Received: 14 March 2022 | Revised: 4 June 2022 | Accepted: 9 June 2022

DOI: 10.1111/mec.16575

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Richness and resilience in the Pacific: DNA metabarcoding enables parallelized evaluation of biogeographic patterns

Susan Kennedy[1,2] | Jerilyn Calaor[3] | Yazmín Zurápiti[1] | Julian Hans[2] | Masashi Yoshimura[4] | Juanita Choo[5] | Jeremy C. Andersen[6] | Jackson Callaghan[7] | George K. Roderick[8] | Henrik Krehenwinkel[2,†] | Haldre Rogers[3,†] | Rosemary G. Gillespie[8,†] | Evan P. Economo[1,9,†]

[1]Biodiversity and Biocomplexity Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

[2]Department of Biogeography, Trier University, Trier, Germany

[3]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa, USA

[4]Environmental Research Support Section, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

[5]Science and Technology Group, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

[6]Department of Environmental Conservation, University of Massachusetts Amherst, Amherst, Massachusetts, USA

[7]Department of Integrative, Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA

[8]Department of Environmental Science, Policy and Management, University of California, Berkeley, California, USA

[9]Radcliffe Institute for Advanced Study, Harvard University, Cambridge, Massachusetts, USA

Correspondence
Susan Kennedy, Biodiversity and Biocomplexity Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan.
Email: susanrkennedy@gmail.com

Funding information
Okinawa Institute of Science and Technology Graduate University, Grant/Award Number: Subsidy funding

Handling Editor: Michael Hickerson

Abstract

Islands make up a large proportion of Earth's biodiversity, yet are also some of the most sensitive systems to environmental perturbation. Biogeographic theory predicts that geologic age, area, and isolation typically drive islands' diversity patterns, and thus potentially impact non-native spread and community homogenization across island systems. One limitation in testing such predictions has been the difficulty of performing comprehensive inventories of island biotas and distinguishing native from introduced taxa. Here, we use DNA metabarcoding and statistical modelling as a high throughput method to survey community-wide arthropod richness, the proportion of native and non-native species, and the incursion of non-natives into primary habitats on three archipelagos in the Pacific – the Ryukyus, the Marianas and Hawaii – which vary in age, isolation and area. Diversity patterns largely match expectations based on island biogeography theory, with the oldest and most geographically connected archipelago, the Ryukyus, showing the highest taxonomic richness and lowest proportion of introduced species. Moreover, we find evidence that forest habitats are more resilient to incursions of non-natives in the Ryukyus than in the less taxonomically rich archipelagos. Surprisingly, we do not find evidence for biotic homogenization across these three archipelagos: the assemblage of non-native species on each

[†]Joint senior authors.

island is highly distinct. Our study demonstrates the potential of DNA metabarcoding to facilitate rapid estimation of biogeographic patterns, the spread of non-native species, and the resilience of ecosystems.

KEYWORDS
conservation biology, insects, invasive species, island biogeography

1 | INTRODUCTION

Islands are indispensable to the study of life. They simultaneously serve as cradles of biodiversity, contributing approximately 20% of known species on Earth (Fernández-Palacios et al., 2021; Kier et al., 2009), and as ideal study systems for the characterization of universal biological processes at smaller and more manageable scales (Vitousek, 2002; Warren et al., 2015). Understanding the processes by which island communities assemble, interact, persist and change is a critical frontier in research, and is increasingly urgent given the many anthropogenic pressures that threaten island biodiversity (Delgado et al., 2017; Harter et al., 2015; Russell & Kueffer, 2019). A long tradition of both theory and empirical research has identified several factors determining the diversity of species and higher clades found on islands (Losos & Ricklefs, 2010; MacArthur & Wilson, 1967; Valente et al., 2020; Whittaker & Fernández-Palacios, 2007). All else equal, larger islands harbour more diversity through effects on extinction and/or speciation (Borregaard et al., 2017), while more isolated islands generally have lower diversity due to reduced colonization rates (Barreto et al., 2021; Kaloveloni et al., 2018). Moreover, the life cycle of islands probably has an effect on diversity: for example, continental islands may have higher diversity than oceanic islands (Castro-Urgal & Traveset, 2014; Kitayama, 1996), and the diversity of oceanic islands should generally increase with island age (Gillespie & Baldwin, 2010).

Although these principles should determine the patterns we see in natural systems, a major reorganization of communities is underway due to anthropogenic impacts. Humans have had a drastic effect on most island ecosystems (Wood et al., 2017), leading to changes in the composition of biological communities that play out against a background of "natural" biogeographic patterns. In some cases, the introduction of non-native species has resulted in extreme biotic turnover. For example, Cicconardi et al. (2017) found a striking degree of overlap in soil arthropod communities in three distinct, geographically distant archipelagos where a few haplotypes from non-native species made up nearly 30% of the Collembola in their samples. Such findings suggest that island biota may be in the process of becoming increasingly homogenized by the establishment of common invasive species. While such invasions may obscure pre-existing diversity patterns, biogeographic theory may nevertheless illuminate the extent to which different island communities are impacted by human activity. The biotic resistance hypothesis (Simberloff & Von Holle, 1999) proposes that communities with high native diversity and a corresponding scarcity of niche space are less invasible than those with lower diversity and more available niche space (Beaury et al., 2020; Mungi et al., 2021). Hence, oceanic islands with lower native diversity due to their geographic isolation, young geologic age or small size may be subjected to a greater degree of invasion and biotic homogenization than more connected, older or larger islands (Moser et al., 2018). Furthermore, continental islands, with their history of geographic connection to the mainland and consequently higher expected native diversity, should show a greater degree of biotic resistance to invasion than oceanic islands.

Pacific islands present a promising opportunity to test the susceptibility and resilience of island systems to biotic homogenization. Most Pacific archipelagos have high levels of endemism (Keppel et al., 2014) and were colonized by humans relatively recently, in the last 3500 years (Hunt et al., 2017). Many islands have been subjected to strong anthropogenic pressures, beginning with extensive habitat modification shortly after the arrival of humans (Boivin et al., 2016). Some archipelagos, including the Ryukyus, the Marianas and Hawaii, additionally experienced intense disturbance during World War II. These common features and histories make these islands a well-suited system for identifying how biological communities respond to similar anthropogenic impacts on archipelagos with different biogeographic characteristics. Importantly, the three aforementioned archipelagos represent a gradient of geologic ages and degrees of geographic isolation. The Ryukyus are the oldest and most geographically connected, having originated as part of the Asian continent during the Permian Period (299–251 million years ago [Ma]; Hanzawa, 1935) and later separated into islands by tectonic activity approximately 1.5 Ma (Osozawa et al., 2012). Unlike hotspot archipelagos, the Ryukyus are continental in origin (Hanzawa, 1935) and therefore should harbour diversity accumulated through vicariance as well as dispersal from the mainland. The Marianas are intermediate in age, with the southern arc of inhabited islands having originated 15–20 Ma (Pacific Islands Benthic Habitat Mapping Center, 2022) as a result of volcanic activity from the subduction of the Pacific plate under the Mariana plate, and are approximately 3000 km from continental Asia. Hawaii, created by volcanic activity over a hotspot, is the youngest archipelago, with its high islands ranging from ~5–<1 Myr old (Clague & Sherrod, 2014). Hawaii is also the most isolated at over 4000 km from the nearest continent, North America. Thus, these three archipelagos are expected to support biological communities with varying levels of diversity, which may in turn play a

role in determining their relative resilience to biological invasions. Thus far, it has been difficult to characterize patterns of diversity in whole communities due to the sheer magnitude of the task, especially in hyperdiverse groups such as arthropods. For these groups, taxonomic impediments are strong, limiting both characterization of native biotas and our ability to distinguish native from non-native taxa (Graham et al., 2017). However, the advent of DNA metabarcoding, alongside new computational approaches to analysing sequence data, has substantially enhanced our ability to address both of these questions (Kirse et al., 2021; Lopez-Vaamonde et al., 2019; Noguerales et al., 2021).

Here, we use DNA metabarcoding to test the extent to which these three archipelagos match our expectations regarding their native diversity and resilience to invasion based on their relative areas, ages and degrees of isolation. For each archipelago, we focus on one island and compare data from both intact and degraded forests. Okinawa-jima is situated in the centre of the Ryukyus Archipelago (Japan) and covers 1199 km^2 of area. Rota (Commonwealth of the Northern Mariana Islands), of the Marianas Archipelago, is situated immediately northeast of Guam, has an area of 85.5 km^2, and contains the highest proportion of intact limestone forest of all the Mariana islands (Falanruw et al., 1989). The Big Island of Hawaii (hereafter "Hawaii"; USA) is the youngest island in the Hawaiian Archipelago at <500,000 years (Clague & Sherrod, 2014) and the largest island included in this study, covering 10,438 km^2. We test our expectations using arthropods, which comprise one of the most taxonomically and functionally diverse groups on Earth (Giribet & Edgecombe, 2019) and perform critical roles in ecosystems as herbivores, predators, parasites, decomposers and pollinators (Wong et al., 2019). We hypothesize the following: (H1) Arthropod community diversity on a given island is a function of the age and size of that island, and inversely related to its degree of geographic isolation. If H1 is supported, then Hawaii (the youngest and most isolated) and Rota (Marianas; intermediate age and isolation but smallest size) should have lower native arthropod γ- and α-diversity than Okinawa-jima (Ryukyus; the oldest and least isolated). Although Hawaii is the largest of the three islands, we still expect relatively low native diversity there due to its very young age. (H2) Older, more connected and larger islands should have greater resistance to invasion and biotic homogenization than younger, more isolated and smaller ones. If H2 is supported, then Hawaii and the Marianas should have a significantly higher proportion of non-native taxa than the Ryukyus, and should also show evidence for biotic homogenization (decreased β-diversity between sites) while the Ryukyus should show comparably distinctive arthropod assemblages at different sites. (H3) Given the findings of Cicconardi et al. (2017), biological invasions may have led to an overall homogenization of arthropod assemblages such that degraded forests across archipelagos are more similar to each other than to intact forests on the same archipelago. If H3 is supported, then β-diversity between intact and degraded sites within an island should be higher than β-diversity between degraded sites across archipelagos. In addition, we would expect to find the same invasive or cosmopolitan arthropod species on all archipelagos.

