Conopomorpha cramerella (Lepidoptera: Gracillariidae) in the Malay Archipelago: Genetic Signature of a Bottlenecked Population?

LEO H. SHAPIRO,[1,2] SONJA J. SCHEFFER,[1] NAVIES MAISIN,[3] SMILJA LAMBERT,[4] HUSSIN BIN PURUNG,[5] ENDANG SULISTYOWATI,[6] FERNANDO E. VEGA,[7] PAUL GENDE,[8] SAMSON LAUP,[8] ADE ROSMANA,[9] SYLVIA DJAM,[9] AND PRAKASH K. HEBBAR[10]

Ann. Entomol. Soc. Am. 101(5): 930–938 (2008)

ABSTRACT *Conopomorpha cramerella* (Snellen) (Lepidoptera: Gracillariidae) is a devastating pest of cacao, *Theobroma cacao* L. (Sterculiaceae), in Southeast Asia, particularly in the Malay Archipelago. We surveyed genetic variation at two unlinked loci, mitochondrial cytochrome oxidase I (COI) and nuclear elongation factor-1α (EF-1α), in *C. cramerella* from throughout most of their known geographic range. Given the enormous area sampled, COI variation is extremely low; EF-1α variation may be low as well, but this is more difficult to assess due to the lack of appropriate data sets for comparison. Our results strongly suggest that sampled *C. cramerella* populations have experienced at least one bottleneck in their recent past, although the possibility that COI variation has been reduced by a selective sweep cannot be excluded based on available data. We suggest that one or more bottlenecks likely occurred when *C. cramerella* from an as yet unknown source population, either within or outside the Malay Archipelago, became established on cacao, which is not endemic to this region (*Conopomorpha* is an Old World genus and cacao originated in the New World). Identification of the source of this pest could be important in efforts to identify natural enemies for biological control.

KEY WORDS *Theobroma cacao*, *Conopomorpha*, genetic bottleneck, cytochrome oxidase I, elongation factor-1α

Determining the geographic distribution of intraspecific genetic variation is a primary aim of phylogeography (Avise 2000), allowing inferences regarding population history and processes. Major geographic barriers or sharp ecological boundaries that reduce gene flow may result in distinctly structured populations (reviewed by Avise 2000). With greater isolation over longer periods, we expect to see a stronger genetic signature. Across island archipelagos, for example, we expect to see strong phylogeographic structure and significant isolation by distance in all but the most vagile terrestrial organisms (Heaney et al. 2005). In fact, as Darwin first recognized, isolation on islands often leads to speciation, with each species found on just one or a few islands in an archipelago (Gillespie and Roderick 2002).

Conopomorpha cramerella (Snellen) (Lepidoptera: Gracillariidae) is a devastating pest of cacao, *Theobroma cacao* L. (Sterculiaceae), in Southeast Asia, particularly in the sprawling Malay Archipelago (Malaysia, the Philippines, Borneo, and the islands of Indonesia). This tiny moth was first described (as *Gracilaria* [sic] *Cramerella* [sic]) from established cacao plantations in Java in 1904, but at that time it had already been familiar for decades to cacao growers in Indonesia, notably in Sulawesi, where it caused a rapid decline of the cocoa industry in the mid-1800s. Reductions in yields of cocoa pod beans due to *C. cramerella* damage may reach 40–50% (Day 1985, Anonymous 2003). In the 1990s, *C. cramerella* nearly caused a collapse of the cocoa industry in Malaysia (Shapiro and Rosenquist 2004), and it is now the primary limitation to growing cacao in Indonesia, Malaysia, and the Philippines (Anonymous 2006). The economic impact of this pest on cacao-dependent economies is enormous. In Indonesia, which is the world's third largest cocoa producer, accounting for ≈15% of world cocoa production (World Cocoa Foundation 2008), cacao is grown by 500,000 smallholders, and the Indonesian crop has an annual gross production value of $700 million; the industry is currently losing an estimated $300 million each year as a result of yield losses

[1] Systematic Entomology Laboratory, USDA–ARS, BARC-W, 10300 Baltimore Ave., Beltsville, MD 20705.
[2] Corresponding author, e-mail: leo.shapiro@ars.usda.gov.
[3] Malaysian Cocoa Board, Cocoa Research and Development Center, 91012 Tawau, Sabah, Malaysia.
[4] Mars Snack Food Australia, Bauarat, 3353 Victoria, Australia.
[5] Effem Foods S.E. Asia-PT Effem, Makassar, South Sulawesi, Indonesia.
[6] Indonesian Coffee and Cocoa Research Institute, Jl. PB Sudirman, 90, Jember, Java, Indonesia 68118.
[7] Insect Biocontrol Laboratory, USDA–ARS, BARC-W, 10300 Baltimore Ave., Beltsville, MD 20705.
[8] Cocoa and Coconut Institute, Rabaul, East New Britain, PNG.
[9] Plant Protection Department, Hassanudin University, Makassar, South Sulawesi, Indonesia.
[10] Mars Inc., Hackettstown, NJ 07840.

Fig. 1. Collecting sites for *C. cramerella* in the Malay Archipelago and New Guinea. Black circles indicate approximate sampling sites; circles that seem to be placed in the water represent sites on offshore islands too small to detail on this map.

and quality discounts (≈20% of international prices) due to *C. cramerella* (Chisholm et al. 2006). In contrast to Indonesia, Malaysia is not among the world's top producers of cocoa beans (it accounts for <1% of world production), but it is a major cocoa processor, with estimated 2006 export earnings from cocoa beans and cocoa products of ≈$600 million (Malaysian Cocoa Board 2008) and inconsistent bean quality due to *C. cramerella* has a severe impact on the processing industry. *C. cramerella* has only recently emerged as a threat in Papua New Guinea, where cacao is the second most important export crop after coffee, and it is now viewed as a potentially severe threat to cocoa production in Papua New Guinea.

