MR. JONATHAN BRADLEY REIL (Orcid ID : 0000-0001-9511-960X)

Article type : Original Article

Title: **Transpacific coalescent pathways of coconut rhinoceros beetle biotypes: resistance to biological control catalyzes resurgence of an old pest**

Running Title: **Coalescence of resurgent beetle pest biotypes**

J. Bradley Reil[1], Camiel Doorenweerd[1], Michael San Jose[1], Sheina Sim[1,2], Scott Geib[2], Daniel Rubinoff[1,3]

[1]Department of Plant and Environmental Protection Sciences, 310 Gilmore Hall, 3050 Maile Way, University of Hawaii at Manoa, Honolulu, Hawaii 96822

[2]Tropical Crop and Commodity Protection Research Unit, Daniel K Inouye U.S. Pacific Basin Agricultural Research Center, USDA, Agricultural Research Services, Hilo, HI, USA

[3]Corresponding Author: rubinoff@hawaii.edu

ABSTRACT

Biological control agents have several advantages over chemical control for pest management, including the capability to restore ecosystem balance with minimal non-target effects and a lower propensity for targets to develop resistance. These factors are particularly important in the invasive species control. The coconut rhinoceros beetle (*Oryctes rhinoceros* Linnaeus) is a major palm pest that invaded many Pacific islands in the early 20th century through human-mediated dispersal. Application of the *Oryctes* nudivirus in the 1960's successfully halted the beetle's first invasion wave and made it a textbook example of successful biological control. However, a recently discovered *O. rhinoceros* biotype that is resistant to the nudivirus appears to be correlated with a new invasion wave. We performed a

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14879

population genomics analysis of 172 *O. rhinoceros* from seven regions, including native and invasive populations, to reconstruct invasion pathways and explore correlation between recent invasions and biotypes. From ddRAD sequencing, we generated datasets ranging from 4,000 to 209,000 loci using STACKS and IPYRAD software pipelines and compared genetic signal in downstream clustering and phylogenetic analyses. Analysis suggests that the *O. rhinoceros* resurgence is mediated by the nudivirus-resistant biotype. Genomic data has proven essential to understanding the new *O. rhinoceros* biotype's, invasion patterns and interactions with the original biotype. Such information is crucial to optimization of strategies for quarantine and control of resurgent pests. Our results demonstrate that while invasions are relatively rare events, new introductions can have significant ecological consequences, and quarantine vigilance is required even in previously invaded areas.

KEYWORDS: Invasion dynamics; biological control resistance; biotypes; *Oryctes rhinoceros*; *Oryctes* nudivirus; RADseq

INTRODUCTION

Understanding the pathways and ecological interactions governing biological invasions and post-invasion success are fundamental to excluding exotic species, mitigating their damage and anticipating future invasions. Biological control agents, including natural predators, parasitoids, and pathogens, have the potential to restore balance in disturbed ecosystems with minimal non-target impacts (Lacey *et al*. 2001; Hoddle 2004; Hajek *et al*. 2007). Evolution of resistance to biological control is thought to be less common than resistance to synthetic pesticides, because biocontrol agents exert weaker selective pressures, retain the potential for coevolution, and exhibit *modus operandi* that require multiple genome changes to counter (Jervis 1997; Holt & Hochberg 1997; Hufbauer & Roderick 2005; Tomasetto *et al*. 2017). Nevertheless, biocontrol-resistant pest biotypes have arisen in several

systems (e.g. Bravo *et al*. 2011; Di Giallonardo & Holmes 2015; Tomasetto *et al*. 2017), and the sustainability of biological control agents remains heavily debated (Lacey *et al*. 2001; Messing & Wright 2006; Glare *et al*. 2012; Lacey *et al*. 2015; Hajek *et al*. 2016). A detailed understanding of mechanisms by which an invasive pest species develops resistance to established biological control is a critical aspect of long-term management since the acquisition of resistance allows pests to resurge, resulting in range expansion and increasing damage (Owen & Zelaya 2005; Romero *et al*. 2007; Furlong *et al*. 2013).

Biotypes are conspecifics that appear similar but exhibit variation in one or more functional traits, such as host preference or pathogen resistance (Diehl & Bush 1984), and the term is now widely applied (Pekár *et al*. 2012; Majtánová *et al*. 2016; González-Torralva *et al*. 2017). When biotypes are identified in pest species, tracking and containment of these different types becomes critical for management. Biotypes with varying levels of resistance to control agents may carry other traits that facilitate range expansion and secondary invasion waves, as seen in both plants (Erfmeier *et al*. 2011) and insects (Horowitz *et al.* 2005; Delatte *et al*. 2009; Dennehy *et al*. 2010). In agricultural systems, for example, biotypic variation can facilitate expansion across a landscape of crop strains and cultivars (Hartman *et al.* 2001; Hill *et al*. 2010). When multiple biotypes invade a region, hybridization between previously controlled populations and new invaders can generate recombinant biotypes with enhanced plasticity, fostering increased virulence and invasiveness (Ellstrand & Schierenback 2000). Thus, even in previously invaded regions, novel biotype introductions may prove economically and environmentally damaging.

The coconut rhinoceros beetle *Oryctes rhinoceros* (Linnaeus) is a major palm pest which has been spread by human commerce and travel to many parts of the tropical Pacific and Indian Oceans (reviewed in Bedford 1980). Adult beetles feed on developing palm fronds, leading to nutrient deficiency, delayed maturity and, in cases of severe infestation,

tree death (Hinckley 1966; Hinckley 1973; Young 1975; Liau & Ahmad 1991). Yield loss in agricultural settings can be severe, with regional palm tree mortality reaching 50 to 100% (Gressitt 1953; Khoo *et al*. 1991; Dhileepan 1992; Manjeri *et al*. 2014). The beetle also damages tourist economies indirectly by degrading the palm-dependent tropical aesthetic that tourism industries rely on in destinations such as Guam and Hawaii (Smith and Moore 2008).

Originally native to a region between India and Indonesia, including the Philippines, parts of Southern China, and Taiwan (Catley 1969; Bedford 1980), *O. rhinoceros* has spread widely (Fig. 1). In the early 1900's, the beetle began invading Pacific island nations, including American Samoa (Jepson 1912; Bedford 1980), Palau (Gressitt 1953), and Fiji (Swaine 1966), facilitated by increased trade and military activity in the region. To control this spread, the *Oryctes* nudivirus (OrNV) was isolated from Malaysian populations and distributed in the early 1960's (Huger 2005; Bedford 2013). Successful establishment of the virus and declines in beetle damage throughout islands in the South Pacific and Indian Ocean lead to it being hailed as a biocontrol success, allowing many impacted regions to recover (Lacey *et al*. 2001; Huger 2005; Jackson 2009). However, in the last two decades, this biocontrol system has apparently broken down (Jackson *et al*. 2005; Jackson 2009), and there is a modern resurgence of *O. rhinoceros* including newly established populations in Guam (Smith and Moore 2008) and Oahu, Hawaii (Hawaii Dept. of Ag. 2014). In 2017, interceptions occurred as far east as Mexico (Jiménez *et al*. 2017). Coincidentally, an OrNV resistant *O. rhinoceros* biotype was recently described and identified throughout much of the beetle's native and invasive range (Marshall *et al*. 2017).

The resurgence of *Oryctes rhinoceros* provides a special opportunity to investigate the mechanisms and biogeographic patterns of novel resistance leading to a biological control breakdown after decades of successful suppression. Understanding the distribution of a resistant biotype and its interactions with established susceptible biotypes is essential to

elucidating the dynamics of biotypes and resistance in the context of controlling and preventing future invasions by species following the same paradigm. Ideally, to best understand this phenomenon, we would study a resurgent invasion associated with a novel biotype as it is occurring, revealing the patterns of genomic interaction between new, resistant, and older, susceptible, genotypes. Island systems provide an optimal substrate since timing of both initial and secondary invasion waves is available and islands provide inherent isolation between invaded regions. *Oryctes rhinoceros* presents a special case for understanding biotype-level invasion dynamics because the original invasion, subsequent control, and recent release from control are all well-documented. This beetle is especially relevant because it contains invasive populations that are still susceptible to the original biological control which can be contrasted with native range populations and invasive populations just now exhibiting resistance. Using both Sanger and genomic sequencing techniques, we examine the relationships between beetle populations across the native and invasive range of the insect to understand 1) How are beetle populations related to each other and what are the patterns of invasion? 2) What is the association between the resistant biotype and invasion resurgence? 3) What is the interaction between old susceptible and newly introduced, resistant, biotypes?

MATERIALS and METHODS

Sample Collection

We gathered 172 *O. rhinoceros* adults and larvae from the beetle's current range (Fig. 1), including samples from native populations in Thailand; Hainan Province, China; and Taiwan (Nantou and Kaohsiung counties), and introduced populations from the Republic of Palau; American Samoa; Guam; and Oahu, Hawaii (Supp. Table 1). All samples were collected by collaborators and were either placed immediately into ethanol or within 24 hours

after mortality. In some cases, pre-dissected adult or larval legs (2–4 legs per sample) were provided in 70–90% EtOH. Samples were stored at -20°C prior to DNA extraction.

DNA Extraction and Library Preparation

DNA extraction was performed on adult or larval legs using previously published methods (Sim and Geib 2017). A single leg, or two in the case of larvae, was homogenized in tissue lysis buffer using a 2010 Geno/Grinder Automated Tissue Homogenizer and Cell Lyser (SPEX SamplePrep, Metuchen, New Jersey, USA) for 30 seconds at 1500 rpm. The homogenate was then incubated in a 55°C water bath for 3 hours. Incubation was followed by extraction on a Kingfisher Flex 96 automated extraction instrument (Thermo Scientific, Waltham, MA) using standard protocols with a NucleoMag Tissue Kit (MACHEREY-NAGEL, Düren, Germany). The quantity and quality of the extracted DNA sample was determined using the High Sensitivity Genomic DNA Analysis Kit on a Fragment Analyzer (Advanced Analytical, Ankeny, IA).

