) reported evidence only of limited and historically recent intraspecific genetic diversity among CPB sampled across much of its distribution in the Malay Archipelago. Their sequence analysis of polymorphic and independent mitochondrial and nuclear genes indicated the presence only of shallow population-level genetic variation in the widespread *C. cramerella* pest. Further, there was no apparent genetic segregation between CPB lineages sampled from cacao or rambutan hosts (taking into consideration that CPB were sparsely sampled from the latter host). The geographically broad but sparse sampling of CPB across >5000 km of the Malay Archipelago by Shapiro *et al*. (2008) potentially under-sampled the true diversity of the pest. Unpublished COI sequence analyses (refer [I Valenzuela comm.], p. 28, Gende 2012) of three putative CPB biotypes sampled in East New Britain during 2009 were genetically sorted to *C. cramerella* and two unrelated and unidentified moth species. These two unidentified CPB biotypes were trapped outside of cacao plantations, and their (unreported) status as pests to the cacao industry remains unclear. One factor which potentially confounded earlier CPB identifications concerns field-based species identification of immature moths sampled from cacao. Handlers lacking specialised taxonomic expertise may mistakenly identify immature stages of various Lepidoptera species present in cacao fruits to the common *C. cramerella* pest. For example, *T. zophophanes* are morphologically distinct from *C. cramerella* as adults; but differences between the two species during pre-adult stages of development are less obvious and require microscope use and a morphology-based key for their separation. Presence of *T. zophophanes* among larvae in cacao fruit was only detected in a single instance in the current study prior to the DNA barcoding, with other larvae initially assumed at field stations to be *C. cramerella* (noting that several untrained locals assisted with sample collection). Specimens which are phenotypically ambiguous due to their life stage or physical condition are liable to be misidentified. The facility of DNA barcoding in matching phenotypically ambiguous specimens to species is well documented. This is particularly so where reference DNA barcode libraries obtained by broad taxonomic and geographic sampling show evidence of defined genetic separations among taxonomically described species (deWaard *et al*. 2010; Gopurenko *et al*. 2013; Fletcher *et al*. 2017).

One of the problems affecting this and other studies using DNA barcoding for query-based species diagnostics of phenotypically ambiguous specimens is the delivery of species-level assignments using incomplete reference DNA barcode libraries (Ekrem *et al*. 2007). For example, comparative DNA barcode sequences at public repositories were available for four of the 14 taxonomically described *Conopomorpha* species and five unspecified *Conopomorpha* taxa (Fig. 3). Subsequently, among our sequence queries of test specimens, identity of one of the larvae found in a cacao pod at the offshore province of East New Britain was genetically matched (99.1% similarity) to an unspecified *Conopomorpha* species sampled earlier from Madang Province on mainland PNG. This unspecified *Conopomorpha* is registered with barcode index number BOLD:AAI5875 and includes prior DNA barcodes of 10 specimens (Fig. 3). DNA barcodes of six adult moths we incidentally trapped in plantations were genetically distinct from *C. cramerella* but unmatched to existing accessions at sequence repositories (Table 2). The poor physical condition of these incidental specimens negated their explicit identification via traditional taxonomic process; subsequently, their identity at lower taxonomic levels remains unclear. Future availability of matching DNA barcodes sourced from vouchered and described specimens will help to resolve species identity of all these additional taxa. In the interim, the BIN system as implemented in BOLD allows indexed registration of their DNA barcodes to terminal molecular taxonomic units. In the absence of available taxonomic description, BINs can provide valuable ecological and/or geographical information of unspecified taxa matched by their DNA barcodes (Zahiri *et al*. 2014). In some instances, this may also inform their probable genus and/or sub-genus relationships based on nearest genetic neighbour relationships with described taxa (Wilson *et al*. 2011).

In conclusion, DNA barcoding identified the majority of Lepidoptera sampled from three cacao-producing provinces in PNG to *C. cramerella*. The species showed limited and closely related genetic diversity across the disjunct provinces, largely consistent with that evidenced from elsewhere in the species distribution. This genetic pattern of reduced diversity in PNG, particularly in the offshore provinces, possibly resulted through dispersal or transfer of genetically impoverished founder moths among provinces. Transfer of infested inter-plantation stock materials and/or other practices lacking adequate bio-sanitary precautions likely facilitated the spread of this pest in PNG. Our evidence of several closely related mitochondrial matri-lineages of *C. cramerella* infesting cacao were not consistent with earlier suggestions of host-specific biotypes within