Female *C. cramerella* lay eggs in cacao pod surface furrows; after hatching, larvae bore into the pods to feed on pulp and placenta tissue in the pod wall, causing premature ripening that results in small, flat beans and other damage. Efforts to control this pest have been limited in their success. Chemical pesticides are often too expensive, are ecologically disruptive, and may become less effective as local populations evolve resistance. Modifications of cultural practices (such as regular complete harvesting of pods and removal of husks) to interrupt the life cycle of *C. cramerella* have proven somewhat effective, but they are difficult to implement in practice. Control efforts using pheromone traps in the 1980s yielded inconsistent results, although recent work is much more promising (Zhang et al. 2008). Attempts to develop practical biological control methods using natural enemies have met with limited success. Rear-and-release trial using the egg parasitoid *Trichogrammatoidea bactrae fumata* Nagaraja (Hymenoptera: Trichogrammatidae) resulted in significant control, but the program was too complex and expensive to maintain and expand; efforts to make this approach more efficient are ongoing. Attempts to establish several other parasitoid wasps also have been unsuccessful (Vos et al. 2003).

Surprisingly, some fundamental questions about the basic biology of *C. cramerella* have gone unaddressed despite the longstanding interest in developing practical and effective control methods. Currently, no information is available regarding geographic patterns of genetic variation across the broad distribution of *C. cramerella;* in fact, it has even been unclear whether *C. cramerella* is a single species or a complex of unrecognized cryptic species. Another uncertainty is the origin of *C. cramerella*, which is of particular interest in the search for natural enemies. Cacao is originally from the New World (Sauer 1993) and *Conopomorpha* is confined to the Old World (De Prins and De Prins 2005), so the original host(s) of *C. cramerella* cannot be cacao. Several other native and introduced hosts have been reported for putative *C. cramerella* in cacao-growing regions, including rambutan (*Nephelium lappaceum*, Sapindaceae), nam-nam (*Cynometra cauliflora*, Fabaceae), kola (*Cola* spp., Sterculiaceae), and taun (*Pometia pinnata*, Sapindaceae).

We set out to address three questions: 1) Is *C. cramerella* on cacao truly a single species, or is it a complex of two or more species, as some researchers have suspected (e.g., Rita and Tan 1987)?; 2) What is the geographic population structure of *C. cramerella* across its distribution?; and 3) Can we use the patterns of genetic variation we find across the distribution of *C. cramerella* to infer the origin of *C. cramerella* populations now established on cacao?

Materials and Methods

We obtained *C. cramerella* larvae, pupae, and adults stored in alcohol from across most of their known distribution, with a maximum distance between samples of >5,000 km (Fig. 1; Table 1). DNA was extracted using DNeasy tissue kits (QIAGEN, Valencia, CA). An entire individual was used for most extractions, but we have retained as vouchers other individuals collected at the same time from the same trees. We amplified and sequenced a 624-bp portion of the mitochondrial gene cytochrome oxidase I (COI) from 92 individuals using primers 1500F (5′-ATTGGAACTTTATATTTTATATTTGG-3′) and 2392R (5′-CCTGTAGGAACAGCAATAATTATTG-3′), as well as two additional

Table 1. COI and EF-1α haplotype/allele frequencies for each sampling region

Geographic region	No. COI individuals. (*N*)	COI haplotypes (*N*)	No. EF-1α individuals (*N*)	EF-1α alleles (*N*)
Philippines				
Luzon	4	CO-A (4)	6	A (4); D (1); G (2); H (5)
Palawan	5	CO-A (3); CO-B (2)	6	A (5); D (2); I (3); J (1); W (1)
Mindanao	5	CO-A (5)	7	A (6); C (1); D (4); E (2); J (1)
Borneo				
Sabah (Malaysia)	5	CO-A (2); CO-B (3)	4	A (5); G (1); J (2)
Kalimantan (Indonesia)	4	CO-A (1); CO-C (3)	1	A (1); J (1)
Sulawesi (Indonesia)				
North Sulawesi	7	CO-A (7)	6	A (8); C (2); I (1); P (1)
South Sulawesi	12	CO-A (12)	7	A (10); E (2); J (2)
N. Maluku (Moluccas, Indonesia)	6	CO-A (6)	5	A (6); B (2); F (1); K (1)
New Guinea				
Irian Jaya (Indonesia)	6	CO-A (3); CO-B (3)	5	A (1); D (1); F (4); I (1); J (2); L (1)
East New Britain (PNG)	11	CO-B (11)	4	F (6); T (2)
Peninsular Malaysia	3	CO-A (1); CO-B (1); CO-C (1)	0	
Sumatra (Indonesia)	4	CO-A (3); CO-D (1)	3	A (5); V (1)
Java (Indonesia)	8	CO-A (7); CO-E (1)	2	M (1); N (1); P (1); T (1)
Bali (Indonesia)	8	CO-A (7); CO-F (1)	4	A (5); D (1); Q (1); S (1)
Flores (Indonesia)	4	CO-A (4)	4	A (5); B (2); U (1)
Total	92	6 haplotypes	64	21 alleles

The six COI haplotypes are labeled A through F. The 21 EF-1α alleles are labeled A through W (omitting O and R).

internal sequencing primers, CO2183F (5′-CAACATTTATTTTGATTTTTTGG-3′) and CO2191R (5′-CCCGGTAAAATTAAAATATAAACTTC-3′). For many samples, we also amplified and sequenced an 848-bp portion of the nuclear gene EF-1α (all coding sequence) using primers 40F (5′-GTCGTGATCGGACACGTCGATTCCGG-3′) and 71R (5′-CTTGCCCTTGGTGGCCTTCTCGG-3′), as well as three additional internal sequencing primers: EF46F (5′-TGAGGAAATCAAGAAGGAAG-3′), EF53R (5′-GCGAACTTGCAAGCAATGTGAGC-3′), and EF61R (5′-GATGGTTCCAACATGTTGTC-3′).