Samples were used to construct double-digest restriction-site associated DNA libraries following Peterson *et al*. (2012). To prepare a library, 175 ng of DNA from each individual was digested using the restriction enzymes NlaIII and MluCI. One of 48 uniquely barcoded adapters was ligated to the restriction overhang, generating inline barcodes. Sub-pools of samples containing these 48 barcodes were generated and size-selected using a 1.5% agarose gel cassette on a Blue Pippin electrophoresis unit (Sage Science, Beverely, MA) with a target size selection of "narrow 400 bp". The final PCR amplification step was run for 10 cycles, during which a second barcode was added in the Illumina i7 location for each sub-pool, and PCR products were cleaned using solid-phase reversible immobilization (SPRI) beads at a 1.5:1 ratio of PEG containing bead solution to sample volume (DeAngelis *et al.*

1995; Rohland and Reich 2012). The final libraries were analyzed for quantity and size distribution using the NGS Fragment Analysis Kit on a Fragment Analyzer and pooled at equal molar ratios to generate the final library. Four libraries were created in this manner, including a total of 172 unique samples. The ddRAD libraries were each subjected to 100 bp single-end sequencing on a single lane of an Illumina HiSeq 4000 sequencer.

Read clustering and Loci and SNP selection

Twenty-one of the original 172 individual beetles were excluded from our analyses based on a 75% missing data threshold across all loci. Raw Illumina sequencing reads were then processed using STACKS v. 1.35 (Catchen *et al*. 2013) and IPYRAD v. 0.7.13 (Eaton & Overcast 2016) pipelines, which differ in locus assembly procedures. Whereas STACKS filtering is stricter, not allowing indels and utilizing an "off-by-N" similarity threshold when assembling loci, or "stacks" (Catchen *et al*. 2011), IPYRAD is more lenient, allowing for indels and variable numbers of nucleotide polymorphisms in the resulting loci (Eaton 2014; Eaton & Overcast 2016). We sought to recover loci informative across ecological and evolutionary timescales to fully explore the data, with STACKS theoretically geared to the former and IPYRAD the latter. Figure 2 provides an overview of the different datasets produced for subsequent population genetic and phylogenetic analyses. All data matrices are available at Dryad doi:10.5061/dryad.f4g56

Reads were processed using IPYRAD to conduct a *de novo* assembly of loci suited for comparisons across variable samples by allowing for indels and lower similarity thresholds in the retained loci (Eaton 2014). Due to this more extensive locus recovery and associated downstream computational limitations, the IPYRAD datasets were only produced for a 44-taxon subset selected by choosing the seven individuals for each region with the highest

overall raw read counts. All available individuals were included for Hainan (n=5) and Thailand (n=2). The selection was reiterated when we found that Palau exhibited multiple, distinct, population signatures, and the best individuals from each of these groups (minimum three) were included for a total of seven individuals (Fig. 2b&c; colored bars). The pipeline was completed using default parameter settings, which enforces minimum stack depth of six. IPYRAD was also run on the full set of 151 samples, but this information was only used to investigate overall locus recovery when processing the full set versus a subset of samples (Fig. 2b; grey versus pink bars). At a minimum of four shared taxa per locus (sh4), IPYRAD recovered 312,021 filtered loci compiled into both a full-loci data matrix (*IPY.sh4loci*) (Fig. 2b; pink bars) and a SNP data matrix (*IPY.sh4snp*) (Fig. 2b; blue bars). 209,493 sites were retained in the final matrix from polymorphic loci, selecting the site with the highest coverage. A secondary data matrix of loci was pursued with a strict locus-sharing parameter (*IPY.sh35loci*), such that a minimum of 35 shared taxa were required for a locus to be preserved (sh35). This resulted in a more evenly distributed coverage across samples, allowing us to test proportions of shared loci between populations.

In STACKS, *de novo* assembly of loci was completed using the *denovo_map.pl* wrapper program. For initial cataloging of loci, a minimum of 3 identical reads was required to form a "stack", allowing 2 mismatches between loci for a given individual during processing and 3 mismatches. The STACKS core program *populations* was used for additional filtering and to call a single SNP from each locus. SNP-level filters were used to extract variable sites, set minimum stack depth ($n > 10$) and set minimum proportion population participation (prop. pop. > 0.5) for any given locus. SNPs with call-rates < 80% were removed. The primary STACKS data matrix (*STA.151*) included 7,907 SNP sites for the full set of 151 retained samples (Fig. 2c). A secondary data matrix for the 44-taxon subset (*STA.44*) was also generated for comparison with the IPYRAD dataset. Both GENEIOUS v. 7.1.9 (Kearse *et al*.

2012) and PGDSPIDER v. 2.1.1.2 (Lischer & Excoffier 2011) were used to convert data between different file formats.

Biotype Assessment

Two *O. rhinoceros* biotypes are described in the literature, with biotype G exhibiting resistance to the *Oryctes* nudivirus while biotype S remains susceptible (Marshall *et al.* 2017). Marshall *et al*. (2017) identified a single diagnostic SNP able to discern *O. rhinoceros* biotype S and biotype G. Using previously described methods (Reil *et al*. 2016), we targeted this SNP – an A > G transition within the *COI* region – using primers LCO1490 (5'-GCT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') which amplify 658 base pairs including the polymorphism (Folmer *et al.* 1994). All individuals were genotyped, except for Oahu where a total of 60 samples were available, and the first 30 samples received were genotyped (see Supp. Table 1). For individuals shared between this and the Reil *et al*. (2016) study, Sanger sequences for the *COI* region were available. Individuals unique to this study were sequenced from extracts obtained for genomic library prep. Both forward and reverse primers were supplied for sense and antisense sequencing. Sequencing services were used for all samples, either through Eurofins (www.eurofins.com) or the ASGPB Laboratory at the University of Hawaii Manoa (www.hawaii.edu/microbiology/asgpb). We failed to obtain sequences for two individuals from Guam and four individuals from Hawaii. DNA sequences for all other individuals were aligned in GENEIOUS, followed by biotype assessment.

Shared loci and descriptive statistics

Using *IPY.sh4loci* and *IPY.sh35loci*, locus-sharing plots were generated in R v. 3.2.2 (R Core Team 2015) with the RADAMI v. 1.1-2 package (Hipp 2017) to visualize the proportions of shared loci among individuals and populations. A principle component analysis (Supp. Fig. 1) was conducted in R v. 3.2.2 on *STA.151* using the SNPRELATE package (Zheng *et al*. 2012). Genetic variability for each population was assessed on a per-locus basis by determining mean allele number (A), mean number of effective alleles (Ae), observed heterozygosity (Ho), expected heterozygosity (He), and unbiased expected heterozygosity (uHe) calculated in GENALEX v. 6.503 (Peakall & Smouse 2006, Peakall & Smouse 2012). Population pairwise Fst was also estimated in GENALEX using an AMOVA approach with *STA.151*.

Population assignment

Population structure was analyzed with the *STA.151* data matrix using both STRUCTURE v. 2.3.4 (Pritchard *et al*. 2000) and FASTSTRUCTURE v. 1.0 (Raj *et al*. 2014), the latter of which uses an algorithm adjusted for large SNP counts. In STRUCTURE, a range of K values 1 to 10 was explored using 100 runs for each value. FASTSTRUCTURE subsets – single replicates for a range of K values – were run 10 times for all K values 1 to 10 using a simple prior. The optimal K values for each subset were determined using FASTSTRUCTURE's built-in *chooseK* function. The most commonly designated K values were selected for a comprehensive analysis involving 100 clustering runs per value, again with a simple prior. For STRUCTURE, optimal cluster assignment was determined using Evanno's method (Evanno *et al*. 2005). Post-run processing and visualization was performed in R v. 3.4.1 using the POPHELPER package v. 2.2.1 (Francis 2017). CLUMPP (Jakobsson & Rosenberg 2007) was

used to realign clusters across multiple FASTSTRUCTURE and STRUCTURE runs, and to generate consensuses among the 100 runs for each of the optima values. Bar plots were constructed from the resulting consensus file.

Network and Phylogenetic Analyses

Using the *STA.151* data matrix, we performed a NeighborNet analysis (Bryant and Moulton 2004) implemented in SPLITSTREE4 v. 4.14.5 (Huson and Bryant 2006) to generate an unrooted genetic network. This method uses aspects of Neighbor-Joining (Saitou and Nai 1987) and SplitsTree to create a network that visualizes multiple hypotheses simultaneously. Phylogenetic trees were generated for the *STA.44*, *STA.151* and *IPY.sh4snp* data matrices using a maximum likelihood approach in RAXML v. 8.2.10 (Stamatakis 2014). RAXML was run on the CIPRES computing cluster (Miller *et al*. 2010) with the *RAxML-HPC2 Workflow on XSEDE* tool. Given that each of these data matrices used only a single SNP from each locus, data was analyzed under a GTRCAT model with site-specific evolutionary rate categories and a Lewis correction for ascertainment bias (Lewis 2001). Ambiguous and undefined sites within a data matrix can lead to apparently invariable sites being modeled as variable by RAXML; since ascertainment bias likelihood correction assumes a complete lack of invariable sites, these sites must be removed for RAXML to proceed without error. This was achieved in R v. 3.2.2 (R Core Team 2015) using the PHRYNOMICS v. 2.0 (Banbury & Leache 2014) and PHYLOTOOLS v. 0.1.2 (Zhang *et al*. 2012) packages to convert phylip formatted data matrices into "snp" objects which can be filtered to remove potentially invariable sites and converted back to phylip format for use in RAXML. Best tree calculations were performed 10 times, and the tree with the highest likelihood was chosen for optimization and use as the final best tree. Branch support was determined for each tree via bootstrapping under the same

model. The extended majority rule consensus criterion was used to terminate the bootstrapping process, up to a maximum of 1000 replicates. Unaddressed parameters were left at default settings.

Variation in maximum likelihood phylogenies for the different datasets indicated potential gene tree incongruity, specifically among *O. rhinoceros* from Palau. To investigate this, SVDQUARTETS analyses (Chifman & Kubatko 2014; Chifman & Kubatko 2015) implemented in PAUP* v. 4.0b (Swofford 2003) were performed on the CIPRES computing cluster (Miller *et al*. 2010) for *STA.44* and *IPY.sh4snp*. Quartet-based approaches are resilient to missing data and improve phylogenetic inference under varied coalescent models (Eaton *et al*. 2017). SVDQUARTETS analyzed 101,813 quartets (75% of combinations possible with 44 taxa) with branch support determined from 500 standard bootstrap replicates; remaining parameters were left at default settings.

