OXFORD

Original Article

Genetic Differentiation and Demographic Trajectory of the Insular Formosan and Orii's Flying Foxes

Kung-Ping Lin, Shu-Miaw Chaw, Yun-Hua Lo, Teruo Kinjo, Chien-Yi Tung, Hsi-Chi Cheng, Quintin Lau, Yoko Satta, Masako Izawa, Shiang-Fan Chen, and Wen-Ya Ko

From the Department of Life Sciences and Institute of Genome Sciences, National Yang Ming Chiao Tung University, Taipei, Taiwan (Lin, Lo, and Ko); Biodiversity Research Center, Academia Sinica, Taipei, Taiwan (Chaw); Okinawa Zoo and Museum, Okinawa, Japan (Kinjo); Cancer Progression Research Center, National Yang Ming Chiao Tung University, Taipei, Taiwan (Tung); Endemic Species Research Institute, Nantou, Taiwan (Cheng); Department of Evolutionary Studies of Biosystems, SOKENDAI (The Graduate University for Advanced Studies), Hayama, Japan (Liu and Satta); Kitakyushu Museum of Natural History and Human History, Fukuoka, Japan (Izawa); and Center for General Education, National Taipei University, New Taipei City, Taiwan (Chen).

Address correspondence to M. Izawa, S-F. Chen, and W-Y. Ko at the address above, or e-mail: izawa@sci.u-ryukyu.ac.jp; schen@mail.ntpu.edu.tw; wenyako@gmail.com.

Abstract

Insular flying foxes are keystone species in island ecosystems due to their critical roles in plant pollination and seed dispersal. These species are vulnerable to population decline because of their small populations and low reproductive rates. The Formosan flying fox (*Pteropus dasymallus formosus*) is one of the 5 subspecies of the Ryukyu flying fox. *Pteropus dasymallus formosus* has suffered from a severe decline and is currently recognized as a critically endangered population in Taiwan. On the contrary, the Orii's flying fox (*Pteropus dasymallus inopinatus*) is a relatively stable population inhabiting Okinawa Island. Here, we applied a genomic approach called double digest restriction-site associated DNA sequencing to study these 2 subspecies for a total of 7 individuals. We detected significant genetic structure between the 2 populations. Despite their contrasting contemporary population sizes, both populations harbor very low degrees of genetic diversity. We further inferred their demographic history based on the joint folded site frequency spectrum and revealed that both *P. d. formosus* and *P. d. inopinatus* had maintained small population sizes for a long period of time after their divergence. Recently, these populations experienced distinct trajectories of demographic changes. While *P. d. formosus* suffered from a drastic ~10-fold population decline not long ago, *P. d. inopinatus* underwent a ~4.5-fold population expansion. Our results suggest separate conservation management for the 2 populations—population recovery is urgently needed for *P. d. formosus* while long-term monitoring for adverse genetic effects should be considered for *P. d. inopinatus*.

Subject Area: Conservation Genomics and Biodiversity
Key words: Pteropodidae, conservation genetics, RADSeq, Taiwan, Ryukyu Islands

Island species are often characterized by restricted geographic distributions and small population sizes, which are considered to be more prone to adverse genetic factors, such as low genetic diversity and inbreeding depression, than continental species (Ellstrand and Elam 1993; Frankham 1996, 1997, 2005). Both of these genetic factors can increase the extinction risks of the threatened island species. Loss of genetic diversity can result from inbreeding, population decline, or restricted gene flow, and is known to reduce the long-term evolutionary potential of a species (Frankham et al. 1999; Frankham et al. 2002). Inbreeding was shown to have deleterious effects on multiple aspects of reproductive and survival characteristics in naturally outbreeding species (Lynch and Walsh 1998). Therefore, resolving the magnitude of these genetic factors has been considered important for the development of conservation plans, especially for threatened insular species with limited gene flow and small population size. In addition, defining diverged evolutionary units for a threatened population is also considered critical to prevent potential outbreeding depression in crosses between units (Frankham 2010). Thus, distinct evolutionarily significant units (ESUs) with significant genetic and ecological divergence in the wild need to be managed separately. To provide the above critical information for conservation management, genomic data is especially useful as, by gathering numerous neutral loci, they can provide precise population genetics estimates of genetic diversity, inbreeding level, and demographic independence among the target populations (Allendorf et al. 2010; Funk et al. 2012).

Bats in the genus *Pteropus* of the family Pteropodidae are commonly known as flying foxes or fruit bats. *Pteropus* is primarily an insular genus with most species distributed on islands (Vincenot et al. 2017). Flying foxes are found in tropical and subtropical regions, ranging from islands in East Africa to the south-central Pacific. In these island ecosystems, *Pteropus* spp. play critical roles in pollination and seed dispersal for both endemic plants and plants of economic importance to humans (Cox et al. 1991; Fujita and Tuttle 1991). However, the populations of many *Pteropus* spp. are declining due to human disturbances (Mickleburgh et al. 1992). The major anthropogenic menaces to flying foxes include habitat loss caused by deforestation and urbanization, overhunting for bushmeat, and conflicts with fruit growers. Moreover, the consequences of population decline in these species are expected to be exacerbated by their generally low reproductive rates, long maturation times, and small population sizes (Mickleburgh et al. 2009; Aziz et al. 2017). Hitherto, 7 of 65 flying fox species became extinct in the past 200 years, and 34 of the remaining species are currently considered threatened, including *Pteropus dasymallus*, the Ryukyu flying fox, which has been classified as Vulnerable (Vincenot 2017).

The Ryukyu flying fox is indigenous to Taiwan, the Ryukyu Islands, and several islands off the north coast of the Philippines (Kinjo and Nakamoto 2015). There are 5 recognized subspecies of *P. dasymallus* that can be distinguished by their geographical distributions and by slight differences in coloration pattern (Yoshiyuki 1989). Among them, the Formosan flying fox (*P. d. formosus*) is the most threatened subspecies and considered to be on the verge of extinction by the IUCN. Other subspecies include the Orii's flying fox (*P. d. inopinatus*; near threatened) with a population size of ~5000 individuals on Okinawa Island (Funakoshi et al. 2006, 2012), the Yaeyama flying fox (*P. d. yayeyama*; near threatened) with a few thousand individuals on Yaeyama Islands, and the Erabu flying fox (*P. d. dasymallus*; critically endangered) and Daito flying fox (*P. d. daitoensis*; critically endangered) on the Osumi Islands and Daito Islands, respectively (Figure 1).