2 | MATERIALS AND METHODS

2.1 | Field sites and sampling design

Sampling was performed in 2019 on the Ryukyus (April and October), Marianas (June), and Hawaii (August; Figure 1, Table 1). The April sampling round on the Ryukyus was used to establish and standardize the protocol; therefore, April samples were excluded from all analyses except for identifying species that were detected on multiple archipelagos (H3; see Table S1). Six forest sites were sampled on each island. Of these, three were classified as "intact" forest (dominated by native vegetation and usually under legal protection for conservation purposes) and three as "degraded" forest (secondary forest on previously cleared sites, dominated by non-native vegetation; see Table 1 for descriptions of the sites). In the case of the Ryukyus, we had just two intact sites and four degraded sites due to the difficulty of accessing intact forest on Okinawa-jima. The variable defined by intact or degraded forest is hereafter referred to as "forest type." Within each site, three sampling areas were designated approximately 100 m apart. Using a randomly selected compass bearing, a 20-m transect was set up at each of these areas. A GPS waypoint was recorded at the starting point using a GPSmap 62stc (Garmin). Two types of sampling were performed at points every 5 m along the transect: (1) litter sifting, wherein all leaf litter above the soil surface in a 1 m^2 quadrat was collected and sifted with a litter sifter (Marizete do Brasil, Ilhéus, Brazil) before undergoing 48 h of extraction in a mini-Winkler funnel (Marizete do Brasil) with a collecting cup at the bottom full of absolute ETOH; (2) vegetation beating (hereafter "beating"), in which canopy and understory vegetation surrounding the transect point was struck with a stick for 20 s while a fabric "beating sheet" (BioForm) was held underneath to catch falling arthropods, which were then collected by aspirator (Roppon-Ashi) into sterile 50-ml Falcon tubes (Thermo Fisher Scientific) and immediately preserved in absolute ETOH. Litter sifting samples were collected before beating samples in order to minimize chances of foliage arthropods falling into the leaf litter and being erroneously included in the litter sifting samples. Beating samples were collected separately for each point, whereas litter samples were pooled by transect, resulting in a total of five beating and one litter sample per transect. Table 2 summarizes sample sizes by collection method. ETOH-preserved samples were maintained at room temperature until DNA extraction.

A separate, brand new set of collecting equipment was used on each archipelago to prevent cross-contamination during sampling. After completing each transect, all equipment was thoroughly shaken out and visually inspected to ensure that no arthropods remained inside. On Hawaii, due to the threat of Rapid Ohia Death (Fortini et al., 2019; Roy et al., 2020), equipment was disinfected with 70% isopropanol between sites.

2.2 | DNA extraction

DNA was extracted nondestructively using the Qiagen PureGene tissue kit (Qiagen) following the manufacturer's protocol. When

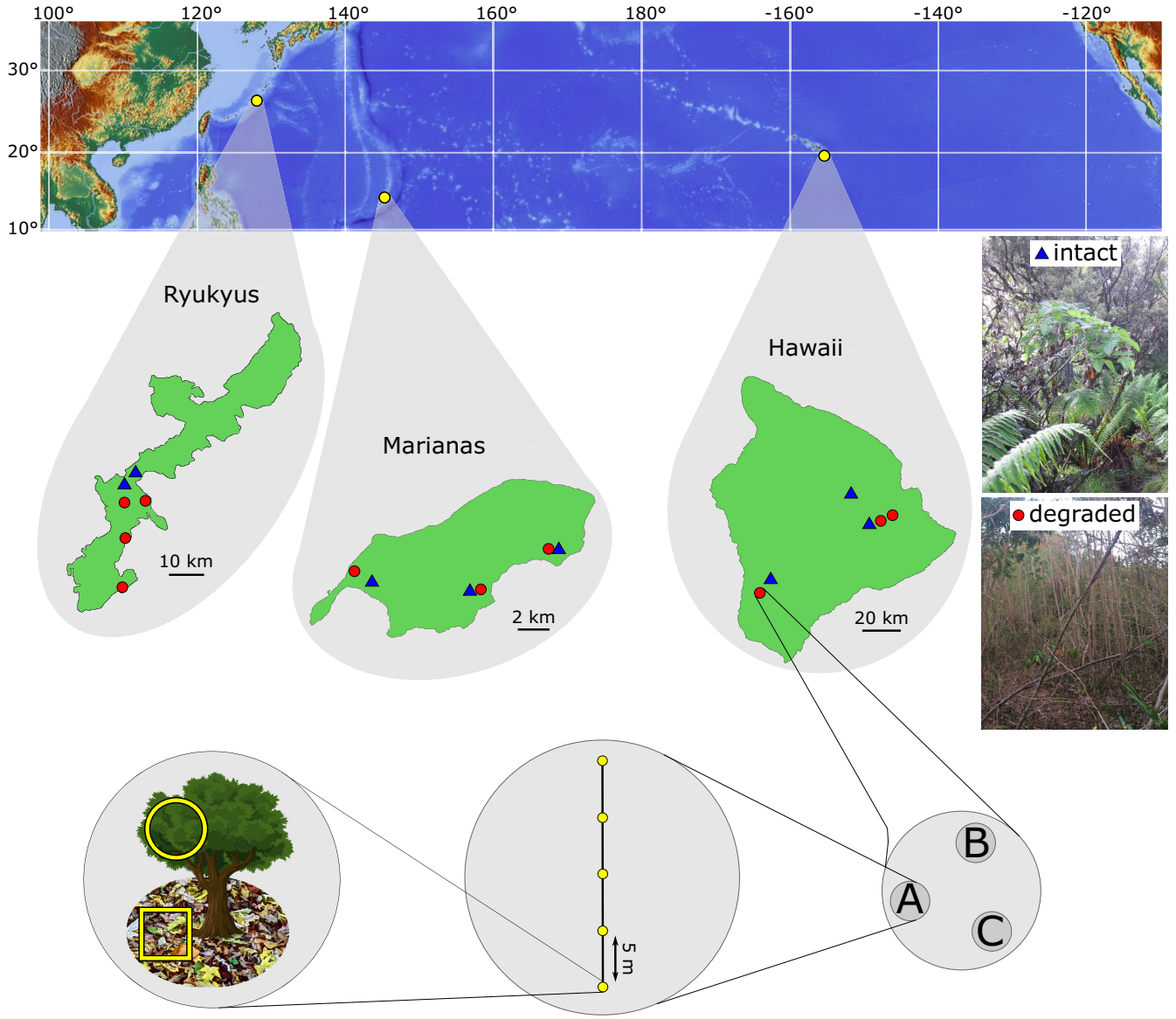

FIGURE 1 Study sites and sampling scheme. One island from each of three Pacific archipelagos – the Ryukyus (Okinawa-jima), the Marianas (Rota), and Hawaii (the Big Island) – was sampled in both intact (blue triangles) and degraded (red circles) forests. Three transects (A, B and C) were set up at each site. Leaf litter sifting and vegetation beating were performed every 5 m along the transects. Maps are adapted from maps-for-free.com

samples contained large specimens (>1 cm body length), tissue from those specimens was subsampled to prevent the DNA of those specimens from overwhelming the sample. For arachnids (e.g., opilionids) or adult insects, a leg was removed with sterile forceps and included in the sample. For large caterpillars and millipedes, a 1 cm piece of tissue was subsampled. Litter sifting samples were sorted by hand under a stereoscope, using sterile forceps, to remove soil and plant matter which might otherwise inhibit downstream PCR. However, for the majority of samples, no such preprocessing was necessary. ETOH was decanted off of the samples, which were then left to air-dry for 1–2 h with a Kimwipe on top to prevent introduction of contaminants. Then, 2 ml of cell lysis buffer (10 mM Tris pH 8, 100 mM NaCl, 10 mM EDTA pH 8 and 0.5% SDS dissolved in H_2O) and 10 µl of 20 mg/ml proteinase K (Thermo Fisher Scientific) were pipetted onto each sample. The samples were then incubated without shaking in a VWR Incu-Line (Avantor) at 55°C overnight (16–20 h). Subsequently, the lysate was pipetted off of the specimens and transferred to 2-ml tubes (Sarstedt), and the specimens were preserved in absolute ETOH. A 600-µl aliquot of lysate was used for subsequent DNA extraction. Coprecipitant GlycoBlue (15 mg/ml, Thermo Fisher Scientific) was included in the DNA precipitation step in a 1:600 ratio by volume with the isopropanol. DNA quality was checked on 1.5% agarose gels and found to be adequate for the purpose of this study (bands >10 kb). DNA extracts were aliquoted into 96-well PCR plates and stored at −20°C.