Migration and Demographic History

TREEMIX analysis was used to assess gene flow and migration patterns between populations using the *STA.151* dataset for model scenarios from 0 to 10 migration events (numk = 0 to 10) (Pickrell & Pritchard 2012). Block size was set to 50; chromosome number was not used as a variable in our analysis. As the most centrally situated region in the *O. rhinoceros* native range, Thailand was selected as the outgroup for purposes of rooting. Variance was calculated using the script TreemixVarianceExplained.R (https://github.com) and graphed in R v. 3.2.2 to evaluate the optimal number of migrations necessary to explain the data. We used MOMENTS to explore demographic history, population growth patterns and migration in the two distinct Palau populations (Jouganous *et al.* 2017). A simple demographic model was constructed in MOMENTS which estimates the relative time the two populations diverged (T1), effective population sizes (nu1 and nu2) and relative migration

rates between the two populations (m12 and m21). Both simple and growth restricted models were tested. Pairwise comparisons were run for 100 replicates with random starting points for the five parameters estimated. From each model, the replicate with the highest log likelihood score was selected and its parameter estimates used as a starting point for 100 additional replicates. To estimate statistical support, we ran bootstrap analyses for each model using Godambe Information matrix implemented in MOMENTS.

RESULTS

Illumina sequencing generated 415,428,692 raw reads, with an average read count of 2,415,283 per individual. The IPYRAD pipeline, used on the 44 taxa subset, assembled the ddRAD reads into 471,066 pre-filtered loci. After filtering, 312,021 loci were retained and compiled into the three different IPYRAD datasets. The SNP site with the highest coverage from each polymorphic locus was included in the IPYRAD SNP data matrix comprised of 209,493 SNPs. STACKS assembled raw reads into 2,450,740 cataloged loci. The first variable site from each of the polymorphic loci that passed filtration was retained providing an alignment with 7,907 SNPs. STACKS produced smaller datasets on average due to more rigorous filtering and assembly but had more evenly distributed coverage across taxa (Fig. 2c). IPYRAD coverage appeared to be strongly correlated with raw read count for most individuals (Fig. 2b). The overall proportion of missing data for datasets *IPY.sh4snp*, *STA.151*, and *STA.44* is 65.8%, 28.1% and 27.6% respectively.

Population Structure

The optimal number of clusters for STRUCTURE runs was K=2 or 7 based on Evanno's method (Evanno *et al*. 2005). Under K=2, STRUCTURE included all individuals from Guam and Oahu in one cluster and placed the remaining populations into a second (Fig. 3), with minor admixture. At K=7, STRUCTURE assigns all but one invasive population (Palau), as well

as Taiwan and Hainan, to non-admixed, discrete clusters (Fig. 3). Individuals from Thailand had alleles admixed with at least five clusters, though they shared a majority of their diversity with the Hainan cluster. In Palau, five of twenty-three individuals formed a separate cluster with minor admixture from a second cluster that included the region's remaining individuals. Given this division, Palau was considered to have two populations, Palau population 1 (Palau S1, n=5) and Palau population 2 (Palau S2, n=18), in subsequent analyses.

FASTSTRUCTURE results mostly mirrored those of STRUCTURE, except at K=7 where there are notable differences in the assignment of individuals from Palau (Fig. 3). The five individuals in Palau S1 were primarily assigned to the broader Palau cluster. The cluster which defined Palau S1 in STRUCTURE was represented as admixture in these individuals encompassing about 25% of markers. Samples from Thailand exhibited lower marker diversity, and Hainan displayed admixture of a cluster that otherwise delineated the population on American Samoa.

Network and Phylogenetic Analyses

NeighborNet analysis identified eight clusters within *O. rhinoceros* (Fig. 4a). Guam and Oahu form distinct clusters sister to each other. American Samoa, Taiwan, Hainan, and Thailand form individual clusters, with Hainan and Thailand sister to each other and American Samoa sister to them. Palau splits into two distinct clusters matching the populations identified by the STRUCTURE analysis for K=7. Most of the variation illustrated by the tree network involves the position of Palau clusters. A series of networks place the Palau populations sister to each other in either recent (Fig. 4a; red branches) or deep (Fig. 4a; green branches) time, building on the disparity observed between FASTSTRUCTURE and STRUCTURE assignments for this region.

RAXML analyses yielded a variable set of unrooted phylogenies, depending on the data matrix used (summarized in Fig. 4b). Two major clades are conserved across the tree set: Hainan and Thailand as sister groups, with American Samoa sister to them, and the placement of Guam and Oahu as sister groups. Taiwan specimens occupy a central branch nearly equidistant from these two clades. Variation between topologies was generally focused around the placement of the Palau populations. In *IPY.sh4snp* and *STA.44* trees, the Palau populations are sister to each other, and this clade is placed sister to that of Hainan-Thailand-American Samoa. In the *STA.151* tree, Palau populations are split such that Palau S1 has a similar node position as in the 44-taxon trees, but Palau S2 is sister to Taiwan. In all variations, support is weak surrounding nodes which anchor Palau along the backbone topology.

Quartet analyses performed on the *STA.44* and *IPY.sh4snp* data matrices produced unrooted trees with different topologies (summarized in Fig. 4b). Branch support values highlighted regions of ambiguity in coalescent models between different genes. The *STA.44* SVDQUARTETS tree mirrored the maximum likelihood trees for *STA.44* and *IPY.sh4snp*, with poor support on branches around the Palau split (Branch Support < 70). On the *IPY.sh4snp* tree, Palau populations are split: Palau S1 is sister to the Hainan-Thailand-American Samoa clade and Palau S2 is sister to the Guam-Oahu clade. Taiwan sits between these groupings. The backbone of this tree was well supported for all clades (Branch Support > 90), in contrast to the other phylogenies produced by this study which exhibited weak support (Branch Support < 90) on branches associated with either of Palau's two populations (Fig 4b).

Demographic History

TREEMIX analysis revealed an admixture event between Palau S1 and Palau S2 that improved the tree, increasing variance explained by the model by 8.97% (Fig. 5a&b). Assuming additional admixture events beyond the first did not strengthen the model substantially (0.2% improvement or less). MOMENTS software further reinforced a migration pattern from Palau S1 to Palau S2 (Fig. 5c). Effective population sizes were less than one individual but retain relativistic value. Migration from Palau S1 into Palau S2 (m12) was more extreme than migration in the opposite direction, with an approximately twenty times greater migration rate from Palau S1 to Palau S2 (m12 = 19.98), than from Palau S2 to Palau S1 (m21 = 1.101).

Genetic Variability and Population Differentiation

Populations that diverge from a source population through a bottleneck event will represent a subset of alleles from the source population. We found however that the shared loci analyses of the least-filtered dataset (*IPY.sh4loci*) only reflect the locus coverage per sample, and in the more filtered dataset, with comparable per sample loci coverage (*IPY.sh35loci*), there are no shared locus differences that remain between populations (Supp. Fig. 2 & 3). Genetic variability and diversity statistics performed on *STA.151* revealed that in all populations except Thailand, realized heterozygosity is lower than the sample-size corrected, unbiased expected heterozygosity (Table 1). Hainan and Palau S1 exhibited heterozygosity over the expected values but not over unbiased expected values. Fst values indicate that Guam and Oahu have the highest genetic similarity (Fst = 0.124) among populations sampled (Table 2). Remaining pairwise comparisons suggest less strongly supported patterns. Palau S2 is genetically similar to Guam (Fst = 0.224), Oahu (Fst = 0.237), and Palau S1 (0.240). Palau S2 is the next most similar population for both Guam and Oahu.

Palau S1 is most genetically similar to Palau S2, followed by Guam (Fst = 0.293) and Oahu (Fst = 0.354). American Samoa is most similar to Guam (Fst = 0.228) followed by Thailand (0.241). American Samoa is most similar to Thailand, followed by Guam (Fst = 0.270) and Hainan (Fst = 0.289).

Biotype Mapping

Biotype S or G designations were successfully obtained for 119 samples based on the *COI* SNP (Supp. Table 1). Mapping biotypes to phylogenetic results revealed single-biotype populations in all locations except Palau (Fig. 4b). Palau S1 included only biotype S individuals; Palau S2 included both biotypes dispersed throughout the clade with no apparent pattern (Fig. 4). Populations in Taiwan, Guam, and Hawaii are all biotype G (OrNV resistant). Populations in American Samoa, Hainan and Thailand are all biotype S (OrNV susceptible).

DISCUSSION

Invasion pathways linked to biocontrol resistance

Genomics-based reconstruction of *O. rhinoceros* invasion waves in the Pacific indicates a strong relationship between the G biotype and the resurgence of invasiveness. While the OrNV susceptible S beetle biotype is prominent in regions that were invaded earlier, the invasions of the last two decades were all by the nudivirus-resistant G biotype. This demonstrates the potential for rapid invasion by undetected, resistant biotypes that allow pests to suddenly escape previously effective biological control measures. Furthermore, our results suggest hybridization between beetle biotypes, leading to introgression.

Oryctes rhinoceros invasion pathways across the Pacific are obscured by a lack of variation within and between populations, likely caused by intense bottleneck events, which are evidenced by the absence of detectable variation in locus representation. Nonetheless, we were able to clearly differentiate genetic signatures of the beetle's two documented invasion waves: one from the early- to mid-20th century, prior to the rise of *Oryctes* nudivirus as a control agent in the 1960's (Catley 1969; Bedford 1980), and the modern wave which resulted in newly established beetle populations on Guam and Hawaii, and secondary invasion of Palau, over the last two decades (Marshall et al. 2017). All analyses point to a stepping-stone invasion pattern from Guam into Hawaii. The precursory invasion of islands between mainland Asia and the Hawaiian archipelago is a common pattern for other species as well, with the intervening islands facilitating the trans-Pacific invasion of pests (Moore 2012; Vargas *et al*. 2015).