Because some of our EF-1α sequences showed evidence of heterozygosity or other differences from the modal EF-1α sequence, we cloned a subset of our EF-1α sequences using TOPO TA cloning kits from Invitrogen (Carlsbad, CA). To correct for rare incorporation of incorrect nucleotides that are subsequently preserved by cloning, which produces spurious variants, some other authors (e.g., Villablanca et al. 1998) have conservatively excluded all singleton clones from their data set. For the purposes of our study, however, this would have unacceptably biased our results. Instead, we used comparisons of the (very clean) original diploid electropherograms with putative cloned alleles to weed out cloned sequences that were inconsistent with the original diploid sequences (i.e., polymerase chain reaction [PCR] artifacts). Using this approach, we were able to specify the diploid genotypes for 64 individuals with a high degree of confidence.

For most amplifications we used a touch-down PCR profile as follows: initial denaturation at 92°C (2:00); one cycle of 92°C (10 s)-58°C (10 s)-72°C (1:30) followed by 12 additional cycles with annealing temperature dropped 1°C each cycle; 40 cycles of 92° (10 s)-45° (10 s)-72° (1:30); final extension of 72°C (10:00). PCR products were cleaned using QIAquick PCR purification kits (QIAGEN) and cycle sequenced in both directions using an ABI PRISM cycle sequencing kit, version 3.1 (Applied Biosystems, Foster City, CA). Sequenced products were run out on an ABI 3100 automated sequencer. We aligned and edited our sequences using Sequencher, version 4.5 (Genecodes Corporation, Ann Arbor, MI). For both COI and EF-1α, the regions amplified included only protein-coding sequence and alignments were therefore unambiguous. Genetic distances were calculated using PAUP*, version 4.0b10 for Macintosh (Swofford 2002). We used DnaSP, version 4.50.2 (Rozas et al. 2003) to estimate Watterson's theta (number of segregating sites) and nucleotide diversity and to test for evidence of deviations from neutrality using Tajima's test (Tajima 1989a), as well as the related tests of (Fu and Li 1993, Simonsen et al. 1995). Haplotype/allele networks were constructed with the help of Network, version 4.200, using a median-joining network algorithm to identify all possible shortest trees (Bandelt et al. 1999; program available from www.fluxus-engineering.com). All 92 COI and 64 diploid EF-1α sequences included here have been submitted to GenBank (accession nos. EU644510 through EU644601 [COI] and EU644610 through EU644673 [EF-1α]). GenBank accession numbers for representatives of all COI haplotypes and EF-1α alleles are shown in Table 2.

Results

We found very low levels of variation in both COI and EF-1α among all individuals sampled from across most of the known distribution of *C. cramerella.* For COI, uncorrected pairwise genetic distances ranged between 0 and 0.3%. Among 92 individuals analyzed

Table 2. Geographic locations of each of the six COI haplotypes found from 92 individuals and of each of the 21 EF-1α alleles found from 64 individuals

Haplotype/allele	Frequency	Sampling areas in which detected	GenBank accession no. (example)
COI			
CO-A	65	All areas (from both cacao and rambutan)	EU644563
CO-B	20	Palawan, Sabah, Irian Jaya, East New Britain, Peninsular Malaysia	EU644578
CO-C	4	Kalimantan, Peninsular Malaysia	EU644565
CO-D	1	Sumatra	EU644539
CO-E	1	Java (from rambutan)	EU644577
CO-F	1	Bali (from rambutan)	EU644589
Total	92		
EF-1α			
EF-A	61	All areas except East New Britain, Java	EU668004
EF-B	4	N. Maluku, Flores	EU668005
EF-C	3	Mindanao, N. Sulawesi	EU668006
EF-D	9	Luzon, Palawan, Mindanao, Irian Jaya, Bali	EU668007
EF-E	4	Mindanao, S. Sulawesi	EU668008
EF-F	11	N. Maluku, Irian Jaya, E. New Britain	EU668009
EF-G	3	Luzon, Sabah	EU668010
EF-H	5	Luzon	EU668011
EF-I	5	Palawan, N. Sulawesi, Irian Jaya	EU668012
EF-J	9	Palawan, Mindanao, Sabah, Kalimantan	EU668013
EF-K	1	S. Sulawesi, Irian Jaya	EU668014
EF-L	1	Irian Jaya	EU668015
EF-M	1	Java	EU668016
EF-N	1	Java	EU668017
EF-P	2	N. Sulawesi, Java	EU668019
EF-Q	1	Bali	EU727205
EF-S	1	Bali	EU668020
EF-T	3	East New Britain, Java	EU668021
EF-U	1	Flores	EU727206
EF-V	1	Sumatra	EU708727
EF-W	1	Palawan	EU708728
Total	128		

for COI, we found variation (in all cases involving silent changes) at only five sites in 624 bp (0.8% variable sites). Even at these five sites most individuals shared the same base, and among all COI sequences we found a total of only six haplotypes (Table 2; Fig. 2). Estimated theta per site and nucleotide diversity per site (both corrected for the four-fold relative reduction in effective population size for mitochondrial DNA [mtDNA]) were 0.632 and 0.316%, respectively. No geographic structure was apparent in the COI data set (Table 2). The most common haplotype (COI-A: 71% of sequences) was found in all sampled locations and occurred on both cacao and rambutan. The second most common haplotype (CO-B: 22% of sequences) also was found in geographically disjunct samples ranging from Peninsular Malaysia to East New Britain. The third most common haplotype (4%) was present in both Kalimantan (Borneo) and peninsular Malaysia. The remaining three haplotypes were singletons from Sumatra, Java, and Bali (the latter two from the small fraction of our samples that were from rambutan; see below).