© 2021 Australian Entomological Society

C. cramerella. Our DNA barcoding of moth larvae from infested cacao pods indicated a minor portion in PNG plantations constituted additional species diversity. This included three specimens identified as the tortricid moth *T. zophophanes* and one unspecified *Conopomorpha* species distantly related to *C. cramerella* and requiring taxonomic description beyond the capacity of our report. The extent to which these additional moth species affected cacao plantations in PNG (and elsewhere) remains to be determined. Our research indicated they were likely less abundant than *C. cramerella* and at juvenile stages may only readily discerned by trained personnel or using genetic approaches such as DNA barcoding. Six unrelated adult moths sampled from plantations were unmatched to public DNA barcode records and likely represented incidental captures as opposed to cacao pests. Two of these moths were caught in *C. cramerella* pheromone-baited traps. Given the ease at which additional taxa may be mistaken for *C. cramerella* under field sampling conditions, we argue it is essential that genetic assisted taxonomic assessment of Lepidoptera found affecting cacao is used in the development and assessment of integrated pest management practices specifically designed for control of *C. cramerella*.

ACKNOWLEDGEMENTS

This work was funded by the Australian Centre for International Agricultural Research (SRA project: HORT/2018/114). Colleagues from the PNG Cocoa Board, PNG Department of Primary Industries, National Agricultural Research Institute (NARI) as well as growers are also thanked for their incredible hospitality and accommodating our request to sample insects across the provinces of East New Britain, East Sepik and the Autonomous Region of Bougainville. Field work provisions, sample preparations and laboratory tasks at NSW DPI Agricultural Institutes were provided by Terry Osborne, Ros Mirrington and Sujeewa Rathnayake at Menangle, Orange and Wagga Wagga, respectively. The map in Figure 1 was produced by J. Lye (Cesar Australia). We appreciate improvements of the manuscript following internal review by Bernie Dominiak and external anonymous peer referees.

REFERENCES

Baker GL. 1976. *Cryptophlebia encarpa* (Meyrick) (Lepidoptera: Tortricidae) as a pest of cacao pods in the Northern Province of Papua New Guinea. *Papua and New Guinea Agricultural Journal* **27**, 11–27.

Beevor PS, Mumford JD, Shah S, Day RK & Hall DR. 1993. Observations on pheromone-baited mass trapping for control of cocoa pod borer, *Conopomorpha cramerella*, in Sabah, East Malaysia. *Crop Protection* **12**, 134–140.

Bellis GA, Gopurenko D, Cookson B *et al*. 2015. Identification of incursions of *Culicoides* Latreille species (Diptera: Ceratopogonidae) in Australasia using morphological techniques and DNA barcoding. *Austral Entomology* **54**, 332–338.

Berlocher SH & Feder JL. 2002. Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology* **47**, 773–815.

Bradley JD. 1986. Identity of the South-East Asian cocoa moth, *Conopomorpha cramerella* (Snellen) (Lepidoptera: Gracillariidae), with description of three allied new species. *Bulletin of Entomological Research* **76**, 41–51.

Clement M, Posada D & Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657–1659.

Day RK. 1985. Control of the cocoa pod borer (Conopomorpha cramerella). PhD Thesis, University of London, London, United Kingdom.

Day RK. 1989. Effect of cocoa pod borer, *Conopomorpha cramerella*, on cocoa yield and quality in Sabah, Malaysia. *Crop Protection* **8**, 332–339.

De Prins J & De Prins W. 2019. Global Taxonomic Database of Gracillariidae (Lepidoptera). World Wide Web electronic publication (http://www.gracillariidae.net) [accessed December 20, 2019].

deWaard JR, Mitchell A, Keena MA *et al*. 2010. Towards a Global Barcode Library for *Lymantria* (Lepidoptera: Lymantriinae) Tussock Moths of Biosecurity Concern. *PLoS ONE* **5**, e14280.

Dunn GS. 1955. Miscellaneous pests of cocoa. *Papua New Guinea Agricultural Journal* **9**, 109–119.

Ekrem T, Willassen E & Stur E. 2007. A comprehensive DNA sequence library is essential for identification with DNA barcodes. *Molecular Phylogenetics and Evolution* **43**, 530–542.