Unidirectional biotype introgression

Our results demonstrate that Palau supports two genetically distinct populations of *O. rhinoceros* – one corresponding to the initial invasion wave with biotype S beetles and one corresponding to the recent invasion with biotype G – with gene-flow between them. TREEMIX analysis supports introgression, with an admixture event between Palau S1 and Palau S2 substantially increasing variance explained by the model. The ramifications of this introgression between beetle biotypes in Palau are yet unclear. Currently, the introgression appears to be unidirectional. We found *COI* S-type individuals throughout the Palau S2 clade in the genomic data, despite it originating from the recent biotype G introduction, and no evidence of G genotypes occurring in the S-typed (older introduction) Palau S1 population. This directionality is supported both by TREEMIX and MOMENTS which demonstrated gene flow from Palau S1 to Palau S2 with only minor contributions in the opposite direction. Our genomic analysis of *O. rhinoceros* provides fresh evidence supporting the paradigm that new

introductions can improve the adaptive potential of old pests by boosting genetic diversity (Lambrinos 2004; Kolbe *et al*. 2004; Lavergne & Molofsky 2007; Roman & Darling 2007) and suggest a mechanism by which hybridization may generate recombinant biotypes with enhanced invasiveness and severity (Ellstrand & Schierenback 2000; Anderson *et al.* 2018). While no other biotype G populations in the Pacific exhibited nudivirus infection, some biotype G individuals on Palau had measurable viral loads (Marshall et al. 2017); this points towards adaptation for viral tolerance, and not simply resistance, as a mechanism for biotype G's bolstered invasiveness. Tolerance and resistance are not equivalent traits, with the former more readily achieving fixation in a population (Roy & Kirchner 2000).

Although we cannot differentiate the influences of introgression from a novel evolutionary trajectory, the potential for multiple introductions to enhance adaptability remains a subject for further study. Regardless, there is an immediate need for the development of new robust methods for screening these biotypes. The apparent introgression is likely to jeopardize the reliability of the current diagnostic *COI* genotype assays due to phenomena such as mito-nuclear discordance (e.g. Dupont et al. 2016). Mito-nuclear discordance is known to be caused by asymmetric movement of mtDNA through sex-biased, demographic, and adaptive processes (reviewed in Toews & Brelsford 2012). Because introgression is only detected in the direction of biotype S mitochondria associating with nuclear DNA from the new, resistant, invasive G strain, and not vice-versa, this suggests resistant male beetles may be outcompeting, or are more numerous than, non-resistant males on Palau, causing a mating bias. *Wolbachia* infections are also known to cause mito-nuclear discordance in insects (Gompert *et al.* 2008; Smith *et al.* 2012), but there is no evidence to date supporting rampant infection by these microbial endosymbionts among the Scarabaeidae. As the first location with demonstrated introgression of *O. rhinoceros* biotypes, Palau has proven to be a crossroads between the distinct invasion waves. Intensified

quarantine throughout the region, and particularly in Palau, is paramount, as these processes appear to be actively ongoing. Escaped individuals from regions experiencing introgression will be even more difficult to track and diagnose genetically and may lead to cascading resistance in other regions.

Invasions in the native range?

Our results reveal inconsistencies in genomic diversity between native and invasive ranges of an invasive species, including the identification of invaded areas which were previously considered part of the native range. While *O. rhinoceros* populations in Taiwan were always considered native (Bedford 1980), we found low diversity in this region, suggesting an older undocumented invasion (see Dlugosch & Parker 2008), or displacement of the native population by an invasive biotype. Misinterpretation of native status can distort our fundamental understanding of a pests' evolution, biology, and history, forestalling management, thus facilitating reintroduction and spread (e.g. Holsbeek *et al*. 2008; Liu *et al*. 2007; Gao *et al*. 2012). Genetic diversity in native areas will have to be re-assessed to develop effective biotype-based quarantine protocols to prevent genetic swamping – as is currently ongoing in Palau – that can result in spreading resistance.

Origin of a resistant biotype

The OrNV was first discovered in beetles in Malaysia (Huger 2005), but it is unclear where and when the G biotype first appeared. The biotype may have existed for much of the species' history but only recently escaped isolation due to developing human infrastructure and globalization-mediated release (e.g. Lees *et al*. 2011). The first indicators of potential viral resistance came from repeated *O. rhinoceros* outbreaks in Java and South Sulawesi in the mid-1980s (Holt & Hochberg 1997). Recent inventories using the *COI* resistance marker found biotype G in the native range in both Indonesia and the Philippines (Marshall *et al.*

2017). In Indonesia, biotype S and biotype G were found sympatrically, while in the Philippines all 12 beetle samples were G-type. In the context of this investigation, the lack of shared allelic diversity between Palau and other biotype G populations, coupled with phylogenetic position of samples from this region, suggest that Palau is likely not the source of this resistant strain. Of the regions sampled for this study, Taiwan was the only biotype G dominated region which the literature places in the beetle's native range, and it shares mitochondrial haplotypes with recently established populations (Reil *et al*. 2016). Because the resistant strain is dominant in Taiwan, it appears the most likely source for G-type beetles in the recent invasions of regions we sampled. Still, a more comprehensive sampling of *O. rhinoceros'* native range would allow more finite conclusions to this end.

Impact of analysis method on introgression and secondary invasion detection

The use of genomic data for the reconstruction of invasion pathways is relatively new and the analyses methods are in development, making our comparison of two pipelines broadly relevant. Considering IPYRAD and STACKS for ddRAD sequenced loci assembly and SNP selection, we found the former to produce more extensive (max: 209,493 SNPs) datasets more effective at revealing deeper evolutionary relationships despite uneven data coverage. In contrast, the STACKS datasets were more evenly distributed, contained fewer loci (max: 7,907) and were sensitive to recent introgression, but missed deeper signals of isolation. A combination of IPYRAD's more 'relaxed' assembly and a quartet approach, which is robust to the incongruence of individual locus or gene trees, enabled us to identify the hybridizing coalescent biotype lineages in Palau and revealed the more recent introduction of biotype G during the modern invasion phase. This may be an important consideration for other studies in which multiple signals are present. Additionally, we found that STRUCTURE delimited two distinct *O. rhinoceros* populations in Palau without extensive gene-flow, whereas

FASTSTRUCTURE detected recent hybridization by assigning these groups to largely the same cluster.

Management implications

The propagation of the *Oryctes* nudivirus as a biocontrol agent may have catalyzed the recent success of biotype G over biotype S. This, coupled with anthropogenic movement of *O. rhinoceros* in recent decades, appears to have driven the resistant biotype's 21st-century dispersal. Recent interceptions in Mexico extend this pest's ongoing threat to the Neotropics. *Oryctes rhinoceros* resistance to OrNV, detected as early as 1989, is one of the first clear cases of biological control resistance associated with a diagnosable biotype (Zelazny *et al*. 1989). In the broader context of pest invasion dynamics, our results suggest that biological control resistance can not only promote pest recovery, but also the resurgence of invasive vagility, leading to a new invasion wave nearly half a century after control efforts began.

Despite concerns over robustness and longevity, the use of biocontrol agents is steadily increasing, encouraged by the trend towards pest management practices with minimal non-target and environmental impacts (Glare *et al*. 2012; Lacey *et al*. 2015; Siegwart *et al*. 2015; Cordeau *et al*. 2016). While biological control agents have clear benefits, our results reveal the mechanism of a potential breakdown, and emphasize the importance of continued quarantine against pests, even those already thought to be present in a region. Novel pest bio- and genotypes appear, with the potential to enhance outbreak damage in both invasive and native ranges (e.g. Delatte *et al*. 2009; Stastny & Sargent 2017). In the case of *O. rhinoceros*, it is unclear whether OrNV resistance is the sole phenotypic factor contributing to biotype G's invasion success, but a new approach to monitoring ongoing changes in invasive pest genomics is essential as globalization rapidly facilitates the spread of novel, virulent biotypes. Old pests can become new threats; as species we have

dispersed and managed for decades continue to evolve resistance, the resurgent invasion risk is only going to increase.

ACKNOWLEDGEMENTS

We thank Darcy Oishi, Michael Melzer, Shizu Watanabe, Megan Manley, Daniel Nitta, Will Haines, Nicole Yoneishi, Angela Kauwe, Cheng Zhiqiang, Aubrey Moore, Yolisa Ishibashi, Mark Schmaedick, Hong Thai Pham, Meredith Metotagivale, Niela Leifi, Lanor Laulu, Domenick Skabeikis, Joel Miles, Nelson Masang Jr., Srishti Sami Soni, Christopher Kitalong, Li-Cheng Shih (Chubby), Hsiang-Yu Chuang (Nomis), Yu-Chi Lin (Chichi), Wei Yan, Frank Hsu, Yu-Feng Shu, Chia-Lung Huang (Biogun), and the Hawaii Department of Agriculture for providing *O. rhinoceros* samples, assisting with collection and/or otherwise helpful contributions. Illumina Sequencing was conducted at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10OD018174. This study was funded, in part, by the USDA Farm Bill, USDA National Institute of Food and Agriculture (NIFA), Agriculture and Food Research Initiative (AFRI), Education and Literacy Initiative (ELI) 2017-67012-26087 (FY17-FY19) which was awarded to and partially supported SBS, the ARCS Foundation, and by the University of Hawaii at Manoa College of Tropical Agriculture and Human Resources and USDA Cooperative State Research, Education and Extension (CSREES) project HAW00942-H administered by the College of Tropical Agriculture and Human Resources, University of Hawaii at Manoa.

REFERENCES

Anderson, C. J., Oakeshott, J. G., Tay, W. T., Gordon, K. H., Zwick, A., & Walsh, T. K. (2018). Hybridization and gene flow in the mega-pest lineage of moth, Helicoverpa. *Proceedings of the National Academy of Sciences*, 201718831.

Banbury, B., & Leache, A. (2014) phrynomics: An SNP toolkit. R package version 2.0.

Bedford, G. O. (1980). Biology, ecology, and control of palm rhinoceros beetles. *Annual Review of Entomology*, *25*(1), 309-339.

Bedford, G. O. (2013). Biology and management of palm dynastid beetles: recent advances. *Annual review of entomology*, *58*, 353-372.

Bravo, A., Likitvivatanavong, S., Gill, S. S., & Soberón, M. (2011). Bacillus thuringiensis: a story of a successful bioinsecticide. *Insect biochemistry and molecular biology*, *41*(7), 423-431.