Early studies suggested that Formosan flying foxes are native to Green Island, located in southeast of Taiwan (Kano 1929; Kuroda 1933). Individuals of *P. d. formosus* were also recorded on the east coast of Taiwan but they were presumed to be solitary without a stable and sustainable population. The population size of *P. d. formosus* on Green Island was estimated to be up to a few thousand before a dramatic decline during the 1970s, which led to a near extirpation of the population on the island with only a few individuals observed over the last 2 decades (Lin and Pei 1999; Chen et al. 2009). Overhunting was thought to be the main factor causing this population decline, combined with the effect of habitat loss due to windbreak planting in the coastal region. Although the hunting pressure was greatly reduced after the implementation of a conservation law in 1983, the lost habitat (mostly *Ficus* species) has never recovered. Therefore, habitat abandonment by *P. d. formosus* could also be responsible for the apparent disappearance of this population, especially since flying foxes are generally considered highly mobile species that are capable of long-range migrations and interisland movements (Webb and Tidemann 1996; McConkey and Drake 2007; Brown et al. 2011; Roberts et al. 2012). For example, the Orii's flying fox has been reported to be able to travel 50 km across the sea between islands (Nakamoto et al. 2011). In 2004, a small but stable population (~50 individuals) was newly recorded on Gueishan Island off the northeast coast of Taiwan (Wu 2010). However, the origin of this newly established population is unclear since Gueishan Island is geographically closer to Yaeyama Islands (inhabited by the Yaeyama flying fox) than to Green Island (Figure 1). Due to its small population size, severe population decline, and fragmented habitats, the Formosan flying fox is listed as an Endangered Species by the Wildlife Conservation Act in Taiwan (Cheng et al. 2017). Thus, a precise and detailed conservation policy for this endangered taxon is urgently needed. In contrast, whether to treat *P. d. formosus* as a distinct ESU from other populations on the Ryukyu Islands remains an obscure issue given their high migratory abilities.

Here, we applied double digest restriction-site associated DNA sequencing (ddRADSeq) method to study the conservation genetics of *P. d. formosus* in Taiwan and *P. d. inopinatus* in the main island of Okinawa. First, we assessed the adverse genetic factors of these 2 populations by estimating their levels of nucleotide diversity and inbreeding. Second, we evaluated whether these 2 populations should be treated as distinct ESUs and managed separately by uncovering the pattern of genetic structure between these 2 populations. We further inferred the hypothetical demographic model of these 2 populations based on their joint folded site frequency spectrum and revealed very different trajectories of their recent population history. Our results can provide critical insights into the conservation management decisions of both *P. d. formosus* and *P. d. inopinatus*.

Materials and Methods

Sample Collection, DNA Extraction, and RAD Sequencing

We generated genome-wide single nucleotide polymorphism (SNP) data from 7 individuals. Three individuals are *P. d. formosus*, an endangered population endemic to Taiwan, while 4 are *P. d. inopinatus*, which inhabits Okinawa and has been considered a stable population. The 3 *P. d. formosus* individuals, designated as 178, 617, and 690, were captured from Hualien (eastern Taiwan), Green Island (an island offshore of eastern Taiwan), and the open sea off the coast of Yilan (northeastern Taiwan), respectively. The
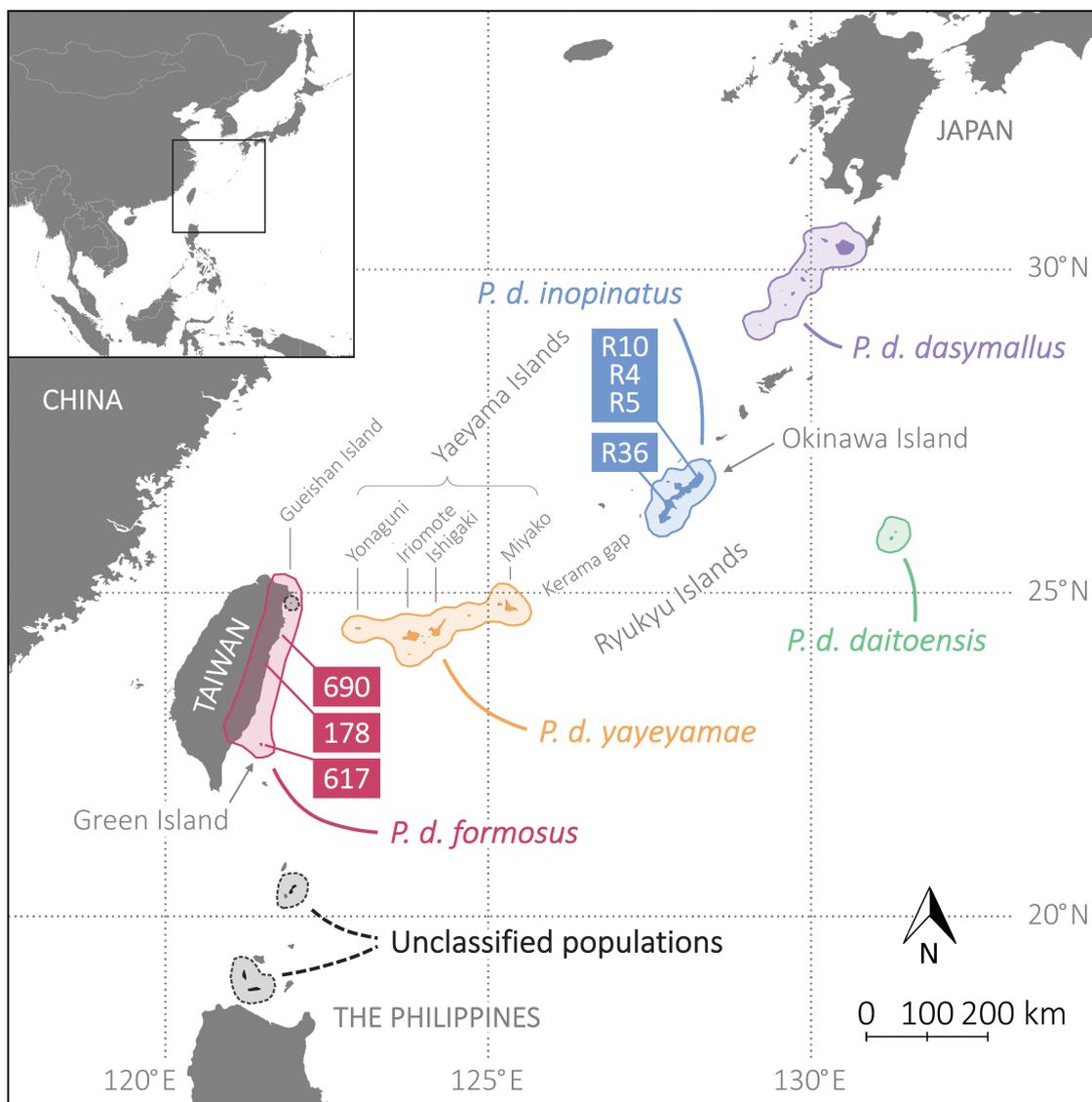
Figure 1. Geographic distribution of *Pteropus dasymallus* subspecies and sample locations. The geographic ranges of the 5 subspecies of *P. dasymallus* are circled. These 5 subspecies include: Erabu flying fox (*P. d. dasymallus*), Daito flying fox (*P. d. daitoensis*), Orii's flying fox (*P. d. inopinatus*), Yaeyama flying fox (*P. d. yayeyamae*), and Formosan flying fox (*P. d. formosus*). The classification of *P. dasymallus* populations in the Gueishan Island and northern islands of Philippines requires further investigation. The original geographic data was downloaded from the IUCN Red List website (Vincenot 2017) and edited by *QGIS* 3.4 (QGIS Development Team 2019). Three samples of *P. d. formosus* were collected from Taiwan and Green Island (designated as 178, 617, and 690). Four samples of *P. d. inopinatus* were collected from Okinawa Island (designated as R10, R36, R4, and R5). See online version for full colors.