FAO. 2004. Medium-term prospects for agricultural commodities: projections to the year 2010. (FAO Commodities and Trade Technical Paper No. 1) Editor-Alexander Sarris, 192 Pages, Food and Agriculture Organization of the United Nations Rome. ISBN 92-5-105077-5; ISSN 1729-9829

Fletcher MJ, Löker H, Mitchell A & Gopurenko D. 2017. A revision of the genus *Orosius* Distant (Hemiptera: Cicadellidae: Deltocephalinae) based on male genitalia and DNA barcoding. *Austral Entomology* **56**, 198–217.

Fletcher TB. 1916. One hundred notes on Indian insects. *Bulletin Agricultural Research Institute, Pusa* **59**, 1–39.

Forster JR & Forster G. 1776. *Characteres Generum Plantarum: quas in itinere ad insulas maris Australis, collegerunt, descripserunt, delinearunt, annis 1772–1775*. B. White, T. Cadell & P. Elmsly, Londini.

Froggatt JL. 1938. Pests of cocoa in the territory of New Guinea. *New Guinea Agricultural Gazette* **4** (4), 66–68.

Froggatt JL. 1940. Entomologist's report. In Annual report for the Department of Agriculture for the year ending 1939. *New Guinea Agricultural Gazette* **6** (2), 9–13.

Gende P. 2012. *The cocoa pod borer,* Conopomorpha cramerella *(Snellen) (Lepidoptera: Gracillariidae) in Papua New Guinea: its distribution and host plant preference*. PhD. Thesis. Faculty of Agriculture and Environment. University of Sydney, Sydney.

Gopurenko D, Bellis G, Yanase T *et al*. 2015. Integrative taxonomy to investigate species boundaries within *Culicoides* (Diptera: Ceratopogonidae): a case study using subgenus *Avaritia* from Australasia and eastern Asia. *Veterinaria Italiana* **51**, 345–378.

Gopurenko D, Fletcher MJ, Löcker H & Mitchell A. 2013. Morphological and DNA barcode species identifications of leafhoppers, planthoppers and treehoppers (Hemiptera: Auchenorrhyncha) at Barrow Island. In: *Records of the Western Australian Museum, Supplement 83: The Terrestrial Invertebrate Fauna of Barrow Island, Western Australia* (eds NR Gunawardene, JD Majer, CK Taylor & MS Harvey), pp. 253–285. Perth: Western Australian Museum.

Gopurenko D & Hughes JM. 2002. Regional patterns of genetic structure among Australian populations of the mud crab, *Scylla serrata* (Crustacea : Decapoda): evidence from mitochondrial DNA. *Marine and Freshwater Research* **53**, 849–857.

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

Hausmann A, Godfray HCJ, Huemer P *et al*. 2013. Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS ONE* **8**, e84518.

Hebert PDN, Penton EH, Burns JM, Janzen DH & Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14812–14817.

Hebert PDN, Ratnasingham S & de Waard JR. 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.

© 2021 Australian Entomological Society

Horak M & Komai F. 2016. *Cryptophlebia* Walsingham, 1900, *Thaumatotibia* Zacher, 1915, and *Archiphlebia* Komai & Horak, 2006, in Australia (Lepidoptera: Tortricidae: Olethreutinae: Grapholitini). *Zootaxa* **4179**, 441–477.

Hübner J. 1825. Verzeichniss bekannter Schmettlinge. 1-431, 1-72 (1816-[1827]).

Lees DC, Kawahara AY, Bouteleux O *et al*. 2013. DNA barcoding reveals a largely unknown fauna of Gracillariidae leaf-mining moths in the Neotropics. *Molecular Ecology Resources* **14**, 286–296.

Linnaeus C. 1753. *Species Plantarum*, Vol. 2. Laurentius Salvius, Stockholm.

Linnaeus C. 1767. *Systema Naturae*, ed. 12, Mantissa Plantarum. Laurentius Salvius, Stockholm, Sweden.

Loke WH, Tuck HC & Lim GS. 2006. *Supporting management of cocoa pod borer,* Conopomorpha cramerella *(Snellen) incursion in Papua New Guinea: Cocoa pod borer scoping study in Papua New Guinea*. Consultancy mission completion report. CAB International South & East Asian Regional Centre, Selangor, Malaysia.

Meyrick E. 1920. *Exotic Microlepidoptera(Marlborough)*. 2(11), 321-352. Taylor & Francis, Fleet Street, London.

Miller P. 1768. The gardeners dictionary: containing the best and newest methods of cultivating and improving the kitchen, fruit, flower garden, and nursery, as also, The 8th edition, revised and altered according to the latest system of botany and embellished with several copper-plates, which were not in some form, v.1, Printed for the author, and sold by John and Francis Rivington [and 23 others]. London.