Bryant, D., & Moulton, V. (2004). NeighborNet: an agglomerative method for the construction of phylogenetic networks. *Molecular biology and evolution*, *21*(2), 255-265.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, genomes, genetics*, *1*(3), 171-182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: an analysis tool set for population genomics. *Molecular ecology*, *22*(11), 3124-3140.

Catley, A. (1969). The coconut rhinoceros beetle Oryctes rhinoceros (L) [Coleoptera: Scarabaeidae: Dynastinae]. *PANS Pest Articles & News Summaries*, *15*(1), 18-30.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*(23), 3317-3324.

Chifman, J., & Kubatko, L. (2015). Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of theoretical biology*, *374*, 35-47.

Cordeau, S., Triolet, M., Wayman, S., Steinberg, C., & Guillemin, J. P. (2016). Bioherbicides: Dead in the water? A review of the existing products for integrated weed management. *Crop protection*, *87*, 44-49.

DeAngelis, M. M., Wang, D. G., & Hawkins, T. L. (1995). Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic acids research*, *23*(22), 4742.

Delatte, H., Duyck, P. F., Triboire, A., David, P., Becker, N., Bonato, O., & Reynaud, B. (2009). Differential invasion success among biotypes: case of Bemisia tabaci. *Biological Invasions*, *11*(4), 1059-1070.

Dennehy, T. J., Degain, B. A., Harpold, V. S., Zaborac, M., Morin, S., Fabrick, J. A., ... & Li, X. (2010). Extraordinary resistance to insecticides reveals exotic Q biotype of Bemisia tabaci in the New World. *Journal of economic entomology*, *103*(6), 2174-2186.

Dhileepan, K. (1992). Insect pests of oil palm (Elaeis guineensis) in India. *Planter*, *68*(7 93), 183-191.

Diehl, S. R., & Bush, G. L. (1984). An evolutionary and applied perspective of insect biotypes. *Annual review of entomology*, *29*(1), 471-504.

Di Giallonardo, F., & Holmes, E. C. (2015). Viral biocontrol: grand experiments in disease emergence and evolution. *Trends in microbiology*, *23*(2), 83-90.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular ecology*, *17*(1), 431-449.

Dupont, L., Porco, D., Symondson, W. O. C., & Roy, V. (2016). Hybridization relics complicate barcode-based identification of species in earthworms. *Molecular ecology resources*, *16*(4), 883-894.

Eaton, D. A. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics, 30*(13), 1844-1849.

Eaton, D. A., & Overcast, I. (2016). ipyrad: interactive assembly and analysis of RADseq data sets. Available at: http://ipyrad.readthedocs.io/.

Eaton, D. A., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic biology*, *66*(3), 399-412.

Ellstrand, N. C., & Schierenbeck, K. A. (2000). Hybridization as a stimulus for the evolution of invasiveness in plants?. *Proceedings of the National Academy of Sciences*, *97*(13), 7043-7050.

Erfmeier, A., Böhnke, M., & Bruelheide, H. (2011). Secondary invasion of Acer negundo: the role of phenotypic responses versus local adaptation. *Biological Invasions*, *13*(7), 1599-1614.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611-2620.

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*, *3*(5), 294-9.

Francis, R. M. (2017). Pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resources, 17*(1), 27-32.

Furlong, M. J., Wright, D. J., & Dosdall, L. M. (2013). Diamondback moth ecology and management: problems, progress, and prospects. *Annual review of entomology*, *58*, 517-541.

Gao, Y., Reitz, S. R., Wei, Q., Yu, W., & Lei, Z. (2012). Insecticide-mediated apparent displacement between two invasive species of leafminer fly. *PLoS One*, *7*(5), e36622.

Glare, T., Caradus, J., Gelernter, W., Jackson, T., Keyhani, N., Köhl, J., ... & Stewart, A. (2012). Have biopesticides come of age?. *Trends in biotechnology*, *30*(5), 250-258.

Gompert, Z., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2008). Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in Lycaeides. *Molecular ecology*, *17*(24), 5231-5244.

González-Torralva, F., Brown, A. P., & Chivasa, S. (2017). Comparative proteomic analysis of horseweed (Conyza canadensis) biotypes identifies candidate proteins for glyphosate resistance. *Scientific reports*, *7*, 42565.

Gressitt, J. L. (1953). The coconut rhinoceros beetle (Oryctes rhinoceros) with particular reference to the Palau Islands. *Bulletin of the Bernice P. Bishop Museum*, (212), 157.

Hajek, A. E., McManus, M. L., & Junior, I. D. (2007). A review of introductions of pathogens and nematodes for classical biological control of insects and mites. *Biological Control*, *41*(1), 1-13.

Hajek, A. E., Hurley, B. P., Kenis, M., Garnas, J. R., Bush, S. J., Wingfield, M. J., ... & Cock, M. J. (2016). Exotic biological control agents: A solution or contribution to arthropod invasions?. *Biological invasions*, *18*(4), 953-969.

Hawaii Department of Agriculture. (2014). Destructive beetles found on Oahu coconut trees. Department of Agriculture, State of Hawaii. (http://hdoa.hawaii.gov/blog/main/crb/).

Hartman, G. L., Domier, L. L., Wax, L. M., Helm, C. G., Onstad, D. W., Shaw, J. T., ... & Steffey, K. L. (2001). Occurrence and distribution of Aphis glycines on soybeans in Illinois in 2000 and its potential control. *Plant Health Prog*, *10*.

Hill, C. B., Crull, L., Herman, T. K., Voegtlin, D. J., & Hartman, G. L. (2010). A new soybean aphid (Hemiptera: Aphididae) biotype identified. *Journal of Economic Entomology*, *103*(2), 509-515.

Hinckley, A. D. (1966). Damage by the rhinoceros beetle, *Oryctes rhinoceros* (L.), to Pacific island palms. *South Pacific Bulletin*, *16*(4), 51-52.

Hinckley, A. D. (1973). Ecology of the coconut rhinoceros beetle, Oryctes rhinoceros (L.)(Coleoptera: Dynastidae). *Biotropica*, 111-116.

Hipp, A. (2017). RADami: Phylogenetic Analysis of RADseq Data. *R package version 1.1-2*. https://CRAN.R-project.org/package=RADami.

Hoddle, M. S. (2004). Restoring balance: using exotic species to control invasive exotic species. *Conservation Biology*, *18*(1), 38-49.

Holt, R. D., & Hochberg, M. E. (1997). When is biological control evolutionarily stable (or is it)?. *Ecology*, *78*(6), 1673-1683.

Horowitz, A. R., Kontsedalov, S., Khasdan, V., & Ishaaya, I. (2005). Biotypes B and Q of Bemisia tabaci and their relevance to neonicotinoid and pyriproxyfen resistance. *Archives of insect biochemistry and physiology*, *58*(4), 216-225.

Hufbauer, R. A., & Roderick, G. K. (2005). Microevolution in biological control: mechanisms, patterns, and processes. *Biological control*, *35*(3), 227-239.

Huger, A. M. (2005). The Oryctes virus: its detection, identification, and implementation in biological control of the coconut palm rhinoceros beetle, Oryctes rhinoceros (Coleoptera: Scarabaeidae). *Journal of invertebrate pathology*, *89*(1), 78-84.

Huson, D. H. & Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, *23*(2), 254-267. Software available from www.splitstree.org.

Jackson, T. A. (2009). The use of Oryctes virus for control of rhinoceros beetle in the Pacific Islands. In *Use of Microbes for Control and Eradication of Invasive Arthropods* (pp. 133-140). Springer, Dordrecht.

Jackson, T. A., Crawford, A. M., & Glare, T. R. (2005). Oryctes virus—time for a new look at a useful biocontrol agent. *Journal of invertebrate pathology*, *89*(1), 91-94.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801-1806.

Jepson, F. P. (1912). The rhinoceros beetle *(Oryctes rhinoceros, Linn.)* in Samoa. *Fiji Department of Agriculture Bulletin*, (3*)*, 25.

Jervis, M. A. (1997). Parasitoids as limiting and selective factors: can biological control be evolutionarily stable?. *Trends in ecology & evolution*, *12*(10), 378-380.

Jiménez, Q., Martínez, M., Trejo, R., González, V., Guerrero, A., & Chávez, N. (2017). First intercept of the Asiatic coconut rhinoceros beetle *Oryctes rhinoceros* (Linnaeus, 1758) in Mexico. *Revista Mexicana de Ciencias Forestales*, *8*(39), 99-105.

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. (2017). Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. *Genetics*, genetics-117 (Early Online).

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647-1649.

Khoo, K. C., Ooi, P. A. C., & Ho, C. T. (1991). *Crop pests and their management in Malaysia*. Tropical Press.

Kolbe, J. J., Glor, R. E., Schettino, L. R., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, *431*(7005), 177-181.

Lacey, L. A., Frutos, R., Kaya, H. K., & Vail, P. (2001). Insect pathogens as biological control agents: do they have a future?. *Biological control*, *21*(3), 230-248.

Lacey, L. A., Grzywacz, D., Shapiro-Ilan, D. I., Frutos, R., Brownbridge, M., & Goettel, M. S. (2015). Insect pathogens as biological control agents: back to the future. *Journal of invertebrate pathology*, *132*, 1-41.

Lambrinos, J. G. (2004). How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology*, *85*(8), 2061-2070.

Lavergne, S., & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences*, *104*(10), 3883-3888.

Lees, D. C., Lack, H. W., Rougerie, R., Hernandez-Lopez, A., Raus, T., Avtzis, N. D., ... & Lopez-Vaamonde, C. (2011). Tracking origins of invasive herbivores through herbaria and archival DNA: the case of the horse-chestnut leaf miner. *Frontiers in Ecology and the Environment*, *9*(6), 322-328.

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic biology*, *50*(6), 913-925.

Liau, S.S., Ahmad, A. (1991) The control of Oryctes rhinoceros by clean clearing and its effect on early yield in palm to palm replants. *Proceedings PORIM International Palm Oil Development Conference*, *1*, 396-403.

Lischer, H. E., & Excoffier, L. (2011). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298-299.