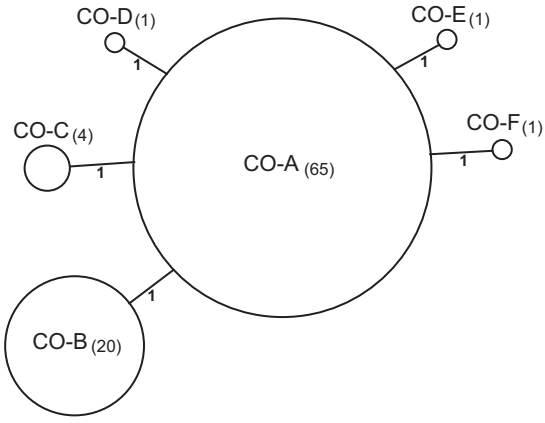

Fig. 2. Haplotype network to show differences among COI haplotypes recovered from 92 individuals. Frequency for each haplotype is shown in parentheses. Bolded numbers adjacent to lines indicate the number of nucleotide differences between the two joined haplotypes.

For EF-1α, uncorrected pairwise genetic distances ranged between 0 and 0.7%, with most between 0 and 0.5%. Among 64 individuals (=128 allele copies) analyzed for EF-1α (nearly all of which were also sequenced for COI), we found variation (in all cases involving silent changes) at 19 sites in 848 bp (2.2% variable sites). At most of these variable positions nearly all individuals were homozygous for the modal base, defining a total of 21 extremely similar alleles (Fig. 3). Estimated theta per site and nucleotide diversity per site was 0.413 and 0.172%, respectively. The

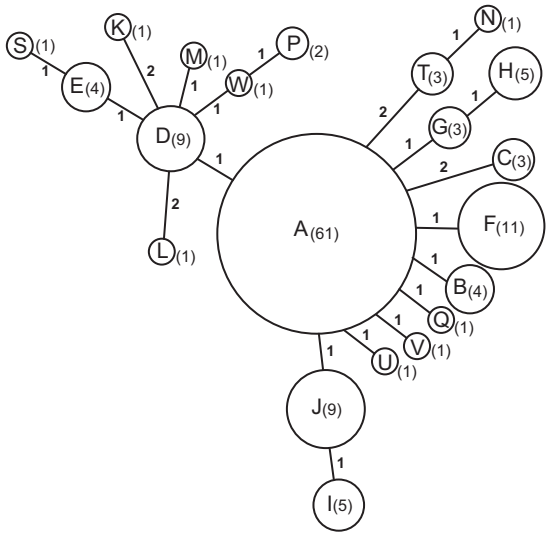

Fig. 3. Network to show differences among EF-1α alleles recovered from subset of 64 individuals (several other equally parsimonious, but nearly identical, networks are also possible). Frequency for each allele (of 128 total) is shown in parentheses. Bolded numbers adjacent to lines indicate the number of nucleotide differences between the two joined alleles.

most common allele (EF-A) had an overall frequency of 48%, and the four most common alleles together accounted for 70% of sampled alleles (Table 2). The remaining 17 alleles were all quite rare: none had a frequency >4%, and most represented <2% of sampled alleles. No geographic structure was apparent in the EF-1α data set. Although it is possible that with more intensive sampling we might find that some alleles are geographically limited in their distribution, in most cases where we detected an allele in more than one population, these populations were in disparate locations (Table 2).

The Tajima neutrality test was borderline significant for EF-1α ($0.10 > P > 0.05$), but two related test statistics, Fu and Li's D* and F*, both yielded nonsignificant values ($P > 0.10$). Conversely, the Tajima test for COI yielded $P > 0.10$, but D* and F* both yielded $0.10 > P > 0.05$.

Nearly all individuals in this study were sampled from cacao due to unfortunate logistical constraints, but our Sulawesi sample included two individuals from rambutan and one from white jambu (*Syzygium* sp., Myrtaceae), and our Java and Bali samples each included four individuals from rambutan. All three of the noncacao Sulawesi individuals had the common genotype for both COI and EF-1α; for both Java and Bali, our noncacao samples included three individuals with the common COI haplotype and one individual with a unique COI haplotype, differing from the common haplotype at just a single position (no EF-1α sequence was collected from these Java and Bali individuals).

Discussion

The low level of genetic variation and lack of apparent geographic structure detected among *C. cramerella* from across the enormous distances and geographically isolated land masses sampled strongly suggest that our samples have a common origin in the relatively recent past. Such a reduction in genetic variation is typically seen after a population bottleneck (Nei et al. 1975), and is often evident in introduced insect populations subsequent to founding events (Ross et al. 1996, Villablanca et al. 1998, Tsutsui et al. 2000, Downie 2002, Oliver 2006). For example, in a phylogeographic study of the seed-feeding torymid wasp *Megastigmus transvaalensis* (Hussey) using 800 bp of COI, Scheffer and Grissell (2003) found that 24 individuals sampled from their native and introduced hosts in Africa represented 22 haplotypes, whereas 20 individuals sampled from introduced hosts in Florida, California, and Hawaii all represented just a single haplotype. Another example of this pattern is provided by the chrysomelid beetle *Leptinotarsa decemlineata* (Say), which reached Europe from North America around the beginning of the 20th century and spread across the continent within a few decades. Grapputo et al. (2005) sampled 58 beetles from 13 North American populations and 51 beetles from eight European populations. They found a total of 20 mitochondrial haplotypes, but all 51 European beetles shared a single haplotype, which was identical to that of the 10 North American beetles sampled from Idaho.