4 *P. d. inopinatus* individuals, designated as R10, R36, R4, and R5, are captive individuals collected from the wild and kept at the University of the Ryukyus. The detailed sampling locations for all individuals are presented in Figure 1 and Supplementary Table S1. For the 3 *P. d. formosus* individuals, frozen muscle tissues were obtained from the Wildlife Cryobank of Taipei Zoo. For each of the 4 *P. d. inopinatus* individuals, skin samples were excised from the wing membranes using a biopsy punch (3 mm in diameter); 8 punches were collected per individual. All biopsies were preserved in Allprotect Tissue Reagent (Qiagen) and stored at −20 °C. Genomic DNA was extracted from samples using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's instructions. To better lyse the tissues, the wing membrane samples were incubated in buffer ATL and proteinase *K* at 56 °C for approximately 17 hours. An additional 20 µL of proteinase *K* was then added and

further incubation for a few hours was conducted depending on lysis conditions.

Considering that *P. d. formosus* and *P. d. inopinatus* are morphologically similar taxa and that few genetic variants have been discovered among *P. dasymallus* subspecies at mitochondrial loci (Hidetoshi Ota, personal communication, 2010), we applied a genomic approach, double digest restriction-site associated DNA sequencing (ddRADSeq) method, which is designed for sequencing orthologous loci across individual genomes of the same or closely related species (Peterson et al. 2012). For a given individual genome, 2 restriction enzymes, EcoRI and MseI, were used to cleave the DNA. The cleaved DNA fragments were ligated with adaptors that contain barcodes and sequences complementary to the PCR primers. After these ligated fragments were amplified by PCR, fragments with lengths between 250 base pairs (bp) and 450 bp were retained and

sent for paired-end sequencing. Sequencing was carried out using the Illumina HiSeq™ 2000 sequencing system in the core facility center of Genomics company in New Taipei City.

Quality Controls and RADSeq Assembly

The demultiplexed paired-end sequence reads were processed with *SOAPnuke2* to discard low-quality reads and unwanted sequences (Chen et al. 2018). First, reads with correct barcodes were retained exclusively and their barcodes were trimmed. Reads without correct restriction enzyme cleavage sites were subsequently removed from the dataset. Low-quality reads were also removed if the proportions of base pairs with Phred quality scores lower than 15 were >40% or the proportions of unidentified base pairs were >10%. Next, we removed the restriction enzyme cleavage sites from all remaining reads. The remaining forward reads were trimmed to 87 bp to accommodate the differences in sequence lengths caused by barcode trimmings and restriction enzyme cleavage site removals.

We performed paired-end RAD sequence assemblies using *Stacks* 2.2 (Catchen et al. 2011). This software consists of several core programs. The first program, *ustacks*, established "stacks," the piles of reads with exact matches that serve as putative alleles, for each individual by aligning the raw forward reads with each other. The minimum number of reads (read depth) that is required to start a stack was governed by *m*. Next, similar stacks were aligned against each other within each individual to form putative loci (pairs of putative alleles) if the number of mismatches between them was ≤ *M* (maximum number of allowed mismatches). The second program, *cstacks*, was performed to align the putative loci across all 7 individuals to synthesize a catalog with the maximum number of allowed mismatches defined by the parameter *n*. The putative loci from all individuals were subsequently matched against the catalog to identify the unique loci combinations on each RAD locus across all individuals; this procedure was conducted using the *sstacks* program. Thereafter, the reads from both ends were merged into contigs using the *gstacks* program. Lastly, the genetic variants and assembled contigs were called using the *populations* program. We executed the above programs using a built-in pipeline, *denovo_maps.pl*, with different sets of *m*, *M*, and *n*; secondary read incorporations were disabled and the deleveraging algorithm was enabled. Only sites that were sequenced in all individuals were assembled to prevent systematic biases on population genetics statistics, including estimates of nucleotide diversity and Tajima's D (Arnold et al. 2013; Gautier et al. 2013). RAD loci without correct matches with reverse reads were omitted from the remaining studies.