Liu, S. S., De Barro, P. J., Xu, J., Luan, J. B., Zang, L. S., Ruan, Y. M., & Wan, F. H. (2007). Asymmetric mating interactions drive widespread invasion and displacement in a whitefly. *Science*, *318*(5857), 1769-1772.

Majtánová, Z., Choleva, L., Symonová, R., Ráb, P., Kotusz, J., Pekárik, L., & Janko, K. (2016). Asexual reproduction does not apparently increase the rate of chromosomal evolution: karyotype stability in diploid and triploid clonal hybrid fish (Cobitis, Cypriniformes, Teleostei). *PloS one*, *11*(1), e0146872.

Manjeri, G., Muhamad, R., & Tan, S. G. (2014). Oryctes rhinoceros beetles, an oil palm pest in Malaysia. *Annual Research & Review in Biology*, *4*(22), 3429.

Marshall, S. D., Moore, A., Vaqalo, M., Noble, A., & Jackson, T. A. (2017). A new haplotype of the coconut rhinoceros beetle, Oryctes rhinoceros, has escaped biological control by Oryctes rhinoceros nudivirus and is invading Pacific islands. *Journal of Invertebrate Pathology, 149*, 127-134.

Messing, R. H., & Wright, M. G. (2006). Biological control of invasive species: solution or pollution?. *Frontiers in Ecology and the Environment*, *4*(3), 132-140.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE),* 14 Nov. 2010, New Orleans, LA (pp. 1-8).

Moore, A. (2012, February). Guam as a Source of New Insects for Hawaii. In, *Current and Emerging Insect Pests in Hawai'i and the Pacific*. Symposium conducted at the XIV Pacific Entomology Conference, Challenges and Changes in Entomology, Honolulu, Hawaii.

Owen, M. D., & Zelaya, I. A. (2005). Herbicide-resistant crops and weed resistance to herbicides. *Pest management science*, *61*(3), 301-311.

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

Peakall, R. & Smouse P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics 28*, 2537-2539.

Pekár, S., Šmerda, J., Hrušková, M., Šedo, O., Muster, C., Cardoso, P., ... & Sentenská, L. (2012). Prey-race drives differentiation of biotypes in ant-eating spiders. *Journal of Animal Ecology*, *81*(4), 838-848.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS one*, *7*(5), e37135.

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS genetics*, *8*(11), e1002967.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raj, A., Stephens, M., & Pritchard, J. K. (2014) fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics,197*, 573-589.

Reil, J. B., San Jose, M., & Rubinoff, D. (2016). Low Variation in Nuclear and Mitochondrial DNA Inhibits Resolution of Invasion Pathways across the Pacific for the Coconut Rhinoceros Beetle (Scarabeidae: Oryctes rhinoceros). *Proc. of the Hawaiian Ent. Soc., 48*, 57-69.

Rohland, N., & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research*, *22*(5), 939-946.

Roman, J., & Darling, J. A. (2007). Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, *22*(9), 454-464.

Romero, A., Potter, M. F., Potter, D. A., & Haynes, K. F. (2007). Insecticide resistance in the bed bug: a factor in the pest's sudden resurgence?. *Journal of medical entomology*, *44*(2), 175-178.

Roy, B. A., & Kirchner, J. W. (2000). Evolutionary dynamics of pathogen resistance and tolerance. *Evolution*, *54*(1), 51-63.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, *4*(4), 406-425.

Siegwart, M., Graillot, B., Blachere Lopez, C., Besse, S., Bardin, M., Nicot, P. C., & Lopez-Ferber, M. (2015). Resistance to bio-insecticides or how to enhance their sustainability: a review. *Frontiers in plant science*, *6*, 381.

Sim, S. B., & Geib, S. M. (2017). A Chromosome-Scale Assembly of the Bactrocera cucurbitae Genome Provides Insight to the Genetic Basis of White Pupae. *G3: Genes, Genomes, Genetics*, g3-117.

Smith, M. A., Bertrand, C., Crosby, K., Eveleigh, E. S., Fernandez-Triana, J., Fisher, B. L., ... & Hrcek, J. (2012). Wolbachia and DNA barcoding insects: patterns, potential, and problems. *PloS one*, *7*(5), e36514.

Smith, S. L., & A. Moore. 2008. Early detection pest risk assessment: coconut rhinoceros beetle. (http://guaminsects.net/uogces/kbwiki/images/1/13/CRB_Pest_Risk_Assessment.pdf). Google Scholar

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313.

Stastny, M., & Sargent, R. D. (2017). Evidence for rapid evolutionary change in an invasive plant in response to biological control. *Journal of evolutionary biology*, *30*(5), 1042-1052.

Swaine, G. (1966). Fiji's campaign against the coconut rhinoceros beetle. *International Pest Control*, *8*, 6-9.

Swofford, D. L. (2003). PAUP*: phylogenetic analysis using parsimony, version 4.0 b10.

Tomasetto, F., Tylianakis, J. M., Reale, M., Wratten, S., & Goldson, S. L. (2017). Intensified agriculture favors evolved resistance to biological control. *Proceedings of the National Academy of Sciences*, *114*(15), 3885-3890.

Toews, D. P., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*(16), 3907-3930.

Vargas, R. I., Piñero, J. C., & Leblanc, L. (2015). An overview of pest species of Bactrocera fruit flies (Diptera: Tephritidae) and the integration of biopesticides with other biological approaches for their management with a focus on the Pacific Region. *Insects*, *6*(2), 297-318.

Young, E. C. (1975). A study of rhinoceros beetle damage in coconut palms. *Technical Paper, South Pacific Commission,* (170), 63.

Zelazny, B., Alfiler, A. R., and Lolong, A. (1989) Possibility of resistance to a baculovirus in populations of coconut rhinoceros beetle (Oryctes rhinoceros). *Boletin Fitosanitario de la FAO (FAO)*.

Zhang, J., Pei, N., & Mi, X. (2012) phylotools: Phylogenetic tools for Eco-phylogenetics. R package version 0.1.2. http://CRAN.R-project.org/package=phylotools.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326-3328.

Data Accessibility:

-Raw ddRAD Sequencing Reads. NCBI SRA accession **SRP132406**

-*COI* Sanger Sequences for Biotyping samples exclusive to this report. GenBank Accessions **MH800532-MH800566** (*in review to be released upon completion*)

-*COI* Sanger Sequences for Biotyping samples shared and accessioned earlier for use in Reil et al. 2017. GenBank Accessions **KY197991, KY197993, KY197995, KY197997-KY198028, KY198030-KY198069, KY198075-KY198083, KY198085, KY198087, KY198088, KY198093-KY198095, KY198098-KY198107**

-A selection of relevant scripts used in R and Python for data processing, visualization and analysis. **Dryad DOI: https://doi.org/10.5061/dryad.f4g56**

-Raw tree files used in the assembly of figure 4. **Dryad DOI: https://doi.org/10.5061/dryad.f4g56**

-Data matrices and alignments generated from STACKS and IPYRAD, in the case of ddRAD data, and from GENEIOUS, in the case of Sanger data, used for project analyses. PGDSPIDER was used for conversion between file formats when necessary. A VCF file containing all SNPs for all loci identified and cataloged by IPYRAD has also been included. **Dryad DOI: https://doi.org/10.5061/dryad.f4g56.**

Author Contributions:

Authors listed made the following contributions. J.R., S.S., M.S, S.G project conception and development; J.R., S.S., and S.G. sample preparation and laboratory work; J.R. performed data processing and analysis with input and instruction from S.S., M.S., S.G. and C.D.; J.R., S.S., M.S., C.D., S.G., and D.R. data interpretation; J.R. wrote the paper with input from C.D. and the other authors.

APPENDIX A: TABLES AND FIGURES

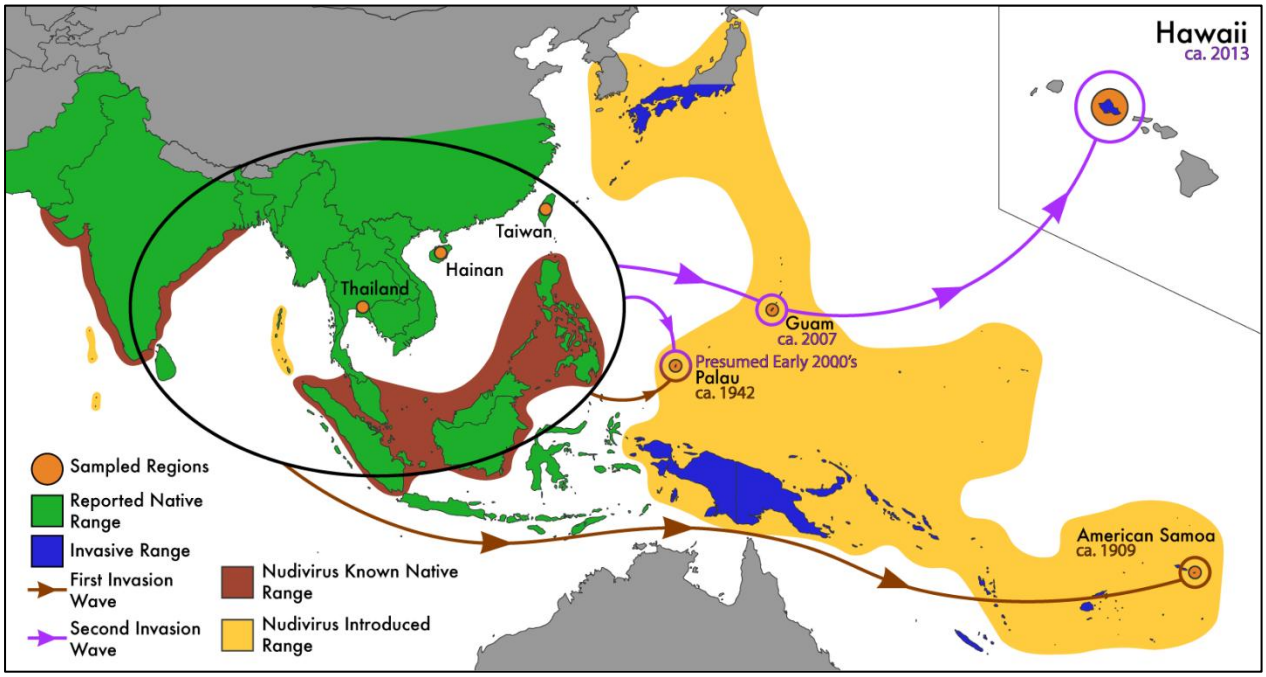

Figure 1. Map of the Pacific and Indian oceans displaying rhinoceros beetle sampling location, basic details of known invasion events, and *Oryctes* nudivirus distribution. Areas in green are in the beetle's native range; regions in dark blue include areas in the beetle's invasive range.