TABLE 1 Sampling site metadata. Latitude, longitude and elevation are the centroids of the starting points of the three transects at each site (datum = WGS 84). All dates sampled were in 2019

Site code	Archipelago	Site name	Latitude	Longitude	Elevation (m)	Forest type	Dates sampled (MM/DD)
R01, R09	Ryukyus	OIST Forest	26.4641	127.8403	100	Intact	04/15, 09/30, 10/01
R02, R11	Ryukyus	Takiyanbaru	26.4136	127.7891	137	Intact	04/16, 10/03
R04, R10	Ryukyus	Tamagusuku	26.1419	127.7839	109	Degraded	04/18, 10/02
R05, R12	Ryukyus	Nakagusuku	26.2851	127.7951	115	Degraded	04/19, 10/04
R07, R13	Ryukyus	Yacho No Mori	26.3763	127.8704	18	Degraded	04/23, 10/07
R08, R14	Ryukyus	SE Botanical Garden	26.3780	127.8063	51	Degraded	04/24, 10/08
M01	Marianas	Palii	14.1385	145.2234	216	Intact	06/08, 06/09
M02	Marianas	Bird Sanctuary Intact	14.1563	145.2673	116	Intact	06/09, 06/10
M03	Marianas	Old Japanese Road	14.1410	145.1569	143	Intact	06/11
M04	Marianas	Songsong Overlook	14.1458	145.1468	98	Degraded	06/12
M05	Marianas	Bird Sanctuary Degraded	14.1559	145.2641	105	Degraded	06/13
M06	Marianas	KC's	14.1390	145.2245	198	Degraded	06/14
H01	Hawaii	Army Road	19.5599	−155.2308	1121	Intact	08/03, 08/04
H02	Hawaii	Kona Hema Intact	19.2201	−155.8301	1132	Intact	08/07, 08/08
H03	Hawaii	Kona Hema Degraded	19.1958	−155.8560	656	Degraded	08/09
H04	Hawaii	Pu'u Maka'ala Low	19.5728	−155.1934	898	Degraded	08/16
H05	Hawaii	Upper Waiakea Low	19.5912	−155.1553	622	Degraded	08/26, 08/30
H06	Hawaii	Kipukas	19.6766	−155.2891	1285	Intact	08/31

TABLE 2 Sample sizes for the two collection methods (beating and litter sifting) at each scale of sampling. Three transects were sampled at each site. Six sites (three intact and three degraded) were sampled per archipelago, with the exception of the Ryukyus, where two intact and four degraded sites were sampled due to the difficulty of accessing intact sites

	Number of samples per…				
Collecting method	Transect	Site	Forest type, Ryukyus	Forest type, Marianas/Hawaii	Archipelago
Beating	5	15	Intact: 30; Degraded: 60	45	90
Litter sifting	1	3	Intact: 6; Degraded: 12	9	18

2.3 | PCR, library preparation and sequencing

Samples were amplified for two overlapping fragments of the mitochondrial cytochrome oxidase I (COI) barcoding region, using the following primer pairs: (1) mlCOIintF (5′-GGWACWGGWTGAACWGTWTAYCCYCC-3′; Leray et al., 2013)/Fol-degen-rev (5′-TANACYTCNGGRTGNCCRAARAAYCA-3′; Yu et al., 2012) for a 313-bp amplicon (hereafter "MCO"), and (2) ARF1 (5′-GCNCCWGAYATRGCNTTYCCNCG-3′; Gibson et al., 2014)/Fol-degen-rev for a 418-bp amplicon (hereafter "HCO"). Although these primers are very degenerate, we still expected a slight taxonomic bias for each one based on preliminary work (Table S2). Thus, by including both markers, we increased taxonomic recovery. Multiplex PCR was run in 10-μl volumes with an annealing temperature of 46°C, using the Qiagen Multiplex kit. Each reaction contained 5 μl of Qiagen multiplex master mix, 4 μM of ARF1, 2 μM of mlCOIintF, 6 μM of Fol-degen-rev, 2.8 μl of PCR water and 1 μl of template. Primers for HCO were used in double the volume of those for MCO because the shorter MCO fragment was expected to amplify preferentially. Template-free negative controls were run in every plate

to test for reagent contamination and sample cross-contamination. PCR success was verified using gel electrophoresis on 1.5% agarose gels. If a sample failed to amplify, the DNA extract was repurified following the Qiagen PureGene protocol, and PCR was repeated.

Indexing was accomplished in a second round of PCR, using a unique index combination for each sample following Lange et al. (2014) with an annealing temperature of 56°C. Indexed samples were visualized on a 1.5% agarose gel and pooled in approximately equimolar amounts based on gel band intensity; litter sifting samples were pooled in 2× volume relative to beating samples because litter sifting samples contained many more individual arthropods. The resulting pool was cleaned of residual primers with 1x AMPure beads XP (Beckman Coulter Life Sciences) and its concentration measured with a Qubit fluorometer (Thermo Fisher Scientific). It was then sequenced on an Illumina MiSeq (Illumina) at the Max Planck Institute for Evolutionary Biology (Plön, Germany) using a 600-cycle paired-end kit with V3 chemistry.

2.4 | Sequence processing

Sequence reads were demultiplexed by indexing barcode with MiSeq Control Software version 2.6.2.1 (Illumina) with one mismatch allowed. Forward and reverse reads were merged with PEAR version 0.9.11 (Zhang et al., 2014) using a 50-bp overlap and minimum quality of 20. Merged reads were quality filtered to keep only sequences in which at least 90% of bases had a quality of 30, and reads were converted to fasta format, with FastX-toolkit version 0.0.13 (Gordon & Hannon, 2010). Fasta files were demultiplexed by primer sequence, then combined into single files for each primer pair, using awk (Aho et al., 1979) and sed. For HCO, the first 105 bp were trimmed from the 5′ end, leaving only the 313 bp that completely overlapped with MCO. This allowed us to concatenate the MCO and HCO data sets into a single fasta file, containing a total of 5,646,420 sequence reads, for subsequent analysis. This fasta file was dereplicated and clustered into operational taxonomic units (OTUs) using USEARCH version 10.0.240 (Edgar, 2010). Clustering was performed at two levels: first, zero-radius OTUs (hereafter "zOTUs") were generated using the unoise3 command, then 3% radius (97% sequence similarity) OTUs (hereafter "OTUs") were clustered using the command cluster_otus. This was done because OTUs are more likely to correspond to species in the relatively fast-evolving COI, whereas zOTUs correspond roughly to haplotypes (Lim et al., 2022). A de novo chimera removal step is included in the clustering algorithms for both unoise3 and cluster_otus.

The OTU dataset was assigned to taxonomic identifications based on BLAST search (Altschul et al., 1990) of the NCBI database and the BOLD COI database (both downloaded 04/2022) with the top 10 hits retained. Only OTUs for which all 10 hits in both databases matched phylum Arthropoda were kept in the data set. These were then carefully investigated to identify the best match among the hits: we kept the hit with the highest (in decreasing order of importance) length of match, percent similarity, and taxonomic resolution. OTUs that matched equally to multiple different arthropod orders were discarded as likely contaminants (fungi or protists). Subsequently, a neighbour-joining tree was run in MEGA7 (Kumar et al., 2016), and the phylogenetic placement of each OTU was carefully examined to determine whether the taxonomic assignment was correct. If not, the order was changed according to tree topology; if a determination still could not be made, the order and lower taxonomic levels were assigned as "indeterminate." OTUs that fell outside of the main tree topology were discarded as probable nonarthropods. Family and genus were assigned as "indeterminate" for OTUs with <95% match to their corresponding BLAST hit, and species as "indeterminate" for those with <98% match. The zOTU data set was then BLAST searched against the OTU data set, and only zOTUs that matched an OTU with ≥97% similarity were retained. The unique identifiers (UIDs) of those OTUs were pasted onto the UIDs of the corresponding zOTUs using awk and sed. The zOTU dataset was then filtered for NUMTs/pseudogenes (Lopez et al., 1994) in MEGA7 as follows: The sequences were translated using the invertebrate mitochondrial genetic code, and any zOTUs containing a stop codon were deleted. Additionally, zOTUs with <312 or >313 bp, resulting in incomplete codons, were deleted. Finally, a zOTU table (Table S3) – a matrix showing the number of sequence reads for each zOTU recovered from each sample – was made using the otutab command in USEARCH. This zOTU table contained a total of 4,507,835 reads.