In contrast, nonintroduced populations of phytophagous insects sampled across such great distances as our *C. cramerella* typically show far greater variation than we found in the current study. For example, Segraves and Pellmyr (2001) studied the prodoxid moth *Tegeticula maculata* Riley in southern California and adjacent Mexico. They found 18 mitochondrial COI haplotypes among 46 individuals from eight populations (the greatest distance between populations was <800 km). Thirty-two of 755 (4.2%) nucleotide positions were variable, and pairs of haplotypes differed at between one and 15 positions. Six of the eight populations had more than one haplotype, and only one haplotype was found in more than one population. Knowles et al. (1999) analyzed COI from 92 individuals of the widespread chrysomelid beetle *Ophraella communa* LeSage sampled from 10 populations across North America. Of a total 400 bp sequenced, they found 68 variable positions (17%), yielding 48 haplotypes, with sequence divergences ranging from 1.04 to 3.6% (the other species in their study, the far more narrowly distributed *O. bilineata* [Kirby], had unique haplotypes in 19 of 22 individuals sampled). In an analysis perhaps more directly comparable with the present study, Knowles (2001) examined COI from the flightless grasshopper *Melanoplus oregonensis* (Thomas), which is found in isolated "sky island" montane meadows in the northern Rocky Mountains. Analysis of 1275 bp of COI from 124 individuals revealed 292 variable sites (23%), yielding 104 unique haplotypes, none of them geographically widespread.

Other phylogeographic analyses of intraspecific mitochondrial genetic variation on sky islands yield similar results. Masta (2000) examined ≈815 bp of ND1–16S in *Habronattus pugillis* Griswold jumping spiders from 13 mountain ranges in southeastern Arizona. Sampling 86 individuals, she found 160 variable positions (≈20%), yielding 81 haplotypes, with some clear geographic structuring of clades. Similarly, Smith and Farrell (2005a) studied the phylogeography of the cerambycid beetle *Moneilema appressum* LeConte on southwestern sky islands. They collected 64 beetles from 16 locations in New Mexico, Chihuahua, and Durango and sequenced 1413 bp of COI, of which 403 (28%) were variable, yielding 63 unique haplotypes.

Phylogeographic studies of clearly conspecific terrestrial insects from across actual island archipelagos are scarce, presumably because isolation on such islands tends to promote speciation. Jordan et al. (2005) examined patterns of genetic variation in two endemic damselflies, *Megalagrion xanthomelas* (Selys-Longchamps) and *M. pacificum* McLachlan (Odonata: Coenagrionidae), across the Hawaiian Archipelago. They sequenced 663 bp of COII from 130 *M. xanthomelas* and 27 *M. pacificum*, finding 37 variable positions (5.6%), which yielded 23 haplotypes for *M. xanthomelas* and eight haplotypes for *M. pacificum.* Park et al. (2006) sequenced 936 bp of COII + COIII for 27 *Reticulitermes speratus* (Kolbe) (Isoptera: Rhinotermitidae) from across the Korean Peninsula and Japanese Archipelago, finding 53 variable positions (5.7%) that defined 15 haplotypes. Numerous other examples in the literature indicate that the COI variation evident in our data set—from weakly flying *C. cramerella* sampled from widespread islands across enormous expanses of water—is, as would be expected from an introduction or some other form of population bottleneck, extraordinarily low (Juan et al. 1998, De La Rúa et al. 2000, Mun et al. 2003, DeChaine and Martin 2005, Smith and Farrell 2005b).

Although it is clear that the variation in COI is extremely low, it is more difficult to find EF-1α data sets to which our data can be compared. Because EF-1α typically evolves too slowly to be generally useful for intraspecific studies, there do not seem to be comparable studies with which to assess the significance of the low variation and lack of apparent geographic structure in this gene across 5,000 km of archipelago. However, although comparable intraspecific, phylogeographic studies using EF-1α are not available, several interspecific studies have compared levels of sequence divergence between COI and EF-1α. In most interspecific studies, COI tends to show a higher rate of change relative to EF-1α, with COI uncorrected pairwise distances typically ranging from ≈2 to 10 times those for EF-1α (e.g., Rubinoff and Sperling 2002, rate ratio of about two between *Hemileuca* [Saturniidae] moths; Kawakita and Kato 2006, rate ratio of ≈2 to 3 between *Epicephala* [Gracillariidae] moths; Mullen 2006, rate ratio of <2 to >10, with most ≈2 to 4, between *Limenitis* [Nymphalidae] butterflies; Roe and Sperling (2007), rate ratio of three or less between *Dioryctria* [Pyralidae] moths). A notable feature of our dataset is that EF-1α and COI genetic divergences are very similar, but given that our data are from within a single species and absolute divergences for both genes are so low, it seems quite likely that our unusual rate ratio is simply a statistical artifact.