Stacks was run multiple times to search for the optimal set of assembled parameters (*m*, *M*, *n*). First, since *m* controls the starting read depth of each stack, a higher *m* is expected to give higher sample coverage while resulting in fewer assembled loci and SNPs; this trend diminishes with increased *m* (Paris et al. 2017). We therefore plotted the distribution of sample coverage against different values of *m* (while fixing *M* and *n* to 3). An optimal value of *m* was chosen if its corresponding value was close to the plateau of sample coverage on the inward curve. Next, we searched for the optimal values of *M* and *n* while fixing *m* to the previously determined value. We took a likelihood framework to search for the optimal *M* and *n* by incorporating the observed SNP density estimated from the data. For each given pair of (*M*, *n*), the empirical per-site SNP density was calculated by the total number of SNPs in all forward reads divided by the total number of base pairs of all forward reads. The distribution of the number of SNPs on each sequence read can be constructed using a

binomial distribution with the number of trials equivalent to read length and the success probability for each trial equivalent to SNP density. Subsequently, the likelihood estimate of the empirical distribution of the SNP number per read can be calculated based on the multinomial probability mass function,

$$L\left(x \mid q\right) = r! \cdot \prod_{i=0}^{l} \frac{q_i^{x_i}}{x_i!}$$

where x_i is the empirical number of reads with *i* SNPs observed, q_i is the binomial probability of observing *i* SNPs on an *l*-bp long read (*l* = 87 in this study), and *r* is the total number of reads (i.e., Σx_i). For each given pair of (*M*, *n*), we randomly sampled 100 000 reads 20 times from the assembly without replacement. The pair of (*M*, *n*) that gives the highest likelihood estimate is the parameter set that fits best to the empirical distribution of SNP number across all forward reads. In addition, Kullback–Leibler (KL) divergences were also calculated to measure the similarity between the empirical distribution of the number of SNPs per read and the theoretical binomial distribution (Kullback and Leibler 1951). We also constructed an alternative likelihood estimator based on Watterson's theoretical distribution of segregating sites under one panmictic population (Watterson 1975). All the parameter optimization procedures were carried out using our custom Python scripts (available on request).

Since SNPs with excessive heterozygosity across all individuals can result from misaligning reads of paralogous or repetitive loci, we applied the Hardy-Weinberg equilibrium (HWE) test to exclude the assembled reads or SNPs that showed significant deviations from HWE (*P* < 0.05). Two types of heterozygosity filters were applied to the dataset to generate datasets for population genetics analyses that rely on accurate nucleotide diversity estimates or only genetically unlinked SNPs. The heterozygosity filters were conducted using the R package *HardyWeinberg* (Graffelman and Camarena 2008; Graffelman 2015).

Nucleotide Diversity and Inbreeding Coefficient Estimations

We estimated genome-wide nucleotide diversity based on both pairwise differences and Watterson's estimator for both *P. d. formosus* and *P. d. inopinatus*. For each estimate, the per-site nucleotide diversity was calculated for each RAD locus. The mean and standard deviation were then computed based on its empirical distribution across all RAD loci. Tajima's *D* was also calculated using the R package *pegas* 0.12 (Paradis 2010) with parametric *P*-values constructed from the beta distribution (Tajima 1989a). Fixation indices were calculated using *GENEPOP* 1.1.7 implemented in R (Raymond and Rousset 1995; Rousset 2008). For comparison, we also computed the above summary statistics for the RADSeq data of 3 widespread bat species (*Lasiurus cinereus*, *L. borealis*, and *Lasionycteris noctivagans*) that were obtained from Sovic et al. (2016).

Genetic Structure Detection

We next performed both principal component analysis (PCA, conducted using *PLINK* 1.9; Chang et al. 2015; Purcell and Chang 2020) and *STRUCTURE* 2.3.4 (Pritchard et al. 2000) to detect the genetic structure between *P. d. formosus* and *P. d. inopinatus*. A subset of SNPs was generated by randomly selecting one SNP from each locus to assure that these SNPs are genetically unlinked. For PCA, a Kruskal-Wallis H test was further performed on the

coordinates of the first principal component to detect the significance between 2 assigned groups. For *STRUCTURE* analysis, the data-collecting period was set to 50 000 and the burn-in period was set to 10 000, 30 000, 40 000, 40 000, and 40 000 for $K = 1$, 2, 3, 4, and 5, respectively, where K is the putative number of ancestral populations; these values were determined by observing the trend of several summary statistics throughout the iterations of the program (Pritchard et al. 2000). The best-fit K value was determined using the ΔK estimator (Evanno et al. 2005) and the built-in posterior probability of *STRUCTURE*. To calculate the ΔK estimator, we performed 20 repetitions for each given K value. We summarized the ancestral components of these 20 repetitions using *CLUMPP* 1.1.2 (Jakobsson and Rosenberg 2007), for which the FullSearch algorithm was applied for $K = 1$ and 2 while the Greedy algorithm was applied for $K = 3$, 4, and 5. In addition, we also quantified the degree of population differentiation between the 2 populations using the Φ_{ST} estimator of the Analysis of Molecular Variance (AMOVA) implemented in *pegas* 0.12. The *P*-value of the Φ_{ST} estimator was calculated from 10 000 permutations.