Prior to downstream analysis, the zOTU table was checked for signs of cross-contamination or index bleeding, both of which could lead to spurious zOTU occurrence in samples. To this end, PCR controls were checked for the occurrence of arthropod zOTUs, and the highest number of reads for a single zOTU in a negative control was used as the cleaning threshold for the corresponding data set. For example, in MCO, the maximum number of reads for a zOTU in a negative control was 37, so all occurrences of any zOTU with ≤37 reads were changed to 0 with Find/Replace in Microsoft Excel version 16051.14430.20306.0. After this cleaning step, a total of 4,321,589 reads remained in the data set. The zOTU table was then rarefied to equal coverage per sample per collection method, using thresholds observed from saturation curves (Figure S1) generated using vegan version 2.5-7 (Oksanen et al., 2007) in R version 4.1.2 (R Core Team, 2021): 2000 reads for beating and 4000 for litter sifting samples. Samples were rarefied using GUniFrac version 1.1 (Chen et al., 2012) in R. The two COI markers and collection methods were divided into four separate zOTU tables and rarefied separately. Subsequently, the aggregate function in R was used to calculate the sum of read counts for the two COI markers, resulting in a single zOTU table for each collection method. This was rarefied again to 2000 reads (for beating) or 4000 reads (for litter sifting) in order to retain all samples, including a few that had not passed the threshold for HCO, at equal coverage. Table S4 shows which OTUs were lost during cleaning and rarefaction. After this, the aggregate function was used to calculate the sum of read counts for zOTUs by their mapped OTU, thereby generating a data set for analysis at the OTU level (Table S5).

2.5 | Inferring geographic origins of OTUs

To test H1 and H2, we needed to know which OTUs were native to a given archipelago and which were introduced. The chief impediment was the fact that most OTUs did not match with high similarity to any sequence in GenBank or BOLD, and therefore could not be reliably assigned to taxonomy. Furthermore, even for those OTUs that matched a known taxon with high similarity (≥98%), it was not always possible to ascertain whether that taxon was native to one of the archipelagos. To overcome this obstacle, we used the recently designed software tool NIClassify (Callaghan & Andersen, 2021), which is based on the principles of Andersen et al. (2019). In brief, the software uses a training data set containing species with known native/introduced designations (hereafter "status") to train a classifier, based on sequence divergence within orders, using a random forest approach (Breiman, 2001). The software works optimally with zOTUs because it entails a species delimitation step; therefore, we used the zOTU data set for this analysis. Once trained, the classifier is then used to predict the status of a full data set of zOTUs based on their intraorder sequence divergence. The assumption is that groups of native species within orders will show shallower divergence from one another than from introduced species, which should be more distantly related to both the native taxa and each other (Andersen et al., 2019). We trained a separate classifier for each archipelago (Table S6), using zOTUs that BLAST-matched a sequence resolved to the species level with ≥98% sequence similarity and whose status we were able to assign based on web searches of GBIF (GBIF.org, accessed 2021–2022), ITIS (itis.gov, accessed 2021–2022) and primary literature. To increase the number of species with known status and thereby improve accuracy, we included the April samples in the training of the Ryukyus classifier. Because we were only able to identify 3 species in our dataset that were native to the Marianas, we added 29 sequences representing 7 arthropod species known to be native to the Marianas based on Bourquin (2002). COI sequences for these species were downloaded from the NCBI nucleotide database (accessed 02/2022) and aligned with our zOTU sequences before being added to the training data set. For each archipelago, 1000 classifiers were compared and the best-performing one was used for generating predictions.

2.6 | Statistical analysis

All statistical analyses were performed in R version 4.1.2 (R Core Team, 2021). Analyses were performed on the OTU data set, rather than zOTUs, in order to capture patterns at the (roughly) species level. April samples were excluded. All data sets were tested for violations of assumptions of normality using ggResidpanel version 0.3.0 (Goode & Rey, 2019) prior to subsequent analysis. Exploratory analysis revealed that the OTU assemblages obtained by beating were very distinctive from those obtained by litter sifting (Figure S2); we therefore performed analyses of α- and β-diversity (see below) on data sets separated by collection method.

For H1, we first calculated γ-diversity (OTU richness) of putatively native and putatively introduced species for data pooled by (1) island (all 6 sites), (2) intact forest only, and (3) degraded forest only. We then calculated α-diversity (OTU richness at the site level) using vegan and tested for interisland differences using a linear model and analysis of variance (ANOVA) followed by a Tukey's HSD test (Tukey, 1949). We then repeated this analysis for a dataset containing only putatively native species, and one containing only putatively introduced species, based on NIClassify designations.

For H2, we calculated β-diversity (pairwise Jaccard dissimilarity) between sites using ecodist version 2.0.7 (Goslee & Urban, 2007), then classified each pairwise comparison by the forest types being compared (intact:intact, intact:degraded, degraded:degraded). To evaluate whether some archipelagos had greater homogenization (lower β-diversity) than others, we tested for interisland differences in Jaccard dissimilarity of intact:degraded comparisons as described above. Because we sampled Hawaii on both the windward and leeward sides, which have considerable climatic differences (Frazier & Giambelluca, 2017), we also ran this analysis with all windward:leeward comparisons removed to control for potential effects of climate.

For H3, we analysed the beating and litter sifting data together because we were testing for homogenization on a broader scale. If H3 were supported, we would predict that degraded forests across archipelagos would be more similar to each other (have lower β-diversity) than to intact forests within the same archipelago. We therefore calculated pairwise β-diversity (Jaccard dissimilarity) between sites using ecodist, then tested for differences between intact:degraded comparisons within each archipelago and degraded:degraded comparisons across archipelagos, as described above.

3 | RESULTS

We recovered a total of 4454 zOTUs and 2075 OTUs belonging to the phylum Arthropoda. We were able to assign 1936 OTUs to order, 549 to family, 465 to genus and 282 to species. After cleaning and rarefaction, 2951 zOTUs and 1645 OTUs remained in the data set.

3.1 | Native vs. introduced status of OTUs

Using NIClassify, we were able to predict the status of 1586 OTUs in our data set. The remaining OTUs were excluded from this analysis for one of the following reasons: (1) they were not identifiable to the order level or (2) they represented singleton orders within the island where they were detected, precluding calculation of intraorder sequence divergence. To estimate classifier accuracy, we cross-checked the status predictions against known statuses of the OTUs we had looked up in web searches. We found conflicts between known status and NIClassify prediction in 18 of 95 OTUs for the Ryukyus, 1 of 41 for the Marianas, and 16 of 109 for Hawaii (35 conflicts out of 245 known statuses in total, i.e., 14%). These

conflicts were corrected to reflect the known statuses (Table S7). A breakdown of putatively native and introduced OTUs by order is available in Figure S3.

3.2 | H1: γ- and α-diversity patterns across archipelagos

We recovered the following total numbers of OTUs for each island (i.e., γ-diversity), including the OTUs for which we had not obtained NIClassify status: Ryukyus (excluding April sampling) – 650, Marianas – 332, Hawaii – 406. Based on NIClassify predictions, we recovered 548 (90%) native and 59 (10%) introduced OTUs from the Ryukyus (excluding April sampling), 24 (8%) native and 289 (92%) introduced from the Marianas, and 113 (30%) native and 270 (70%) introduced from Hawaii (Figure 2).

We found a significantly higher site-level OTU richness (α-diversity) for the Ryukyus than for either Hawaii or the Marianas for both beating (linear model, ANOVA $F = 15.38$, $p < .01$) and litter sifting data sets ($F = 10.24$, $p < .01$; Figure 3a and d). We recovered a similar pattern for data sets containing only putatively native species (beating: $F = 138.8$, $p < .01$; litter sifting: $F = 81.33$, $p < .01$; Figure 3b and e). The Marianas had the lowest native richness, Hawaii intermediate and the Ryukyus the highest (Figure 3b). For putatively introduced species, we found the inverse: namely, the Ryukyus had a significantly lower richness than Hawaii or the Marianas, with Hawaii falling intermediate and the Marianas having the highest (beating: $F = 55.63$, $p < .01$; litter sifting: $F = 20.40$, $p < .01$; Figure 3c and f).

3.3 | H2: β-diversity patterns within islands

We found significantly higher pairwise β-diversity between intact and degraded sites in the Ryukyus and Hawaii than in the Marianas for both beating (linear model, ANOVA $F = 21.32$, $p < .01$) and litter sifting data ($F = 14.92$, $p < .01$; Figure 4a and e). We recovered the

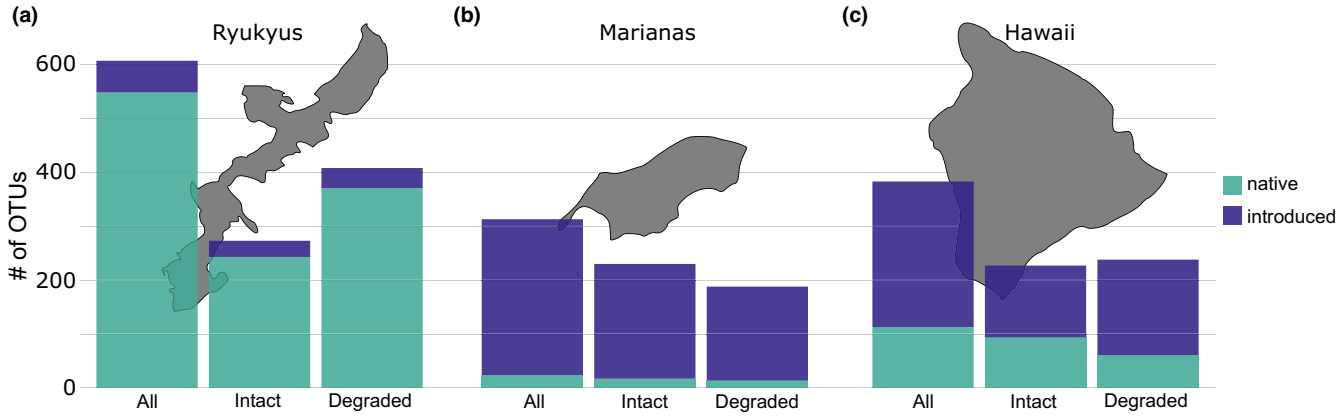

FIGURE 2 γ-Diversity comparison across islands showing total numbers of OTUs designated by NIClassify as native or introduced at all sites, intact sites only and degraded sites only, for (a) Ryukyus, (b) Marianas and (c) Hawaii