It is important to note that although a bottleneck would be expected to reduce diversity of most genes, it would be unsurprising if this effect were less pronounced for nuclear genes than for mtDNA, as may be the case in our data set. The haploid, maternal inheritance of mtDNA is expected to result in a more markedly reduced effective population size (and hence more loss of variation) for COI than for diploid, biparentally inherited nuclear genes such as EF-1α under many (although not all) conditions (Birky et al. 1989, Chesser and Baker 1996). Villablanca et al. (1998), for example, found that invading populations of the tephritid fly *Ceratitis capitata* (Wiedemann) lost a far greater proportion of genetic variation at mitochondrial loci than at nuclear loci and attributed this to the typically larger effective population sizes of nuclear genes relative to mtDNA. Although it seems likely that such an effect explains our data, a possible alternative scenario is that EF-1α variation is in fact not surprisingly low even for the enormous geographic area from which it was sampled, but that COI variation has been reduced by a powerful and geographically widespread selective sweep (e.g., due to *Wohlbachia* infection; Jiggins 2003, Hurst and Jiggins 2005) impacting the entire Malay Archipelago and New Guinea. Distinguishing these two scenarios from the COI data alone is very difficult because demographic phenomena such as a population bottleneck can leave the same genetic signature as a selective sweep acting on mitochondrial symbionts (Tajima 1989b, Hurst and Jiggins 2005). In our particular situation, given the history and biology of *C. cramerella* and cacao, a demographic explanation (population bottleneck due to a founder event on cacao in the sampled region) seems quite likely a priori, but the possibility of a selective sweep cannot be excluded. Because both selection and demographic history may lead to loss of genetic variation and significant neutrality tests, the key consideration is whether these are evident at just a single locus or at multiple unlinked loci (Rokas et al. 2001, Hurst and Jiggins 2005, Nielsen 2005). Thus, if our *C. cramerella* EF-1α variation (along with COI) is truly reduced, this supports the scenario of a population bottleneck, but if it is no lower than expected given our sampling, this would be consistent with the possibility of a selective sweep specifically impacting mtDNA. Results from the neutrality tests for *C. cramerella* COI and EF-1α, which are based on the particular patterns of polymorphism in a set of DNA sequences, are not conclusive on this question, but are suggestive. Although rigorous statistical conclusions are not possible, data from these two genes—one nuclear and one mitochondrial and thus evolving under very different evolutionary conditions—both seem to be borderline deviant from neutrality. Most important, rather than indicate that one gene (e.g., COI) clearly deviates from neutral expec-

tations, whereas the other clearly does not, neutrality tests for the two genes seem to show a similar pattern. This would be expected if the observed loss of genetic variation is due to a shared demographic history impacting both genes, but it would not be expected for a selective sweep impacting just one of the genes. Clearly, however, more data from additional loci would be required to provide conclusive support for one scenario or the other, and because the power of tests used to detect selection/bottlenecks may depend strongly on the particular timing and scale of these events (Simonsen et al. 1995, Depaulis et al. 2003), clarifying this issue may be very difficult.

Barring a selective sweep scenario, a genetic bottleneck in *C. cramerella* could have resulted from either 1) a host shift onto cacao that occurred within the sampled area or 2) an introduction from outside the sampled area (these two possibilities are not mutually exclusive). Because *Conopomorpha* is confined to the Old World and cacao is from the New World, *C. cramerella* must necessarily have fed originally on some other host plant or plants. Whether it underwent a host shift such that the new cacao-feeding population became essentially specialized on this new host or whether it was a simple dietary expansion to include cacao along with other hosts is not known. If the colonization of cacao was a true host shift involving genetic changes and the cessation of gene flow, this shift could have occurred at a location where the original host(s) and cacao already occurred together, with the lineage(s) moving onto cacao representing a subset of the genetic diversity present on the original host(s) in the same area. In this case, we should be able to infer the source by finding greater genetic variation among *C. cramerella* on noncacao host(s) sympatric with cacao, with the sampled *C. cramerella* sequences from the current study nested within this variation (e.g., Downie 2002). In the current study, we unfortunately have only a few samples from noncacao host plants: four each from Java and Bali rambutan, two from Sulawesi rambutan, and one from Sulawesi white jambu. The individual from white jambu and eight of the 10 rambutan individuals had the most common of the six COI haplotypes identified in this study, but the remaining two rambutan individuals (one from Java and one from Bali) represented unique haplotypes. Thus, we have found three COI haplotypes among 10 individuals from rambutan (specifically, the common and widespread *C. cramerella* haplotype plus two others known so far only from rambutan) and just four haplotypes (one of them a singleton) among 81 individuals from cacao sampled from a large number of far flung populations. These results hint at the possibility that *C. cramerella* on rambutan may harbor more genetic diversity than those on cacao, but the question of whether *C. cramerella* haplotypes on cacao represent a subset of those on rambutan or other hosts can only be addressed by far more extensive, carefully designed sampling from both cacao and other hosts.

Alternatively, it is possible that shifting onto cacao may not have required any new adaptation, with *C. cramerella* adding cacao to its repertoire as soon as it became available. In this case, to explain the observed low levels of mitochondrial variation we would infer that *C. cramerella* arrived from some region outside the large area we have genetically characterized and that this subsample of the original genetic variation has subsequently spread across our sampled area (onto both cacao and noncacao hosts) after one or a few initial introductions. If this is the case, we should find greater genetic variation when we are able to sample *C. cramerella* from its currently unknown native range.

Identifying the geographic and host origins of a new pest can be quite challenging. For example, another gracillariid moth, the horse chestnut leaf miner (*Cameraria ohridella* Deschka and Dimic) was first discovered in Macedonia in 1985 (Deschka and Dimic 1986) and subsequently spread rapidly; it now occurs across most of temperate Europe, where it is a serious pest of several important native and exotic ornamental trees. Despite much interest, however, it has not yet been possible to determine the origin of this moth or its natural hosts (Grabenweger and Grill 2000, Freise et al. 2004). As with *C. cramerella*, this lack of information has hindered efforts to identify potentially useful natural enemies. Identifying the original host(s) and geographic origin of *C. cramerella* will require broad sampling of possible noncacao hosts and expansion of our geographic sampling to areas that have not yet been genetically characterized, such as Sri Lanka, Thailand, Vietnam, China, Taiwan, and Hong Kong. Other species in the *Conopomorpha cramerella* complex are known economic pests of litchi (*Litchi chinensis* Sonn.; Sapindaceae), rambutan, and other commercial fruits (Bradley 1986, Menzel 2002), so clarifying the taxonomic diversity, systematic relationships, and host associations within this group will have practical significance extending well beyond cacao.