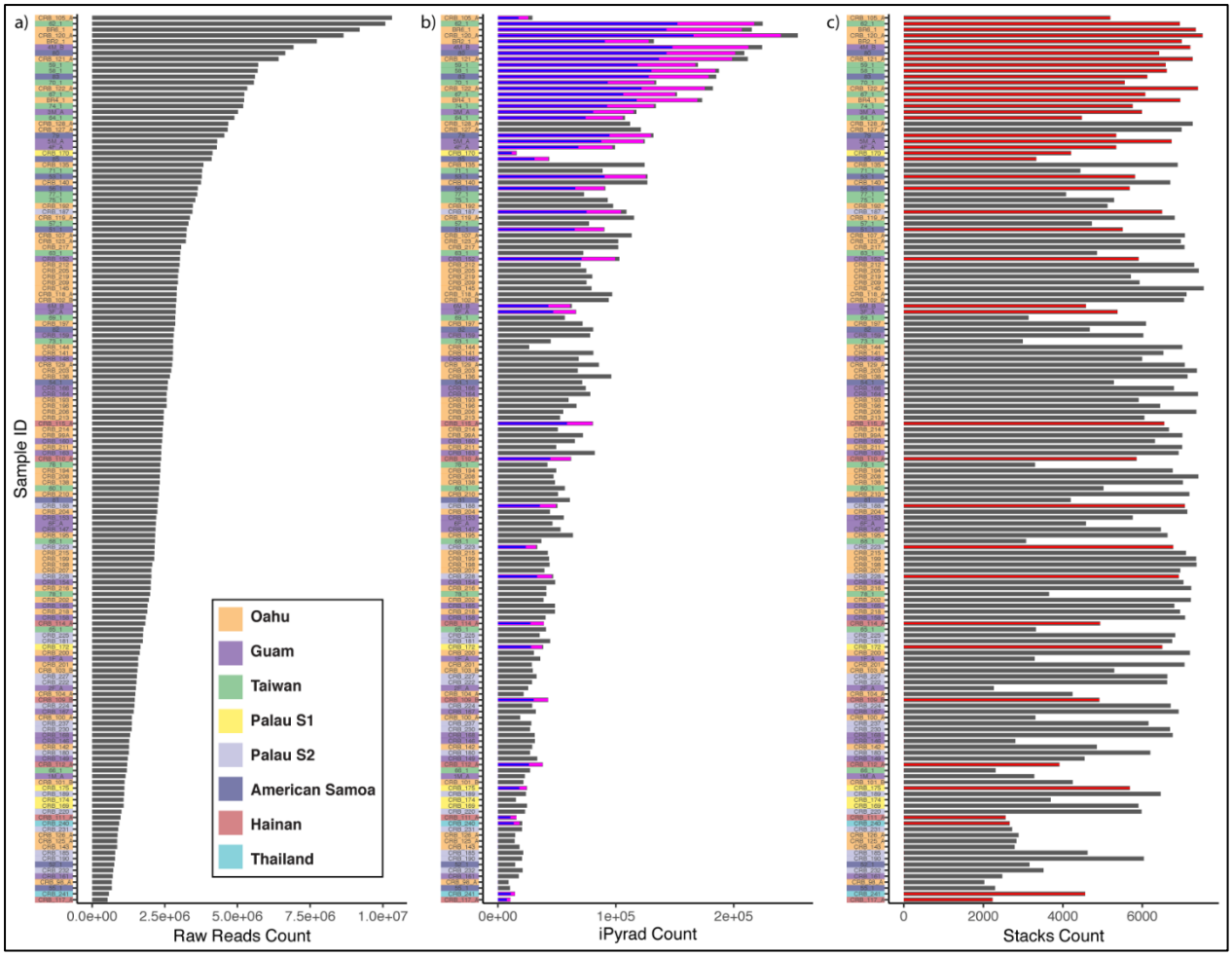

Figure 2. Graphical summary of data included in this study across different datasets. Colored boxes along the y-axis are used to help illustrate the different sampling regions. a) Raw read counts obtained from ddRAD sequencing for 151 retained samples. b) IPYRAD Loci or SNP counts for various datasets and analyses. (Dark Grey) Loci recovered from IPYRAD at minimum shared taxa value 4 when all 151 samples were included; (Pink) Loci recovered from IPYRAD at minimum shared taxa value 4 for the 44-sample subset; (Dark Blue) SNPs recovered from IPYRAD for variable loci at minimum shared taxa value 4 for the 44-sample subset. c) STACKS SNP counts for various datasets. (Dark Grey) SNPs recovered from STACKS for all 151 retained samples; (Red) SNPs recovered from STACKS for the 44-sample subset.

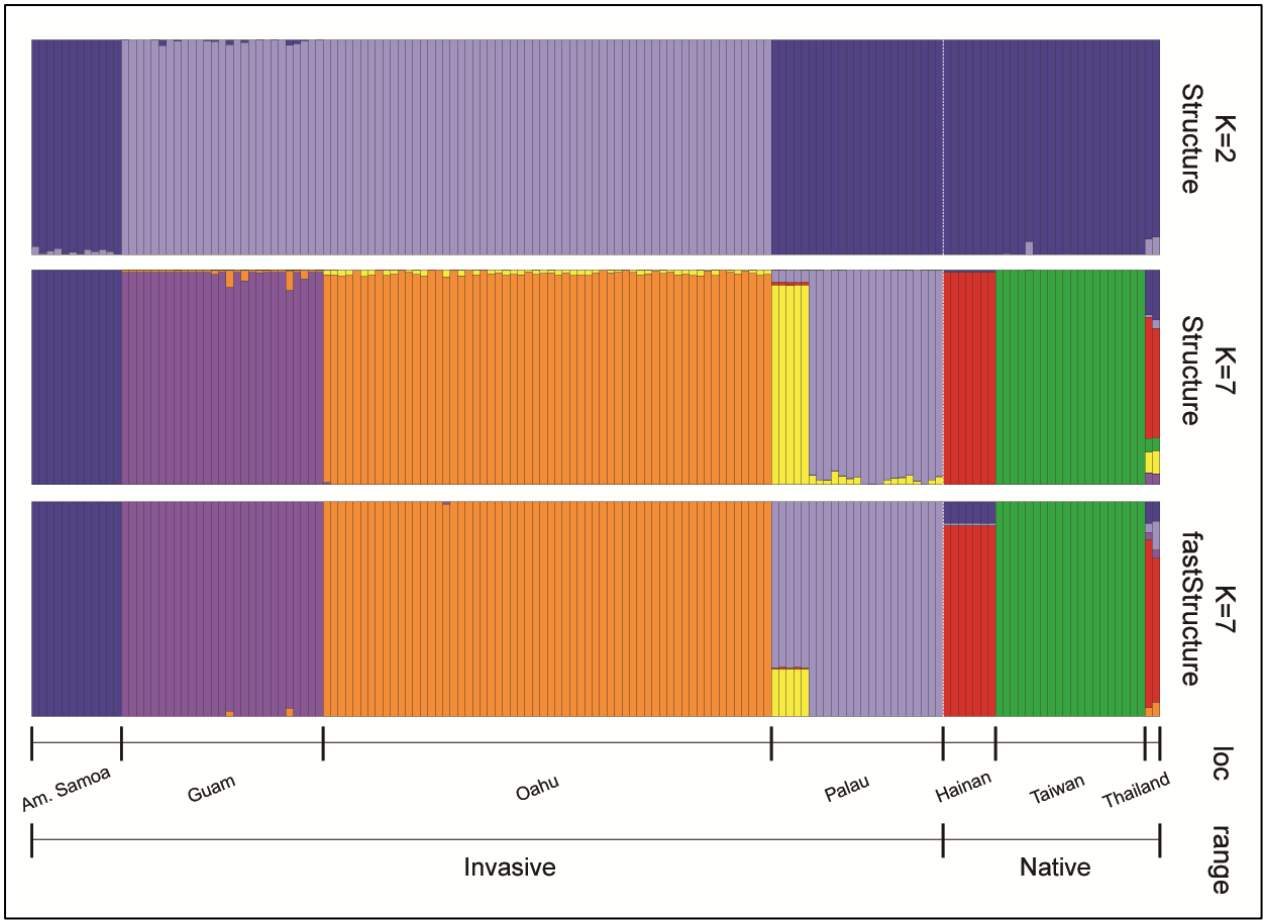

Figure 3. STRUCTURE Plots describing cluster assignment made by STRUCTURE and FASTSTRUCTURE for *O. rhinoceros* samples using the *STA.151* dataset. Each bar represents a single sample. Cluster colors are preserved and used to define similar regional clusters across the figures in this report. At K=2, STRUCTURE and FASTSTRUCTURE made identical assignments with extremely minor differences in admixture, therefore K=2 for FASTSTRUCTURE was omitted. Range designations were made based on historical records of where the beetle natively occurs and where it is known to be invasive.