Demographic History Inference

To infer the demographic history of the 2 *P. dasymallus* populations, we compared the folded site frequency spectrum (fSFS) for each population with the simulated fSFS under the steady-state neutral model. The simulated fSFS of an idealized Wright-Fisher population was constructed from 1000 simulations generated by the coalescent simulator *MSMS* with segregating sites set to the observed values in each population (Ewing and Hermisson 2010). We further applied *fastsimcoal2* 2.6 to infer the demographic history of *P. d. formosus* and *P. d. inopinatus* based on the joint fSFS of these 2 populations (Excoffier and Foll 2011; Excoffier et al. 2013). A number of possible demographic models were tested. These models begin with the simplest model (one population of constant population size over time), then toward several more realistic models to which population splits, gene flow between split populations, and population size changes over time in the progeny populations were introduced. For a given model, the simulated joint fSFS was generated based on 100 000 coalescent simulations assuming the observed total sequence length. The maximum number of cycles of the conditional maximization algorithm (ECM) was set to 40 to search for the estimates that best-fit the observed joint fSFS. This process was repeated 100 times and the estimate with highest likelihood was selected to represent the model. To determine the model that fits our data best, all tested models were compared with each other through the Akaike Information Criterion (AIC; Akaike 1974), calculated using a custom Python script. For the best-fit model, we performed 100 parametric bootstraps to construct the confidence intervals (CIs) of each estimated demographic parameter of interest. The mutation rate (μ) was set to 2.5×10^{-8} per generation per nucleotide site, which was estimated based on human genomic data (Nachman and Crowell 2000). We also repeated all simulations by assuming a different mutation rate of 2.0×10^{-9} according to an estimate from a few introns in vespertilionid and miniopterid bats (Ray et al. 2008). The detailed workflow of all analyses can be found in Supplementary Figure S1.

Results

Determining the Optimal Parameters for RADSeq Assembly

We have successfully obtained the genomic sequence reads for all 7 sampled individuals using ddRADSeq. The total clean base pairs

obtained from each individual ranged from 0.91 Gb to 6.90 Gb with an average of 4.82 Gb. The average number of clean reads per genome was 47.28 million (M), ranging from 7.76 to 68.82 M. The detailed information for each individual genome is given in Supplementary Table S1. To determine the most suitable assembly parameters in *Stacks*, we first plotted the sample coverage against the parameter m (the minimum depth required to start a stack) as shown in Figure 2A. The sample coverage increased with m but gradually reached a plateau. In contrast, the total number of polymorphic loci and monomorphic loci decreased considerably with increased m (Figure 2B). Determining the suitable m value appeared to become a trade-off between the coverage and the number of assembled loci. We set $m = 3$ to be our assembly parameter since its corresponding value on the curve begins to show a slower increment of sample coverage. As a result, the average number of sample coverages among the 7 individuals was 38.49 (range 13.10–51.83; see also Supplementary Table S1) and the number of polymorphic loci was ~28 K (Supplementary Figure S1). Figure 2C shows the plot of log-likelihood estimates of the empirical distribution of SNPs per read based on different sets of (M, n) given $m = 3$. The highest likelihood estimate was found to be at $M = 1$ and $n = 1$ within the range of $M = 1–7$ where $n = M$ or $M+1$. A similar result was also observed using the KL divergence method: the similarity between the empirical distribution of SNPs per read and the theoretical distribution based on the estimated SNP density also appeared to be the highest (i.e., the lowest value of KL divergence) at $M = 1$ and $n = 1$ (Figure 2C). We also computed the log-likelihood estimates of the empirical distribution of SNPs per read based on Watterson's theoretical distribution of segregating site, which resulted in the same optimal parameter set (Supplementary Figure S2). Consequently, the parameters for *Stacks* were set to $M = 1$, $n = 1$, and $m = 3$ for contig assemblies and SNP callings. Based on the above assembly parameters, we obtained 188,588 RAD loci for a total of 43,675,221 bp with a mean locus length of 231.6 ± 36.7 bp (Figure 2D). Among them, 32,935 SNPs were identified in 28,047 polymorphic RAD loci after a series of quality control steps. The workflow of sequence assemblies and quality controls is detailed in Supplementary Figure S1.

Estimating Levels of Nucleotide Diversity, Inbreeding, and Genetic Structure

The Watterson's estimator (θ_W) and the pairwise differences (θ_π) were estimated for *P. d. formosus*, *P. d. inopinatus*, and the pooled population using 32,935 SNPs and 188,588 loci (Table 1). The estimates of Watterson's θ_W appeared to be almost the same but considerably low in both populations (2.14×10^{-4} and 2.13×10^{-4} per site for the *formosus* and *inopinatus* populations, respectively). Similar θ_π values were also observed in the *formosus* (2.17×10^{-4}) and *inopinatus* (2.09×10^{-4}) populations. Both θ_W and θ_π estimates increased slightly when the 2 populations were pooled (2.39×10^{-4} and 2.27×10^{-4}, respectively). We further computed Tajima's D for the 2 populations separately. Although different signs of Tajima's D were observed in the *formosus* (0.11) and *inopinatus* (–0.12) populations, both estimates did not deviate from neutrality ($P = 0.88$ and 0.95, respectively). To estimate the level of inbreeding, we computed the fixation index (F) based on genetically unlinked SNPs (Table 1). Positive F values were observed in both populations in which the average F of the *formosus* (0.12) was slightly higher than that of the *inopinatus* (0.10).