Acknowledgments

We thank Matt Lewis for technical help with data collection and Brian Verrelli for helpful discussions. Cathy Aime, John Brown, Dug Miller, and two anonymous reviewers provided useful comments on the manuscript. We are grateful to Mars, Inc. for financial and logistical support of our work. We also acknowledge support from all the organizations that assisted in obtaining our samples, including Agricultural Cooperative Development International–Volunteers on Overseas Assistance (Washington, DC.), Indonesian Coffee and Cocoa Research Institute (Indonesia), and the Malaysian Cocoa Board (Malaysia).

References Cited

Anonymous. 2003. Strengthening Sulawesi's smallholder cocoa sector. GRO-Cocoa [Global Research On Cocoa] Newsletter. Issue 3 (June 2003). CABI Bioscience, Surrey, United Kingdom.

Anonymous. 2006. CPB in PNG: combating inevitable invasion. GRO-Cocoa [Global Research On Cocoa] Newsletter. Issue 9 (June 2006). CABI Bioscience, Surrey, United Kingdom.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Bandelt, H.-J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16: 37–48.

Birky, C. W., Jr., P. Fuerst, and T. Maruyama. 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121: 613–627.

Bradley, J. D. 1986. Identity of the South-East Asian cocoa moth, *Conopomorpha cramerella* (Snellen) (Lepidoptera: Gracillariidae), with descriptions of three allied new species. Bull. Entomol. Res. 76: 41–51.

Chesser, R. K., and R. J. Baker. 1996. Effective sizes and dynamics of uniparentally and diparentally inherited genes. Genetics 144: 1225–1235.

Chisholm, R. H., J. Toledano, and M. N. Benalt. 2006. "The role of the World Bank in supporting agriculture in Asia: the place for technology and innovation." Presentation at the 2006 Cocoa Symposium held Feb. 9–10, 2006 at The National Academies, Washington, DC. (www.cocoasymposium.com/2006/abstracts).

Day, R. K. 1985. Control of the cocoa pod borer (*Conopomorpha cramerella*). Ph.D. dissertation, University of London, London, United Kingdom.

De La Rúa, P., U. E. Simon, A. C. Tilde, R.F.A. Moritz, and S. Fuchs. 2000. MtDNA variation in *Apis cerana* populations from the Philippines. Heredity 84: 124–130.

DeChaine, E. G., and A. P. Martin. 2005. Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. J. Biogeogr. 32: 1943–1956.

De Prins, W., and J. De Prins. 2005. Gracillariidae (Lepidoptera). *In* World catalogue of insects 6: 1–502. Apollo Books, Stenstrup, Denmark.

Depaulis, F., S. Mousset, and M. Veuille. 2003. Power of neutrality tests to detect bottlenecks and hitchhiking. J. Mol. Evol. 57: S190–S200.

Deschka, G., and N. Dimic. 1986. *Cameraria ohridella* n. sp. aus Mazedonien, Jugoslawien (Lepidoptera, Lithocelletidae). Acta Entomol. Jugosl. 22: 11–23.

Downie, D. 2002. Locating the sources of an invasive pest, grape phylloxera, using a mitochondrial DNA gene genealogy. Mol. Ecol. 11: 2013–2026.

Freise, J. F., W. Heitland, and A. Sturm. 2004. Host plant range of the horse-chestnut leaf miner *Cameraria ohridella* (Lepidoptera: Gracillariidae), a pest of the white flowering horse-chestnut, *Aesculus hippocastanum.* Mitt. Dtsch. Ges. Allg. Angew. Entomol. 14: 351–354.

Fu, Y.-X., and W.-H. Li. 1993. Statistical tests of neutrality of mutations. Genetics 133: 693–709.

Gillespie, R. G., and G. K. Roderick. 2002. Arthropods on islands: colonization, speciation, and conservation. Annu. Rev. Entomol. 47: 595–632.

Grabenweger, G., and R. Grill. 2000. On the place of origin of *Cameraria ohridella* Deschka and Dimic (Lepidoptera: Gracillariidae). Beitr. Entomofaunistk. 1: 9–17.

Grapputo, A., S. Boman, L. Lindström, A. Lyytinen, and J. Mappes. 2005. The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. Mol. Ecol. 14: 4207–4219.

Heaney, L. R., J. S. Walsh, and A. T. Peterson. 2005. The roles of geological history and colonization abilities in genetic differentiation between mammalian populations in the Philippine archipelago. J. Biogeogr. 32: 229–247.

Hurst, G.D.D., and F. M. Jiggins. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts Proc. R. Soc. B 272: 1525–1534.

Jiggins, F. M. 2003. Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. Genetics 164: 5–12.

Jordan, S., C. Simon, D. Foote, and R. Englund. 2005. Phylogeographic patterns of Hawaiian *Megalagrion* damselflies (Odonata: Coenagrionidae) correlate with Pleistocene island boundaries. Mol. Ecol. 14: 3457–3470.

Juan, C., K. M. Ibrahim, P. Oromí, and G. M. Hewitt. 1998. The phylogeography of the darkling beetle *Hegeter politus* in the eastern Canary Islands. Proc. R. Soc. Lond. B 265: 135–140.

Kawakita, A., and M. Kato. 2006. Assessment of the diversity and species specificity of the mutualistic association between *Epicephala* moths and *Glochidion* trees. Mol. Ecol. 15: 3567–3581.