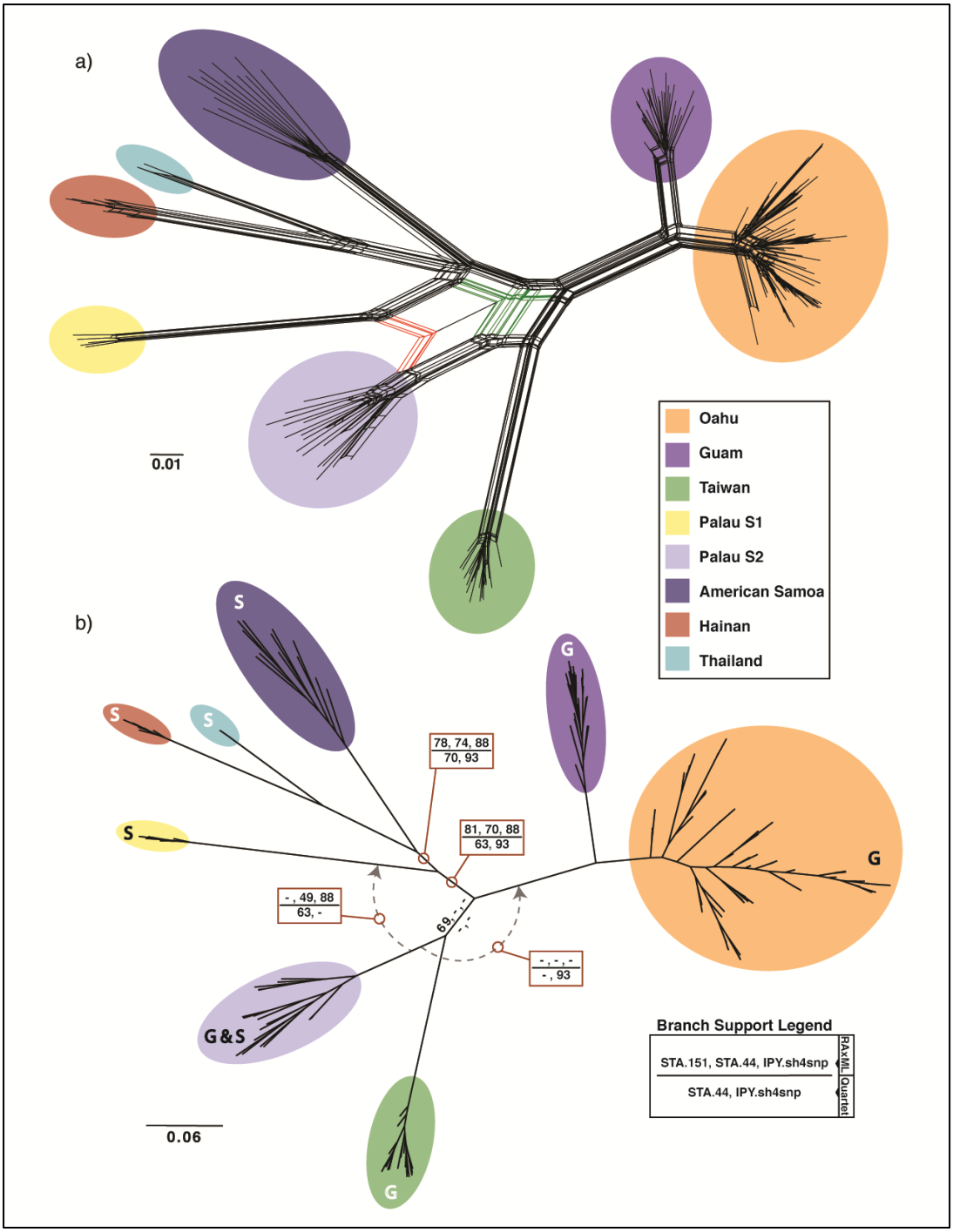

Figure 4. A composite figure describing the results of network and phylogenetic analyses on the different datasets compiled for this report. Scale bars indicate mean number of substitutions per site (in the case of 4b this value only applies to the representative tree). Branches without support values received full support across all analyses. a) A NeighborNet network illustrating multiple hypotheses simultaneously. Red branches highlight hypotheses depicting recent hybridization between beetle populations on Palau. Green branches highlight further areas of ambiguity involving beetle populations on Palau and may be indicative of previous hybridization with a third introduced population or simply hybridization of gene regions informative at deeper time scales. b) A composite tree displaying the RAXML results for *STA.151* (displayed tree) with the branch supports from the RAXML and SVDQUARTETS trees for *STA.44* and *IPY.sh4snp* superimposed onto the tree. Hyphens indicate missing branches in the topology of the respective tree. Beetle biotypes within each clade (biotype G, biotype S, or both) are indicated.

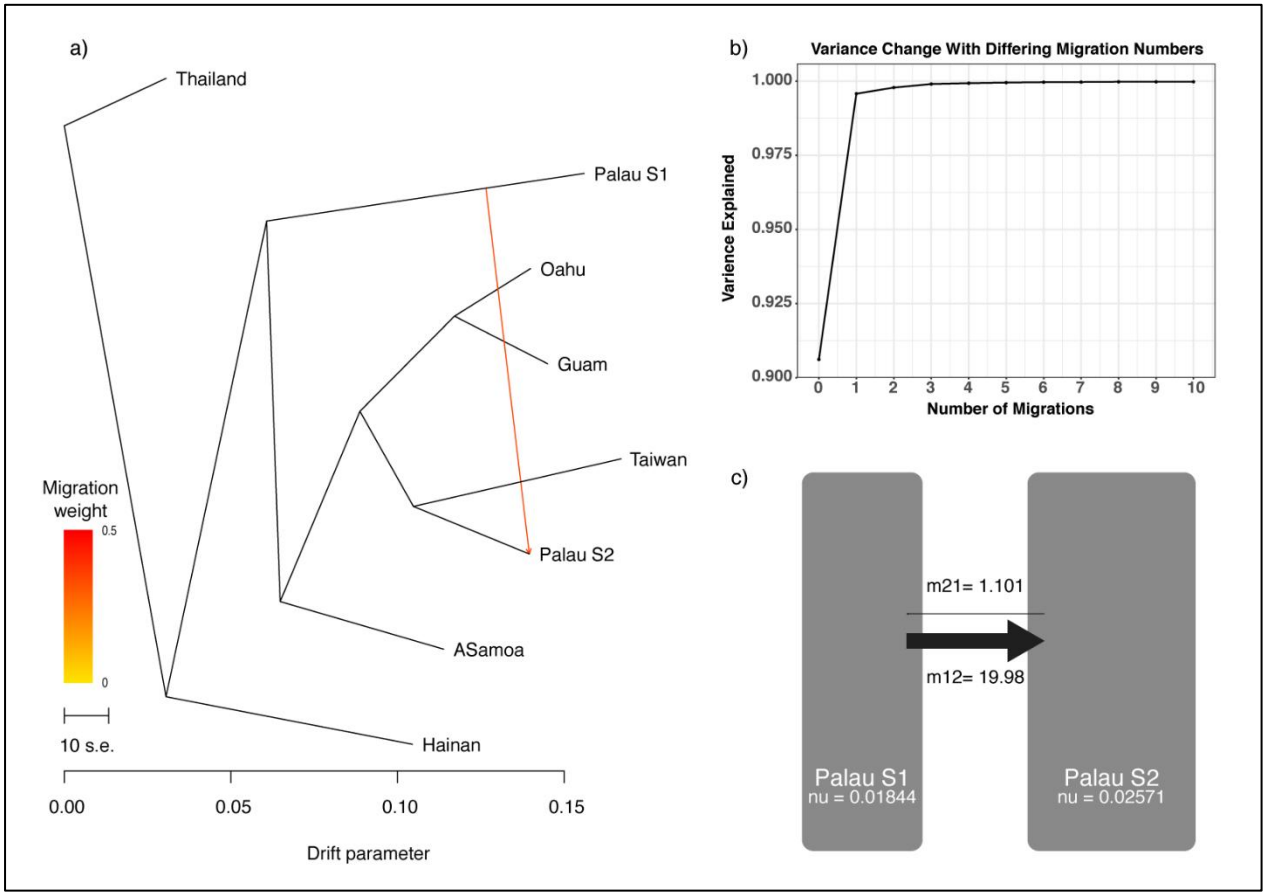

Figure 5. Demographic analyses from TREEMIX and MOMENTS for CRB populations. a) TREEMIX supports admixture between Palau S1 and Palau S2. The tree is rooted at Thailand which is in the geographic center of the beetle's native range. b) Variance explained by models with varying numbers of migrations. A single migration increases the proportion of variance explained from 0.9061 to 0.9958. Assuming additional migrations did not improve the model appreciably. c) MOMENTS software was used to obtain additional details on demographic history for populations involved in this event, illustrated here as grey bins representing Palau S1 and Palau S2. nu = effective population size; m12 = relative migration rate from Palau S1 to Palau S2; m21 = relative migration rate from Palau S2 to Palau S1.

Table 1. Per Loci Summary Statistics for all Sampling Locations

Location	N	A	Ae	Ho	He	uHe
Guam	19.052	1.299	1.167	0.092	0.098	0.101
Am. Samoa	7.316	1.434	1.233	0.144	0.160	0.177
Taiwan	11.663	1.133	1.061	0.055	0.060	0.065
Oahu	48.606	1.281	1.181	0.087	0.104	0.105
Hainan	3.917	1.015	0.969	0.050	0.045	0.054
Palau S1	3.288	1.065	1.008	0.073	0.065	0.078
Palau S2	13.810	1.388	1.222	0.116	0.136	0.142
Thailand	0.913	0.714	0.707	0.060	0.033	0.056

General diversity statistics for loci data calculated for the STA.151 data set evaluated on a per loci basis. N: Average Population Size; A: Number of Different Alleles; Ae: Number of Effective Alleles [1/(Sum pi^2)]; Ho: Observed Heterozygosity (No. of Hets/N); He: Expected Heterozygosity (1-Sum pi^2); uHe: Unbiased Expected Heterozygosity [(2N/(2N-1))*He].

Table 2. Population Pairwise Fst (AMOVA) Values

	Guam	Am. Samoa	Taiwan	Oahu	Hainan	Palau S1	Palau S2	Thailand
Guam		0.001	0.001	0.001	0.001	0.001	0.001	0.001
Am. Samoa	0.228		0.001	0.001	0.001	0.001	0.001	0.001
Taiwan	0.282	0.277		0.001	0.001	0.001	0.001	0.001
Oahu	0.124	0.336	0.395		0.001	0.001	0.001	0.001
Hainan	0.288	0.276	0.366	0.398		0.001	0.001	0.003
Palau S1	0.293	0.355	0.494	0.354	0.409		0.001	0.002
Palau S2	0.224	0.337	0.419	0.237	0.393	0.240		0.001
Thailand	0.270	0.241	0.431	0.414	0.289	0.365	0.369	

Pairwise Fst values evaluated across all population pairs. Fst values located below the diagonal; corresponding p-values located above the diagonal.