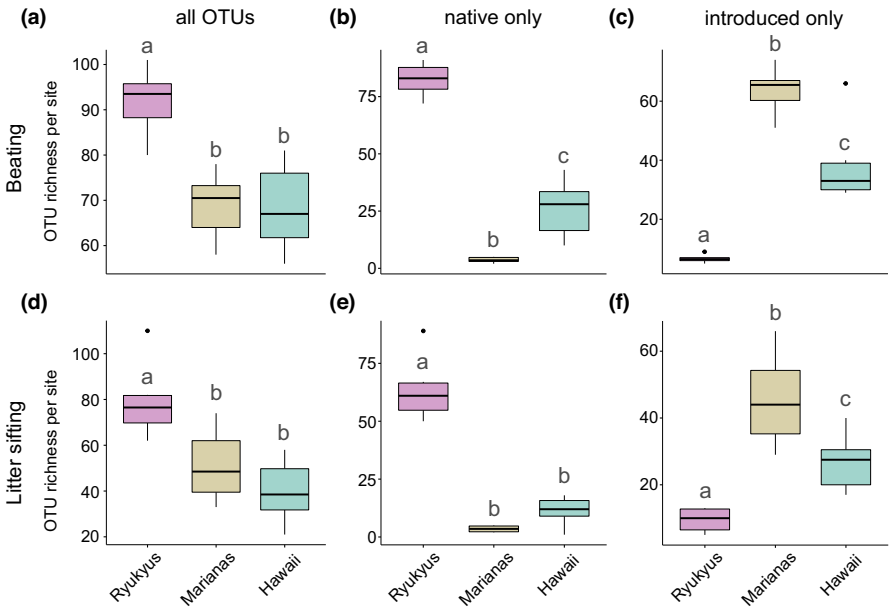

FIGURE 3 Cross-archipelago comparison of α-diversity (OTU richness) at the level of site for (a–c) beating and (d–f) litter sifting samples. (a) and (d) show overall OTU richness. Remaining panels show richness of OTUs designated by NIClassify as (b) and (e) native; (c) and (f) introduced to their respective archipelago. Significant differences in richness between islands (Tukey's HSD) are indicated by different letters (a, b, c) above the boxplots

same pattern after removing all comparisons between the windward and leeward sides of Hawaii (beating: $F = 19.33$, $p < .01$; litter sifting: $F = 13.60$, $p < .01$; Figure S4).

3.4 | H3: Testing for homogenization across archipelagos

We found a significantly higher β-diversity (Jaccard dissimilarity) for between-archipelago comparisons of degraded sites than for any within-island comparison of intact vs. degraded sites (linear model, ANOVA $F = 111.7$, $p < .01$; Figure 5a). Nonmetric multidimensional scaling (NMDS) shows clear clustering of samples by archipelago (Figure 5b). Only two OTUs were recovered from all three archipelagos (Figure 5d). We detected three OTUs on ≥2 archipelagos in intact (Figure 5c) and 35 in degraded forests (Figure 5d). Based on NIClassify predictions, we detected no putatively native (Figure 5e) and 15 putatively introduced OTUs on ≥2 archipelagos (Figure 5f).

4 | DISCUSSION

Our analysis demonstrates the utility of DNA metabarcoding for estimating biogeographic patterns and testing hypotheses across entire communities, circumventing traditional impediments to biodiversity research of hyperdiverse groups such as Arthropoda.

4.1 | H1: Older, more connected and larger islands have higher γ- and α-diversity

In support of H1, we recovered higher γ- and α-diversity in the Ryukyus, the most geographically connected archipelago, than in

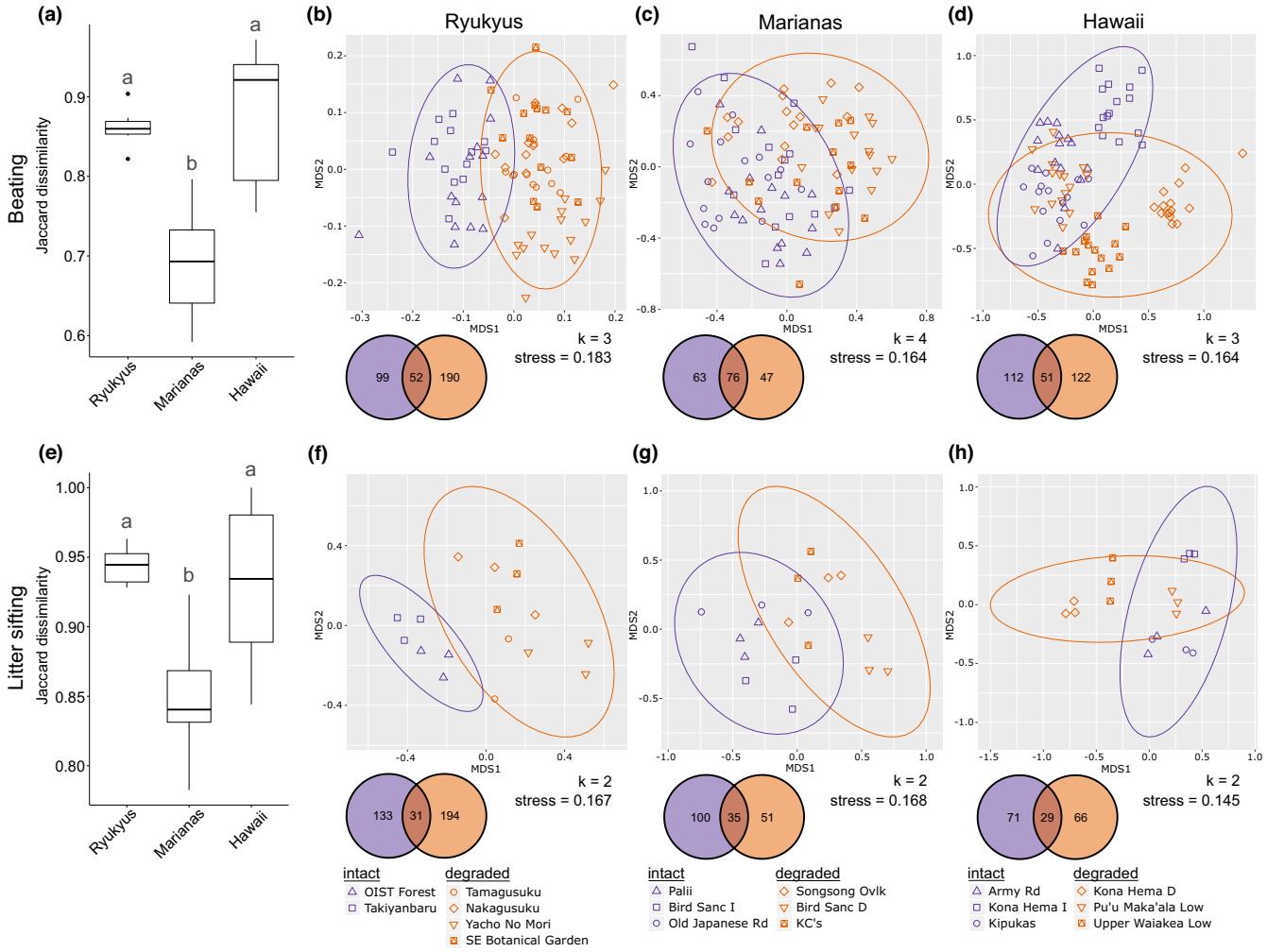

FIGURE 4 Cross-archipelago comparison of β-diversity and taxonomic overlap between intact and degraded forest sites. Boxplots show pairwise Jaccard dissimilarity between intact and degraded sites within islands for (a) beating and (e) litter sifting data. Significant differences (Tukey's HSD) are indicated by different letters (a, b) above the boxplots. NMDS (binary Jaccard) ordinations show samples from intact and degraded sites within each archipelago for (b), (c) and (d) beating and (f), (g) and (h) litter sifting data. Ellipses represent 95% confidence intervals for each forest type. Venn diagrams beneath NMDS plots show number of OTUs recovered from intact (purple) and degraded (orange) sites