Knowles, L. L., D. J. Futuyma, W. F. Eanes, and B. Rannala. 1999. Insights into speciation mode from historical demography in the phytophagous beetle *Ophraella.* Evolution 53: 1846–1856.

Knowles, L. L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. Mol. Ecol. 10: 691–701.

Malaysian Cocoa Board. 2008. Malaysian Cocoa Board. (www.koko.gov.my).

Masta, S. E. 2000. Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): recent vicariance of sky island populations? Evolution 54: 1699–1711.

Menzel, C. 2002. The Lychee Crop in Asia and the Pacific (RAP Publication 2002/16). Food and Agriculture Organization of the United Nations, Regional Office for Asia and the Pacific. Bangkok, Thailand.

Mullen, S. P. 2006. Wing pattern evolution and the origins of mimicry among North American admiral butterflies (Nymphalidae: *Limenitis*). Mol. Phylogenet. Evol. 39: 747–758.

Mun, J., A. J. Bohonak, and G. K. Roderick. 2003. Population structure of the pumpkin fruit fly *Bactrocera depressa* (Tephritidae) in Korea and Japan: Pliocene allopatry or recent invasion? Mol. Ecol. 12: 2941–2951.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29: 1–10.

Nielsen, R. 2005. Molecular signatures of natural selection. Annu. Rev. Genet. 39: 197–218.

Oliver, J. C. 2006. Population genetic effects of human-mediated plant range expansions. Oikos 112: 456–463.

Park, Y. C., O. Kitade, M. Schwarz, J. P. Kim, and W. Kim. 2006. Intraspecific molecular phylogeny, genetic variation and phylogeography *of Reticulitermes speratus* (Isoptera: Rhinotermitidae). Mol. Cell 21: 89–103.

Rita, M. and S. G. Tan. 1987. Biochemical taxonomy of the cocoa pod borer: a prospective view, pp. 119–126. *In* P.A.C. Ooi, L. G. Chan, K. C. Khoo, C. H. Teoh, M. J. Mamat, C. T. Ho, and G. S. Lim [eds.], Management of the cocoa pod borer. The Malaysian Plant Protection Society, Kuala Lumpur, Malaysia.

Roe, A. D., and F.A.H. Sperling. 2007. Population structure and species boundary delimitation of cryptic *Dioryctria* moths: an integrative approach. Mol. Ecol. 16: 3617–3633.

Rokas, A., R. J. Atkinson, G. S. Brown, S. A. West, and G. N. Stone. 2001. Understanding patterns of genetic diversity in the oak gallwasp *Biorhiza pallida:* demographic history or a *Wolbachia* selective sweep? Heredity 87: 294–304.

Ross, K. G., E. L. Vargo, and L. Keller. 1996. Social evolution in a new environment: the case of introduced fire ants. Proc. Nat. Acad. Sci. USA 93: 3021–3025.

Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497.

Rubinoff, D., and F.A.H. Sperling. 2002. Evolution of ecological traits and wing morphology in *Hemileuca* (Saturniidae) based on a two gene phylogeny. Mol. Phylogenet. Evol. 25: 70–86.

Sauer, J. D. 1993. Historical Geography of Crop Plants. CRC, Boca Raton, FL.

Scheffer, S. J., and E. E. Grissell. 2003. Tracing the geographical origin of *Megastigmus transvaalensis* (Hymenoptera: Torymidae): an African wasp feeding on a South American plant in North America. Mol. Ecol. 12: 415–421.

Segraves, K. A., and O. Pellmyr. 2001. Phylogeography of the yucca moth *Tegeticula maculata:* the role of historical biogeography in reconciling high genetic structure with limited speciation. Mol. Ecol. 10: 1247–1253.

Shapiro, H.-Y., and E. M. Rosenquist. 2004. Public/private partnerships in agroforestry: the example of working together to improve cocoa sustainability. Agrofor. Syst. 61: 453–462.

Simonsen, K. L., G. A. Churchill, and C. A. Aquadro. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics 141: 413–429.

Smith, C. I., and B. D. Farrell. 2005a. Phylogeography of the longhorn cactus beetle *Moneilema appressum* LeConte (Coleoptera: Cerambycidae): was the differentiation of the Madrean sky islands driven by Pleistocene climate changes? Mol. Ecol. 14: 3049–3065.

Smith, C. I., and B. D. Farrell. 2005b. Range expansions in the flightless longhorn cactus beetles, *Moneilema gigas* and *Moneilema armatum*, in response to Pleistocene climate changes. Mol. Ecol. 14: 1025–1044.

Swofford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods). Sinauer, Sunderland, MA.

Tajima, F. 1989a. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Tajima, F. 1989b. The effect of change in population size on DNA polymorphism. Genetics 123: 597–601.

Tsutsui, N. D., A. V. Suarez, D. A. Holway, and T. J. Case. 2000. Reduced genetic variation and the success of an invasive species. Proc. Natl. Acad. Sci. USA 97, 5948–5953.

Villablanca, F. X., G. K. Roderick, and S. Palumbi. 1998. Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. Mol. Ecol. 7: 547–560.

Vos, J.G.M., B. J. Ritchie, and J. Flood. 2003. Discovery learning about cacao: an inspirational guide for training facilitators. CABI Bioscience, Surrey, United Kingdom.

World Cocoa Foundation. 2008. World Cocoa Foundation website www.worldcocoafoundation.org/info-center checked April 3, 2008.

Zhang, A., L. F. Kuang, N. Maisin, B. Karumuru, D. R. Hall, I. Virdiana, S. Lambert, H. B. Purung, S. Wang, and P. Hebbar. 2008. Activity evaluation of cocoa pod borer sex pheromone in cacao fields. Environ. Entomol. 37: 719–724.

Received 20 December 2006; accepted 15 May 2008.