Mitchell A & Gopurenko D. 2016. DNA Barcoding the Heliothinae (Lepidoptera: Noctuidae) of Australia and utility of DNA Barcodes for pest identification in *Helicoverpa* and relatives. *PLoS ONE* **11**, e0160895.

Muhamad R & Tan SG. 1987. Biochemical taxonomy of the cocoa pod borer (*Conopomorpha cramerella*): a prospective view. In: *Proceedings of the Symposium on Management of the Cocoa Pod Borer* (eds PAC Ooi, LG Chan, KC Khoo, CH Teoh, MJ Mamat, CT Ho & GS Lim), pp. 119–126. Pertanian Malaysia University, Serdang, Selangor.

Muhamad R, Tan SG, Gan YY, Rita S, Kanasar S & Asuan A. 1989. An electrophoretic study of natural populations of the Cocoa Pod Borer, *Conopomorpha cramerella* (Snellen) from Malaysia. *Pertanika* **12**, 1–6.

Mumford JD. 1986. Control of the cocoa pod borer (*Acrocereops cramerella*): a critical review. In: *Cocoa and Coconuts: Progress and Outlook* (eds E Pushparajah & PS Chew), pp. 277–286. Incorporated Society of Planters, Kuala Lumpur, Malaysia.

Niogret J, Ekayanti A, Ingram K *et al*. 2019. Development and behavioral ecology of *Conopomorpha cramerella* (Lepidoptera: Gracillariidae). *Florida Entomologist* **102**, 382–387.

Niogret J, Ekayanti A, Kendra PE *et al*. 2020. Host preferences of the cocoa pod borer, *Conopomorpha cramerella* (Lepidoptera:Gracillariidae), the main threat to cocoa production in Southeast Asia. *Entomologia Experimentalis et Applicata* **168**, 221–227.

Perry K, Baker G, Powis K, Kent JK, Ward CM & Baxter SW. 2018. Cryptic *Plutella* species show deep divergence despite the capacity to hybridize. *BMC Evolutionary Biology* **18**, 77.

Perry KD, Keller MA & Baxter SW. 2020. Genome-wide analysis of diamondback moth, *Plutella xylostella* L., from Brassica crops and wild host plants reveals no genetic structure in Australia. *Scientific Reports* **10**, 12047.

Picq S, Keena M, Havill N *et al*. 2018. Assessing the potential of genotyping-by-sequencing-derived single nucleotide polymorphisms to identify the geographic origins of intercepted gypsy moth (*Lymantria dispar*) specimens: a proof-of-concept study. *Evolutionary Applications* **11**, 325–339.

Pluzhnikov A & Donnelly P. 1996. Optimal sequencing strategies for surveying molecular genetic diversity. *Genetics* **144**, 1247–1262.

Posada D & Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**, 37–45.

Puillandre N, Dupas S, Dangles O *et al*. 2008. Genetic bottleneck in invasive species: the potato tuber moth adds to the list. *Biological Invasions* **10**, 319–333.

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

Ratnasingham S & Hebert PDN. 2013. A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS ONE* **8**, e66213.

Reynolds OL, Gopurenko D, & Guest D *et al*. 2019. Basic research on the cocoa pod borer in Papua New Guinea to permit effective pest management. Australian Centre for International Agricultural Research (ACIAR), Small Research and development activity, Final Report (project number HORT/2018/114).

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC *et al*. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Molecular Biology and Evolution* **34**, 3299–3302.

Rusnah MS, Muhamad R, Tan SG, Gan YY & Halmy A. 1985. Biological polymorphisms of Peptidase and α-GPDH in natural populations of the Cocoa Pod Borer *Conopomorpha cramerella* (Snellen). *Journal of Plant Protection in the Tropics* **2**, 49–52.

Sam K, Ctvrtecka R, Miller SE *et al*. 2017. Low host specificity and abundance of frugivorous lepidoptera in the lowland rain forests of Papua New Guinea. *PLoS ONE* **12**, e0171843.

Shapiro LH, Scheffer SJ, Maisin N *et al*. 2008. *Conopomorpha cramerella* (Lepidoptera: Gracillariidae) in the Malay Archipelago: genetic signature of a bottlenecked population? *Annals of the Entomological Society of America* **101**, 930–938.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H & Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**, 651–701.

Singh SP & Rethinam P. 2006. Cocoa pod borer threatening the sustainable coconut-cocoa cropping system in Papua New Guinea. *Coconut Research and Development Journal* **22**, 76–109.