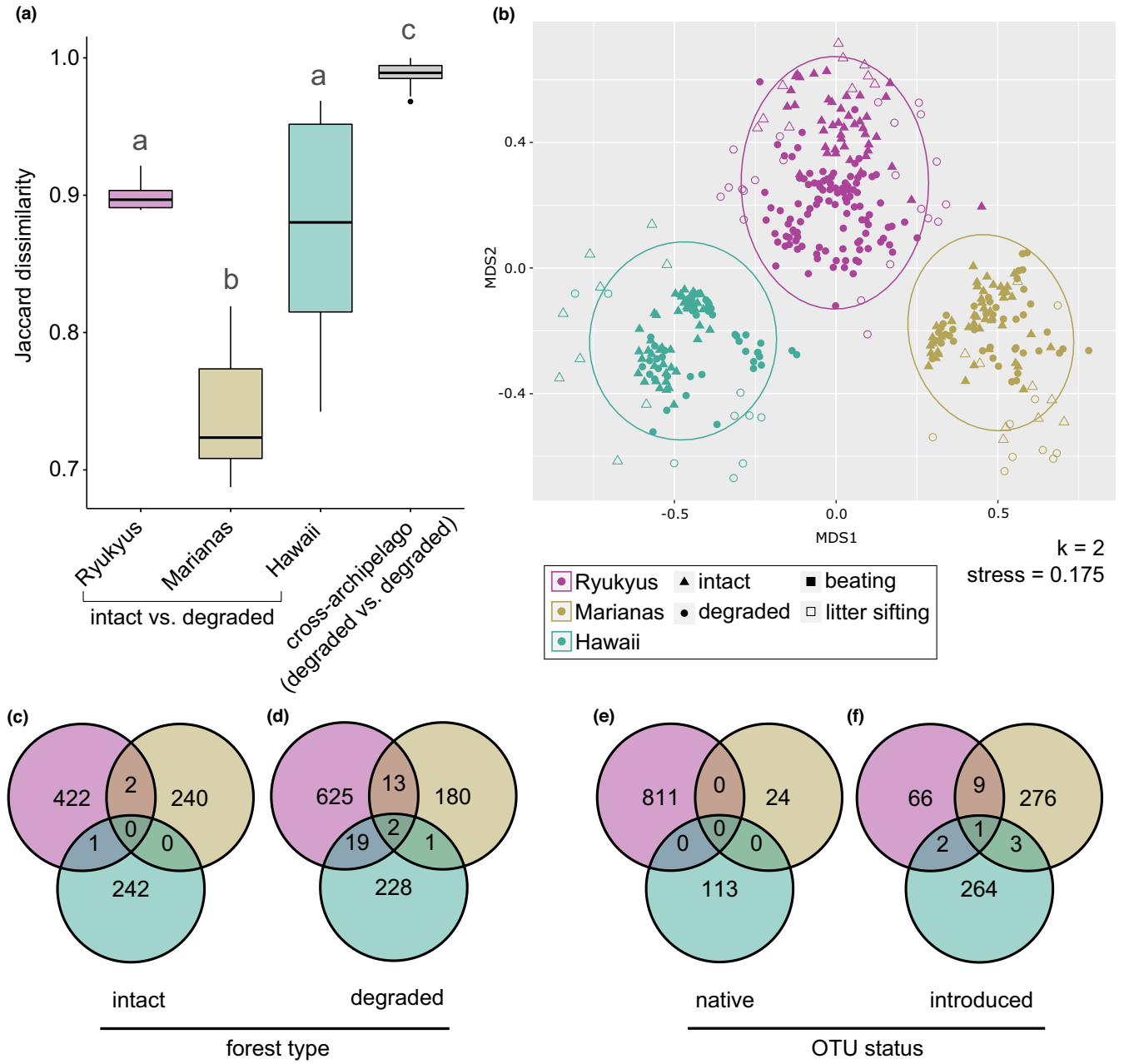

FIGURE 5 β-Diversity between arthropod assemblages on the three archipelagos. (a) Pairwise Jaccard dissimilarity between intact and degraded sites within, and degraded sites between, archipelagos. Significant differences (Tukey's HSD) are indicated by different letters (a, b, c) above the boxplots. April samples are excluded. (b) NMDS (binary Jaccard) of all samples, including those collected in April. Colour indicates archipelago, shape forest type and filled versus open collecting method. Ellipses represent 95% confidence intervals for each archipelago. (c–f) Venn diagrams of all OTUs recovered from the three archipelagos, including April samples, separated by (c, d) forest type: (c) intact, (d) degraded; (e, f) putative status of OTUs based on NIClassify: (e) native and (f) introduced. Archipelagos are shown clockwise from left: Ryukyus (purple), Marianas (yellow), Hawaii (green)

Hawaii or the Marianas (Figures 2 and 3). In addition to their oceanic origin and greater distance from other landmasses, our study islands for the latter two archipelagos were both expected to harbour relatively low native diversity because of their (1) very small size (85.5 km^2) in the case of Rota (Marianas) and (2) very young geological age (<500,000 y; Clague & Sherrod, 2014) in the case of Hawaii. When looking only at putatively native assemblages, we found a similar pattern: a significantly higher α-diversity on the Ryukyus than on either of the other two archipelagos (Figure 3b and e). We also found a significantly higher richness of putatively native species on Hawaii than on Rota (Marianas) for the beating samples (Figure 3b), perhaps due to Hawaii's much larger size. Overall, our results suggest that the "naturally occurring" diversity on these islands indeed matched our expectations. Conversely, we found a significantly higher α-diversity of putatively introduced taxa on both the Marianas and Hawaii than on the Ryukyus (Figure 3c

and f), indicating that the arthropod communities of the Ryukyus may be more resistant to invasion. The statuses we predicted with NIClassify should be interpreted with some caution: very few OTUs in our data set could be assigned to known status based on web searches, which limited the size and potentially the accuracy of our training data sets (Table S6). Our training data were also unbalanced in the data sets for the Marianas, where we had many more known introduced than known native species, and for the Ryukyus, where we had the opposite. Additionally, the sequence divergence patterns on which NIClassify relies might not always accurately reflect status, especially when based on short amplicons (313 bp in our case). Even so, our NIClassify predictions matched our known statuses in the majority of cases (approximately 86%) and recovered such pronounced differences between archipelagos that they probably reflect a true difference, even if the magnitude of that difference is imperfectly characterized. Given the dearth of sequences available in NCBI and BOLD especially for these relatively understudied island systems (Dopheide et al., 2019), the ability to predict the status of an OTU without knowing its exact species is exceptionally useful. NIClassify thus offers a highly valuable tool which we expect will continue to improve in power and accuracy as it is implemented in future metabarcoding studies.

4.2 | H2: Older, more connected and larger islands have higher β-diversity and are more resistant to invasion

By and large, our findings for β-diversity within archipelagos also match expectations. In support of H2, we found a lower β-diversity on Rota, of the relatively isolated Marianas, than on Okinawa-jima, of the relatively connected Ryukyus (Figure 4). However, we also found that the degree of differentiation (Jaccard dissimilarity) between intact and degraded forests was not significantly less in Hawaii than in the Ryukyus (Figure 4a and e). This may be explained by Hawaii's much larger size and topographical complexity, particularly its large elevational range (Table 1), compared with the other two islands. Together with our findings on γ- and α-diversity above, this indicates that the Ryukyus may have a higher resistance to invasion than the other two archipelagos. We found not only a larger number of native species in the Ryukyus, but also a much higher proportion of total diversity: 90% of all species with a predicted status. Furthermore, the Ryukyus have a lower absolute number of introduced species than either the Marianas or Hawaii (Figure 2). Having originated as part of the Asian continent, the Ryukyus may have obtained many of their native taxa by vicariance; in contrast, oceanic islands can only accumulate species via dispersal or in situ speciation, which may make them lag behind continental islands in terms of naturally occurring species richness (Castro-Urgal & Traveset, 2014). This in turn may lead to a higher degree of biotic resistance in continental than in oceanic islands (Longman et al., 2018).

4.3 | H3: Biotic homogenization across archipelagos?

Contrary to our expectations under H3, we did not find strong evidence for biotic homogenization across archipelagos. Although we detected known introduced species on all archipelagos (Table S6), we found little taxonomic overlap, even in degraded forests (Figure 5d). We found only two species on all three archipelagos, *Cardiocondyla obscurior* (Hymenoptera: Formicidae) and *Kallitaxila sinica* (Hemiptera: Tropiduchidae; Table S1). *C. obscurior* is a tramp ant (Heinze et al., 2006) and *K. sinica* an agricultural pest (Englund, 2004). It is important to note, however, that absence of evidence is not evidence of absence. The current study was limited in scale, employing only a small subset of potential collecting methods and forest sites, and did not fully saturate OTU richness at those sites (Figure S5). Thus, these islands may in fact harbour many shared introduced taxa that simply went undetected by our sampling protocol. Nevertheless, the apparently strong differentiation we recovered between island assemblages is striking (Figure 5b). This study served primarily as a pilot, and as such we feel that it succeeded in demonstrating the efficacy and feasibility of using metabarcoding to characterize biogeographic patterns and community resilience in a comparative framework across the Pacific. However, larger-scale studies are still needed to determine the extent to which widespread invasive species may be impacting Pacific islands.