To evaluate the genetic structure between these 2 populations, we performed PCA among all 7 individuals. As a result, the *formosus* and *inopinatus* samples were clearly divided into 2 groups at PC1 ($P = 0.03$
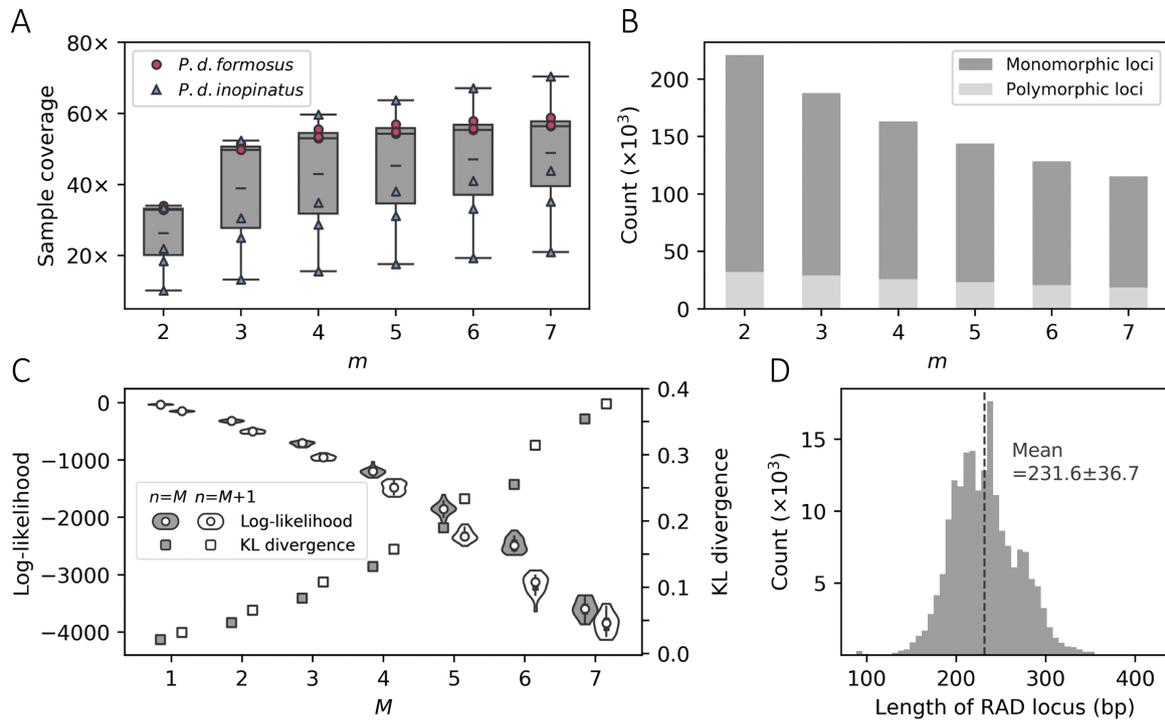
Figure 2. Determining the optimal assembly parameters of *Stacks* and the length distribution of restriction-site associated DNA (RAD) loci. (**A**) The sample coverages under different values of *m* (the minimum read depth required to start a stack) with *M* (the maximum mismatches allowed within each individual) and *n* (the maximum mismatches allowed between each pair of individuals) fixed to 3. The box plot summarizes the coverage of all 7 individuals while the dash indicates the mean value. (**B**) The number of polymorphic loci and monomorphic loci under different values of *m* given *M* = *n* = 3. (**C**) The multinomial log-likelihoods and KL divergences under various (*M*, *n*) values with *m* set to 3. (**D**) The distribution of the lengths of each assembled RAD locus without unidentified nucleotides under the optimal assembly parameters (*m* = 3 and *M* = *n* = 1).

Table 1. Estimates of nucleotide diversity, fixation index, and Tajima's *D* of 2 subspecies of *Pteropus dasymallus*

Group	# Samples	# SNPs	Mean θ_W	Mean θ_π	Mean *F*	*D*
P. d. formosus	3	21106	2.14×10^{-4}	2.17×10^{-4}	0.12	0.11
P. d. inopinatus	4	24020	2.13×10^{-4}	2.09×10^{-4}	0.10	−0.12
All	7	32935	2.39×10^{-4}	2.27×10^{-4}	0.16	−0.21

θ_W, Watterson's estimator; θ_π, pairwise difference; *F*, fixation index; *D*, Tajima's *D*.

using Kruskal-Wallis H test), while PC2 and PC3 reflected only the variances of within-group individuals (Figure 3A and Supplementary Figure S3). We also conducted *STRUCTURE*, a model-based analysis for population substructure detection among these individuals and Figure 3B illustrates the result at *K* = 3, which was the best-fit number of putative ancestral populations (Supplementary Figure S4A and B). Comparable to the result of PCA, explicit genetic structure was detected between the *formosus* and *inopinatus* populations in which distinct patterns of ancestral components were observed between them. In contrast, there was no evident genetic structure within each population. Similar patterns were also observed when *STRUCTURE* was performed at *K* = 2–5 (Supplementary Figure S4C). The degree of genetic structure between these 2 populations was estimated to be 0.178 based on the test statistic Φ_{ST} (AMOVA), revealing a considerable proportion of between-population variance among these individuals (*P* = 0.03; Table 2).

Inferring Demographic History

To uncover their demographic history, we first estimated the fSFS of *P. d. formosus* and *P. d. inopinatus* separately and detected significant differences between them (Figure 4A). The fSFS of *P. d. formosus* possessed excessive intermediate-frequency alleles compared to the simulated fSFS under the neutral model of constant population size over time (*P* = 10^{-62} using chi-square goodness of fit test), which is consistent with the population decline that has been observed in the field during the 1970s. In contrast, the fSFS of *P. d. inopinatus* carried excessive low-frequency alleles in comparison with the simulated steady-state fSFS (*P* = 10^{-19}), which is suggestive of a population expansion.