Slatkin M & Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555–562.

Snellen PCT. (in van Deventer W) 1904. Over de ontwikkelingstoestanden van eenige Microlepidoptera van Java. *Tijdschrift voor Entomologie* **46**, 79–89.

Sonnerat P. 1782. *Voyage aux Indes Orientales*. L'auteur, Paris.

Szent-Ivany JJH. 1961. Insect pests of *Theobroma cacao* in the Territory of Papua New Guinea. *Papua and New Guinea Agricultural Journal* **13**, 127–147.

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.

Turner AJ. 1946. Contributions to our knowledge of the Australian Tortricidae (Lepidoptera). Part II. *Transactions of the Royal Society of South Australia* **70**, 185–220.

Uichanco LB. 1936. Miscellaneous notes on locusts, agriculture and people in Mindanao. *Philippines Agriculture* **25**, 565–588.

Villesen P. 2007. FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes* **7**, 965–968.

von Mueller FJH. 1857. Account of some new Australian plants. *Transactions of the Philosophical Institute of Victoria* **2**, 62–77.

von Mueller FJH. 1864. *Fragmenta Phytographiae Australiae*, Vol. 4, pp. 1858–1882. Auctoritate Gubern. Coloniae Victoriae, Ex Officina Joannis Ferres, Melbourne.

Wilson JJ, Rougerie R, Schonfeld J *et al*. 2011. When species matches are unavailable are DNA barcodes correctly assigned to higher taxa? An assessment using sphingid moths. *BMC Ecology* **11**, 18.

Yen JDL, Waters EK & Hamilton AJ. 2010. Cocoa pod borer *Conopomorpha cramerella* (Snellen) in Papua New Guinea: biosecurity models for New Ireland and the Autonomous Region of Bougainville. *Risk Analysis* **30**, 293–309.

Yong HS, Sarjan R & Ooi PAC. 1987. Allozyme variation in the cocoa pod borer, *Conopomorpha cramerella* (Lepidoptera: Gracillariidae), from different host plants. *Malaysian Journal of Science* **9**, 17–31.

Zahiri R, Lafontaine JD, Schmidt BC, deWaard JR, Zakharov EV & Hebert PDN. 2014. A transcontinental challenge – a test of DNA barcode performance for 1,541 species of Canadian Noctuoidea (Lepidoptera). *PLoS ONE* **9**, e92797.

Zhang A, Kuang LF, Maisin N *et al*. 2008. Activity evaluation of cocoa pod borer sex pheromone in cacao fields. *Environmental Entomology* **37**, 719–724.

Zhu X, Gopurenko D, Serrano M *et al*. 2019. Genetic evidence for plural introduction pathways of the invasive weed Paterson's curse (*Echium plantagineum* L.) to southern Australia. *PLoS ONE* **14**, e0222696.

© 2021 Australian Entomological Society

Accepted for publication 6 July 2021.

SUPPORTING INFORMATION

Additional supporting information may/can be found online in the supporting information tab for this article.

Table S1. Papua New Guinea moth sample information and associated DNA barcode records deposited at BOLD (http://www.boldsystems.org/) under project “Cocoa pod borer in Papua New Guinea” (project code: CPBNG). Specimen sample ID’s at NSW Department of Primary Industries collections linked to BOLD process ID’s and DNA barcode records. DNA barcode sequence length in base pairs (bp) as indicated. Best matched species-level identifications using the BOLD identification engine. Details of PNG province, location and sample site co-ordinates (decimal) as indicated. *Conopomorpha cramerella* haplotypes truncated to 624 BP and matched 100% to equal length haplotypes (CO-A to CO-F) reported by Shapiro *et al*. 2008, except where indicated as novel (CO-G*, CO-H*) or not applicable (N/A).

Table S2. DNA barcode equivalent COI sequences of *Conopomorpha* and other specimens available at GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and BOLD (http://www.boldsystems.org/) repositories (last searched 10/July/2020) used in neighbour joining analyses. Specimen sequences listed by GenBank accession and or by BOLD process ID, except where not reported (N/R). Taxonomic description (species level) as reported for specimen sequence where available and indicating Barcode Index Number (BOLD:BIN) as allocated at BOLD (Ratnasingham & Hebert 2013). Notes column indicate instances where (1) sequence suppressed at GenBank or (2) assigned to species based exclusively on BOLD:BIN association with identified taxa.

© 2021 Australian Entomological Society