AUTHOR CONTRIBUTIONS

Evan P. Economo, Henrik Krehenwinkel, Haldre Rogers and Rosemary G. Gillespie obtained the funding that supported this study. Susan Kennedy, Jerilyn Calaor, Henrik Krehenwinkel, Haldre Rogers, Rosemary G. Gillespie, George K. Roderick, Yazmín Zurápiti and Evan P. Economo conceived and designed the study. Susan Kennedy, Jerilyn Calaor, Henrik Krehenwinkel, Haldre Rogers, Yazmín Zurápiti and Masashi Yoshimura organized and carried out the fieldwork. Susan Kennedy performed the laboratory work. Juanita Choo, Jackson Callaghan and Jeremy C. Andersen contributed scientific and practical expertise. Susan Kennedy, Jerilyn Calaor and Julian Hans performed statistical analyses and produced the figures. All authors wrote, edited and approved the manuscript.

ACKNOWLEDGEMENTS

We thank the Environmental Science Section team at Okinawa Institute of Science & Technology (OIST) as well as Masako Ogasawara, Takuma Yoshida, Dana Drück, Jana Fritsch, Marie Klaka, Judith Paetsch, Jillian Suh-Kurovski, Juan Mungaray, Gaurav Agavekar, Lazzat Aibekova, Julian Katzke and Kelvin Chen for help with fieldwork; Chisa Oshiro, Fumika Azuma, Matthew Knope and Mel Johansen for logistical assistance; Sven Künzel for sequencing our libraries; Natalie Graham for advice on analysis; and three anonymous reviewers for helpful feedback on a previous draft. This study was supported by subsidy funding to OIST. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Raw sequence reads have been made available in the Dryad Digital Repository (https://doi.org/10.5061/dryad.tb2rbp036; Kennedy et al., 2022).

BENEFIT-SHARING STATEMENT

This project is an international collaboration with scientists from the countries providing genetic samples. All collaborators are included as coauthors or listed in the Acknowledgements. The research addresses a priority concern, namely the conservation of island biota and characterization of threats from invasive species.

ORCID

Susan Kennedy https://orcid.org/0000-0002-1616-3985
Henrik Krehenwinkel https://orcid.org/0000-0001-5069-8601
Rosemary G. Gillespie https://orcid.org/0000-0003-0086-7424

REFERENCES

Aho, A. V., Kernighan, B. W., & Weinberger, P. J. (1979). Awk—A pattern scanning and processing language. *Software: Practice and Experience*, *9*(4), 267–279. https://doi.org/10.1002/spe.4380090403

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andersen, J. C., Oboyski, P., Davies, N., Charlat, S., Ewing, C., Meyer, C., Krehenwinkel, H., Lim, J. Y., Noriyuki, S., & Ramage, T. (2019). Categorization of species as native or nonnative using DNA sequence signatures without a complete reference library. *Ecological Applications*, *29*(5), e01914. https://doi.org/10.1002/eap.1914

Barreto, E., Rangel, T. F., Pellissier, L., & Graham, C. H. (2021). Area, isolation and climate explain the diversity of mammals on islands worldwide. *Proceedings of the Royal Society B*, *288*(1965), 20211879. https://doi.org/10.1098/rspb.2021.1879

Beaury, E. M., Finn, J. T., Corbin, J. D., Barr, V., & Bradley, B. A. (2020). Biotic resistance to invasion is ubiquitous across ecosystems of the United States. *Ecology Letters*, *23*(3), 476–482. https://doi.org/10.1111/ele.13446

Boivin, N. L., Zeder, M. A., Fuller, D. Q., Crowther, A., Larson, G., Erlandson, J. M., Denham, T., & Petraglia, M. D. (2016). Ecological consequences of human niche construction: Examining long-term anthropogenic shaping of global species distributions. *Proceedings of the National Academy of Sciences*, *113*(23), 6388–6396. https://doi.org/10.1073/pnas.1525200113

Borregaard, M. K., Amorim, I. R., Borges, P. A., Cabral, J. S., Fernández-Palacios, J. M., Field, R., Heaney, L. R., Kreft, H., Matthews, T. J., Olesen, J. M., & Price, J. (2017). Oceanic Island biogeography through the lens of the general dynamic model: Assessment and prospect. *Biological Reviews*, *92*(2), 830–853. https://doi.org/10.1111/brv.12256

Bourquin, O. (2002). Invertebrates recorded from the Northern Marianas Islands – Status 2002. *Unpublished report. CNMI Invertebrate Collection, CREES-Northern Marianas College, Saipan*, *468*.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Callaghan, J., & Andersen, J. (2021). *NIClassify: A combined GUI toolkit for predictively classifying arthropod sequences as Native/Introduced*. https://github.com/tokebe/niclassify

Castro-Urgal, R., & Traveset, A. (2014). Differences in flower visitation networks between an oceanic and a continental Island. *Botanical Journal of the Linnean Society*, *174*(3), 478–488. https://doi.org/10.1111/boj.12134

Chen, J., Bittinger, K., Charlson, E. S., Hoffmann, C., Lewis, J., Wu, G. D., Collman, R. G., Bushman, F. D., & Li, H. (2012). Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics*, *28*(16), 2106–2113. https://doi.org/10.1093/bioinformatics/bts342

Cicconardi, F., Borges, P. A., Strasberg, D., Oromí, P., López, H., Pérez-Delgado, A. J., Casquet, J., Caujapé-Castells, J., Fernández-Palacios, J. M., & Thébaud, C. (2017). MtDNA metagenomics reveals large-scale invasion of belowground arthropod communities by introduced species. *Molecular Ecology*, *26*(12), 3104–3115. https://doi.org/10.1111/mec.14037

Clague, D. A., & Sherrod, D. R. (2014). Growth and degradation of Hawaiian volcanoes. *Characteristics of Hawaiian Volcanoes*, *1801*, 97–146.

Delgado, J. D., Riera, R., Rodriguez, R. A., Gonzalez-Moreno, P., & Fernández-Palacios, J. M. (2017). A reappraisal of the role of humans in the biotic disturbance of islands. *Environmental Conservation*, *44*(4), 371–380. https://doi.org/10.1017/S0376892917000236

Dopheide, A., Tooman, L. K., Grosser, S., Agabiti, B., Rhode, B., Xie, D., Stevens, M. I., Nelson, N., Buckley, T. R., Drummond, A. J., & Newcomb, R. D. (2019). Estimating the biodiversity of terrestrial invertebrates on a forested Island using DNA barcodes and metabarcoding data. *Ecological Applications*, *29*(4), e01877. https://doi.org/10.1002/eap.1877

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Englund, R. A. (2004). *Report for the 2003 Pacific biological survey, bishop museum Austral Islands, French Polynesia expedition to Tubuai and Rurutu*. Bishop Museum Honolulu.

Falanruw, M. C., Cole, T. G., & Ambacher, A. H. (1989). *Vegetation survey of Rota, Tinian, and Saipan, Commonwealth of the Northern Mariana Islands. Resource Bulletin. PSW-27* (p. 11). US Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station.

Fernández-Palacios, J. M., Kreft, H., Irl, S. D., Norder, S., Ah-Peng, C., Borges, P. A., Burns, K. C., de Nascimento, L., Meyer, J.-Y., & Montes, E. (2021). Scientists' warning – The outstanding biodiversity of islands is in peril. *Global Ecology and Conservation*, *31*, e01847. https://doi.org/10.1016/j.gecco.2021.e01847

Fortini, L. B., Kaiser, L. R., Keith, L. M., Price, J., Hughes, R. F., Jacobi, J. D., & Friday, J. (2019). The evolving threat of rapid 'Ōhi 'a death (ROD) to Hawai'i's native ecosystems and rare plant species. *Forest Ecology and Management*, *448*, 376–385. https://doi.org/10.1016/j.foreco.2019.06.025

Frazier, A. G., & Giambelluca, T. W. (2017). Spatial trend analysis of Hawaiian rainfall from 1920 to 2012. *International Journal of Climatology*, *37*(5), 2522–2531. https://doi.org/10.1002/joc.4862

Gibson, J., Shokralla, S., Porter, T. M., King, I., van Konynenburg, S., Janzen, D. H., Hallwachs, W., & Hajibabaei, M. (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. *Proceedings of the National Academy of Sciences*, *111*(22), 8007–8012. https://doi.org/10.1073/pnas.1406468111

Gillespie, R. G., & Baldwin, B. G. (2010). Island biogeography of remote archipelagoes. In J. B. Losos & R. E. Ricklefs (Eds.), *The theory of Island biogeography revisited* (pp. 358–387). Princeton University Press.

Giribet, G., & Edgecombe, G. D. (2019). The phylogeny and evolutionary history of arthropods. *Current Biology*, *29*(12), R592–R602. https://doi.org/10.1016/j.cub.2019.04.057

Goode, K., & Rey, K. (2019). *ggResidpanel: Panels and interactive versions of diagnostic plots using 'ggplot2'. R package version 0.3.0*.

Gordon, A., & Hannon, G. (2010). Fastx-toolkit. *FASTQ/A short-reads pre-processing tools (unpublished)* http://hannonlab.cshl.edu/fastx_toolkit, 5.