Since these 2 populations share a common ancestor, we further constructed a joint fSFS for our samples to infer their demographic history, time of divergence, and the level of gene flow between populations using *fastsimcoal2* (Figure 4B). The joint fSFS separates private and shared variants of the 2 populations and increases the number of frequency classes. Consequently, the statistical power enhances considerably for detecting population demography and can thus be used to infer complex demographic events such as population divergence and admixture. We performed *fastsimcoal2* likelihood estimations on a series of demographic models listed in Table 3 assuming $\mu = 2.5 \times 10^{-8}$ per site per generation. To determine the best model, the Akaike weights (w_i) of these models were calculated and, as a result, Model F was the model that best-fit the observed
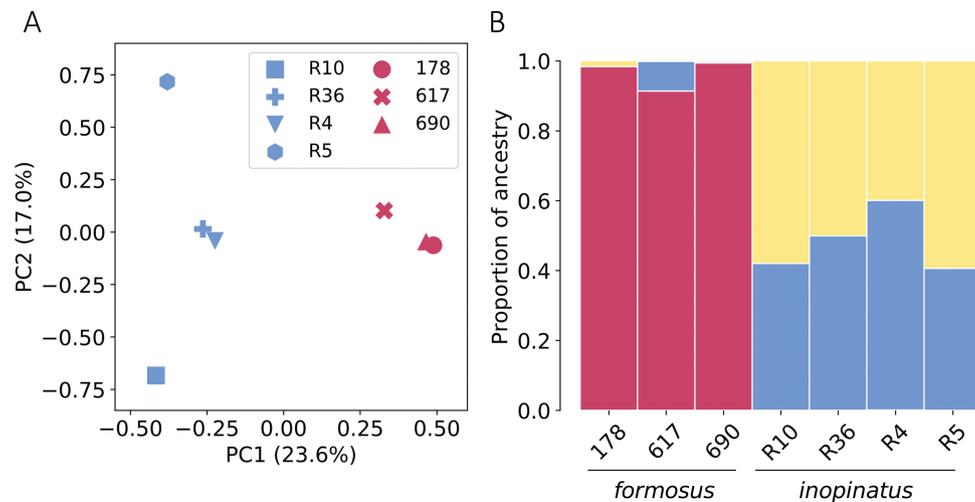
Figure 3. Principal component analysis (PCA) and *STRUCTURE* analysis for detecting genetic structure between *Pteropus dasymallus formosus* and *P. d. inopinatus*. (**A**) PCA was performed with 3 *P. d. formosus* and 4 *P. d. inopinatus* samples. (**B**) *STRUCTURE* result of the 7 *P. dasymallus* samples by assuming *K* = 3 where *K* is the number of ancestral populations. Both PCA and *STRUCTURE* were performed based on a total of 28,074 genetically unlinked SNPs from the 7 samples. See online version for full colors.

Table 2. Analysis of molecular variance (AMOVA) results of *Pteropus dasymallus formosus* and *Pteropus dasymallus inopinatus*

	df	Sum of squares	Variance components	Variance %	*P*-value
Between populations	1	15512.5	1929.2	17.8	0.03
Within populations	5	44490.1	8898.0	82.2	
Total	6	60002.6	10827.2		

joint fSFS among all models ($w_i \approx 1.0$), whereas the other models fit the data significantly worse ($w_i \leq 1.1 \times 10^{-19}$; see Table 3). Model F assumed an ancestral population split into 2 progeny populations (*i.e.*, *formosus* and *inopinatus*) and introduced parameters that describe a recent population decline in *P. d. formosus* and a population expansion in *P. d. inopinatus*, as well as gene flow between these 2 progeny populations (Figure 5A). Under Model F, the maximum likelihood estimate (MLE) of the ancestral effective population size (N_a) of the *formosus* and *inopinatus* populations was estimated to be as large as 9,468 (Table 4). The effective population sizes dropped to 2324 and 2110 for *formosus* ($N_{f.a}$) and *inopinatu* ($N_{i.a}$), respectively, after they split. The split time was dated to be 8518 generations ago. However, the *formosus* population experienced a very recent and severe decline 4 generations ago ($t_{f.de}$), which led to its low current population size (N_f) of 223. In contrast, the *inopinatus* population went through a population expansion 441 generations ago ($t_{i.ex}$) and the current effective population size (N_i) was estimated to be 9547. The level of gene flow between the 2 populations was limited. The estimated migration rate from the *inopinatus* population to the *formosus* population was 1.0×10^{-3} (per generation) and that of the reverse direction was 2.7×10^{-4}. The prior search ranges for these MLEs are provided in Supplementary Table S2. We further constructed the bootstrap distributions (Figure 5B) and computed the 95% CIs for these parameters (Table 4).

We repeated all the simulations for inferring the demographic history with a different mutation rate ($\mu = 2.0 \times 10^{-9}$) that was estimated from 4 introns in vespertilionid and miniopterid bats (Ray et al. 2008), and found that Model F remained the fittest demographic scenario among all models tested. However, by assuming a lower mutation rate, the estimated effective population sizes and dates of demographic events appeared to be 6–66 times larger and

2–300 times older, respectively, than those under the assumption of $\mu = 2.5 \times 10^{-8}$ (Supplementary Table S3).

Discussion

Low Level of Genetic Diversity and High Degree of Inbreeding in *P. d. formosu* and *P. d. inopinatus*