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*(1), 1–19. https://doi.org/10.18637/jss.v022.i07

Graham, N. R., Gruner, D. S., Lim, J. Y., & Gillespie, R. G. (2017). Island ecology and evolution: Challenges in the Anthropocene. *Environmental Conservation*, *44*(4), 323–335. https://doi.org/10.1017/S0376892917000315

Hanzawa, S. (1935). Geological history of the Ryukyu Islands. *Proceedings of the Imperial Academy*, *11*(2), 58–61.

Harter, D. E., Irl, S. D., Seo, B., Steinbauer, M. J., Gillespie, R., Triantis, K. A., Fernández-Palacios, J.-M., & Beierkuhnlein, C. (2015). Impacts of global climate change on the floras of oceanic islands – Projections, implications and current knowledge. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*(2), 160–183. https://doi.org/10.1016/j.ppees.2015.01.003

Heinze, J., Cremer, S., Eckl, N., & Schrempf, A. (2006). Stealthy invaders: The biology of Cardiocondyla tramp ants. *Insectes Sociaux*, *53*(1), 1–7. https://doi.org/10.1007/s00040-005-0847-4

Hunt, T. L., Lipo, C. P., Boivin, N., Crassard, R., & Petraglia, M. (2017). The last great migration: Human colonization of the remote Pacific islands. In *Human Dispersal and Species Movement: From Prehistory to the Present* (pp. 194–216). Cambridge University Press.

Kaloveloni, A., Tscheulin, T., & Petanidou, T. (2018). Geography, climate, ecology: What is more important in determining bee diversity in the Aegean archipelago? *Journal of Biogeography*, *45*(12), 2690–2700. https://doi.org/10.1111/jbi.13436

Kennedy, S., Calaor, J., Zurápiti, Y., Hans, J., Yoshimura, M., Choo, J., Andersen, J. C., Callaghan, J., Roderick, G. K., Krehenwinkel, H., Rogers, H., Gillespie, R. G., & Economo, E. P. (2022). Richness and resilience in the Pacific: DNA metabarcoding enables parallelized evaluation of biogeographic patterns. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.tb2rbp036

Keppel, G., Morrison, C., Meyer, J.-Y., & Boehmer, H. J. (2014). Isolated and vulnerable: The history and future of Pacific Island terrestrial biodiversity. *Pacific Conservation Biology*, *20*(2), 136–145. https://doi.org/10.1071/PC140136

Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., Mutke, J., & Barthlott, W. (2009). A global assessment of endemism and species richness across Island and mainland regions. *Proceedings of the National Academy of Sciences*, *106*(23), 9322–9327. https://doi.org/10.1073/pnas.0810306106

Kirse, A., Bourlat, S. J., Langen, K., & Fonseca, V. G. (2021). Metabarcoding malaise traps and soil eDNA reveals seasonal and local arthropod diversity shifts. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-89950-6

Kitayama, K. (1996). Patterns of species diversity on an oceanic versus a continental Island mountain: A hypothesis on species diversification. *Journal of Vegetation Science*, *7*(6), 879–888. https://doi.org/10.2307/3236463

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Lange, V., Böhme, I., Hofmann, J., Lang, K., Sauter, J., Schöne, B., Paul, P., Albrecht, V., Andreas, J. M., & Baier, D. M. (2014). Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics*, *15*(1), 1–11. https://doi.org/10.1186/1471-2164-15-63

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 1–14. https://doi.org/10.1186/1742-9994-10-34

Lim, J. Y., Patiño, J., Noriyuki, S., Cayetano, L., Gillespie, R. G., & Krehenwinkel, H. (2022). Semi-quantitative metabarcoding reveals how climate shapes arthropod community assembly along elevation gradients on Hawaii Island. *Molecular Ecology*, *2021*(31), 1416–1429. https://doi.org/10.1111/mec.16323

Longman, E. K., Rosenblad, K., & Sax, D. F. (2018). Extreme homogenization: The past, present and future of mammal assemblages on islands. *Global Ecology and Biogeography*, *27*(1), 77–95. https://doi.org/10.1111/geb.12677

Lopez, J. V., Yuhki, N., Masuda, R., Modi, W., & O'Brien, S. J. (1994). Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *Journal of Molecular Evolution*, *39*(2), 174–190. https://doi.org/10.1007/BF00163806

Lopez-Vaamonde, C., Sire, L., Rasmussen, B., Rougerie, R., Wieser, C., Allaoui, A. A., Minet, J., DeWaard, J. R., Decaëns, T., & Lees, D. C. (2019). DNA barcodes reveal deeply neglected diversity and numerous invasions of micromoths in Madagascar. *Genome*, *62*(3), 108–121. https://doi.org/10.1139/gen-2018-0065

Losos, J. B., & Ricklefs, R. E. (Eds.). (2010). *The theory of Island biogeography revisited*. Princeton University Press.

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of Island biogeography* (Vol. *1*). Princeton University Press.

Moser, D., Lenzner, B., Weigelt, P., Dawson, W., Kreft, H., Pergl, J., Pyšek, P., van Kleunen, M., Winter, M., Capinha, C., & Cassey, P. (2018). Remoteness promotes biological invasions on islands worldwide. *Proceedings of the National Academy of Sciences*, *115*(37), 9270–9275. https://doi.org/10.1073/pnas.1804179115

Mungi, N. A., Qureshi, Q., & Jhala, Y. V. (2021). Role of species richness and human impacts in resisting invasive species in tropical forests. *Journal of Ecology*, *109*(9), 3308–3321. https://doi.org/10.1111/1365-2745.13751

Noguerales, V., Meramveliotakis, E., Castro-Insua, A., Andújar, C., Arribas, P., Creedy, T. J., Overcast, I., Morlon, H., Emerson, B. C., Vogler, A. P., & Papadopoulou, A. (2021). Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests. *Molecular Ecology*, *2021*, 1–19. https://doi.org/10.1111/mec.16275

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. (2007). The vegan package. *Community Ecology Package*, *10*(631–637), 719.

Osozawa, S., Shinjo, R., Armid, A., Watanabe, Y., Horiguchi, T., & Wakabayashi, J. (2012). Palaeogeographic reconstruction of the 1.55 ma synchronous isolation of the Ryukyu Islands, Japan, and Taiwan and inflow of the Kuroshio warm current. *International Geology Review*, *54*(12), 1369–1388. https://doi.org/10.1080/00206814.2011.639954

Pacific Islands Benthic Habitat Mapping Center. (2022). *Commonwealth of Northern Mariana Islands (CNMI) and Guam* [Article]. Retrieved from http://www.soest.hawaii.edu/pibhmc/cms/data-by-location/cnmi-guam/

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Roy, K., Jaenecke, K. A., & Peck, R. W. (2020). Ambrosia beetle (coleoptera: Curculionidae) communities and frass production in ʻŌhiʻa (Myrtales: Myrtaceae) infected with Ceratocystis (Microascales: Ceratocystidaceae) fungi responsible for rapid ʻŌhiʻa death. *Environmental Entomology*, *49*(6), 1345–1354. https://doi.org/10.1093/ee/nvaa108

Russell, J. C., & Kueffer, C. (2019). Island biodiversity in the Anthropocene. *Annual Review of Environment and Resources*, *44*, 31–60. https://doi.org/10.1146/annurev-environ-101718-033245

Simberloff, D., & Von Holle, B. (1999). Positive interactions of nonindigenous species: Invasional meltdown? *Biological Invasions*, *1*(1), 21–32. https://doi.org/10.1023/A:1010086329619

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*, 99–114.
Valente, L., Phillimore, A. B., Melo, M., Warren, B. H., Clegg, S. M., Havenstein, K., Tiedemann, R., Illera, J. C., Thébaud, C., Aschenbach, T., & Etienne, R. S. (2020). A simple dynamic model explains the diversity of Island birds worldwide. *Nature*, *579*(7797), 92–96. https://doi.org/10.1038/s41586-020-2022-5
Vitousek, P. M. (2002). Oceanic islands as model systems for ecological studies. *Journal of Biogeography*, *29*(5–6), 573–582. https://doi.org/10.1046/j.1365-2699.2002.00707.x
Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., & Casquet, J. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*(2), 200–217. https://doi.org/10.1111/ele.12398
Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island biogeography: Ecology, evolution, and conservation*. Oxford University Press.
Wong, M. K., Guénard, B., & Lewis, O. T. (2019). Trait-based ecology of terrestrial arthropods. *Biological Reviews*, *94*(3), 999–1022. https://doi.org/10.1111/brv.12488
Wood, J. R., Alcover, J. A., Blackburn, T. M., Bover, P., Duncan, R. P., Hume, J. P., Louys, J., Meijer, H. J., Rando, J. C., & Wilmshurst, J. M. (2017). Island extinctions: Processes, patterns, and potential for ecosystem restoration. *Environmental Conservation*, *44*(4), 348–358. 10.1017/S037689291700039X
Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x
Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, *30*(5), 614–620. https://doi.org/10.1093/bioinformatics/btt593

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kennedy, S., Calaor, J., Zurápiti, Y., Hans, J., Yoshimura, M., Choo, J., Andersen, J. C., Callaghan, J., Roderick, G. K., Krehenwinkel, H., Rogers, H., Gillespie, R. G., & Economo, E. P. (2023). Richness and resilience in the Pacific: DNA metabarcoding enables parallelized evaluation of biogeographic patterns. *Molecular Ecology*, *32*, 6710–6723. https://doi.org/10.1111/mec.16575