In this study, we have successfully applied ddRADSeq to uncover genetic variants from 3 *P. d. formosus* and 4 *P. d. inopinatus* individuals. Given that both subspecies are protected by local conservation laws, it is difficult to obtain any new sample from the natural population. Nonetheless, we were able to obtain a sufficiently large number of SNPs, which allowed us to obtain and compare several key population genetics estimates between these 2 closely related populations. A very low level of nucleotide diversity ($\theta_w = 2.14 \times 10^{-4}$ per site) was detected in the *formosus* population (Table 1), which is ~10-fold lower than the mean θ_w of *Lasiurus cinereus* (2.57×10^{-3}) and *Lasiurus borealis* (2.14×10^{-3}), 2 widespread continental vespertilionid bats in America (Sovic et al. 2016; also see Supplementary Table S4). Surprisingly, *P. d. inopinatus*, a larger population than *P. d. formosus*, also harbors a similarly low level of nucleotide diversity ($\theta_w = 2.13 \times 10^{-4}$). Tajima (1989b) pointed out that changes in nucleotide polymorphism estimates are slow in response to population fluctuations under small sample sizes. The similarity in nucleotide diversity between these 2 populations can be partially attributed to the small sample sizes in both populations. With a small number of sample sizes, genetic variants with intermediate allele frequency in a population are more likely to be sampled than those with low allele frequency. Therefore, the nucleotide diversity estimates are less sensitive for detecting alteration in nucleotide diversity caused by
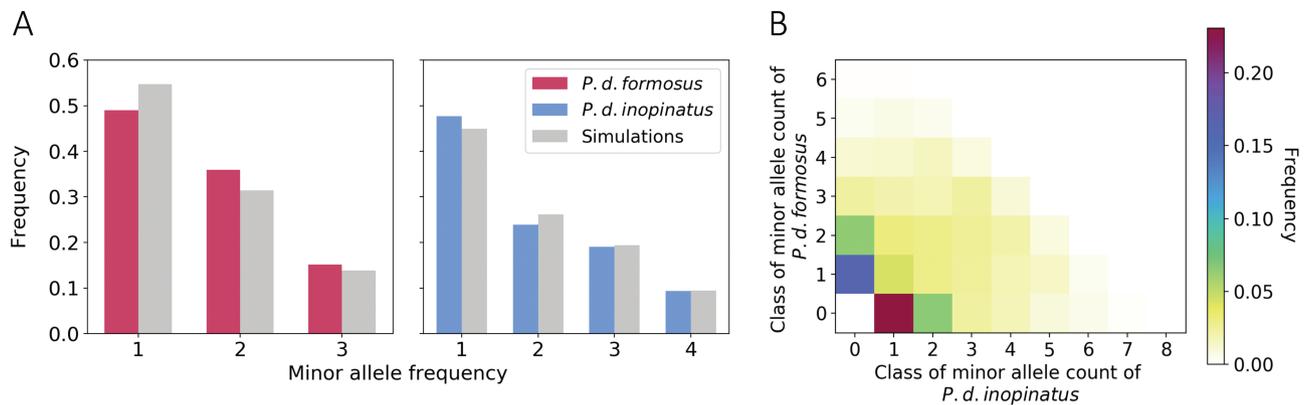
Figure 4. Folded site frequency spectra (fSFS) of *Pteropus dasymallus formosus* and *P. d. inopinatus*. (**A**) The fSFS of 3 *P. d. formosus* and 4 *P. d. inopinatus* were inferred separately based on 21,106 and 24,020 SNPs, respectively. Both spectra were compared to the fSFS of 1000 coalescent simulations with constant population size and the identical segregating sites of the observed fSFS. Chi-square goodness of fit test was performed on each fSFS; both spectra deviate from the steady-state neutral model ($P = 10^{-62}$ for *P. d. formosus* and $P = 10^{-19}$ for *P. d. inopinatus*). (**B**) The joint fSFS of *P. d. formosus* and *P. d. inopinatus* were constructed from 32,935 SNPs. The heatmap indicates the proportions of SNP counts across different joint frequency classes. See online version for full colors.

Table 3. Demographic model comparison of *Pteropus dasymallus formosus* and *Pteropus dasymallus inopinatus*

	A	B	C	D	E	F	G	H
Model illustration								
Max log$L$	−159797	−158076	−157991	−157982	−157991	−157936	−157984	−157982
Parameters	1	4	6	8	8	10	8	10
w_i (AIC)	≈ 0.0	6.3×10^{-59}	7.3×10^{-23}	1.1×10^{-19}	1.9×10^{-23}	≈ 1.0	2.1×10^{-20}	1.9×10^{-20}

Model A illustrates a population with constant size from which samples of *P. d. formosus* and *P. d. inopinatus* were drawn randomly. The rest of models describe one ancestral population split into 2 progeny populations ("*f*" represents *P. d. formosus* and "*i*" represents *P. d. inopinatus*). Models C–F and H introduce gene flow (indicated by arrows) between the 2 progeny populations. Models D–H allow population size changes in at least one progeny population.

low-frequency alleles owing to recent change of population sizes. In this study, we indeed detected signatures of recent population fluctuations in both populations based on the joint fSFS from which the estimated population sizes are similar in both populations prior to their recent demographic changes. Therefore, the estimates of nucleotide diversity could differ considerably between *P. d. formosus* and *P. d. inopinatus* with a larger sample size.

Nucleotide diversity could be underestimated using ddRADSeq due to allele dropout (ADO), that is, restriction enzymes fail to cleave their recognition sites because of mutations (Arnold et al. 2013; Gautier et al. 2013; Cariou et al. 2016). However, the effects of ADO are only significant when the sequence divergence is >2%, and can be mitigated by the use of complete sampling of genetic variants (Cariou et al. 2016). Therefore, our estimates of nucleotide diversity were most likely little affected by ADO.

Our estimated fixation index (*F*) revealed a considerable amount of homozygosity departing from HWE in both populations (0.12 for the *formosus* and 0.10 for the *inopinatus*). Based on the results of PCA and *STRUCTURE*, we did not detect any obvious genetic structure within each population (Figure 3). Therefore, the observed excessive homozygosity in each population can be better explained as the outcome of inbreeding. Insular populations are known to be more inbred (and thus more prone to extinction) than mainland populations due to their smaller population sizes and founder effect (Frankham 2008). In the present study, we detected rather high levels of inbreeding in both *P. d. formosus* and *P. d. inopinatus*, suggesting that both populations might be vulnerable to the risk of inbreeding depression. However, this concern of possible inbreeding depression

is based only on a small sample size number. Therefore, additional samples would be needed to confirm our estimates.

Distinct Recent Demographic History in *P. d. formosus* and *P. d. inopinatus*

We inferred the demographic history based on the joint fSFS of *P. d. formosus* and *P. d. inopinatus* by comparing the likelihood estimates among several candidate models. The best-fit demographic model appeared to be the population split model that assumes a recent population decline in the *formosus* and a recent population expansion in the *inopinatus* irrespective of the underlying assumption of mutation rates (Table 3). It is important to point out that MLEs for most parameters vary considerably with different mutation rates assumed in the model. Particularly, the current population size of *P. d. formosus* (N_f) becomes as large as 14 719 by assuming $\mu = 2.0 \times 10^{-9}$, whereas N_f is only 223 when assuming $\mu = 2.5 \times 10^{-8}$ (Table 4 and Supplementary Table S3). The assumed mutation rate of 2.0×10^{-9} was based on the study of only 4 introns in several species of vespertilionid and miniopterid bats (Ray et al. 2008), while the mutation rate of 2.5×10^{-8} was estimated from many pseudogenes across multiple chromosomes of human genomes. Given that estimates of the nuclear mutation rates in Pteropodidae species are currently unavailable, we are slightly in favor of reporting the MLEs of demographic parameters under the assumption of $\mu = 2.5 \times 10^{-8}$ due to its genome-wide approach. We consider that this higher mutation rate is a conservative approach for estimating the population size, which is also in good agreement with the observed population status of *P. d. formosus* in the field.
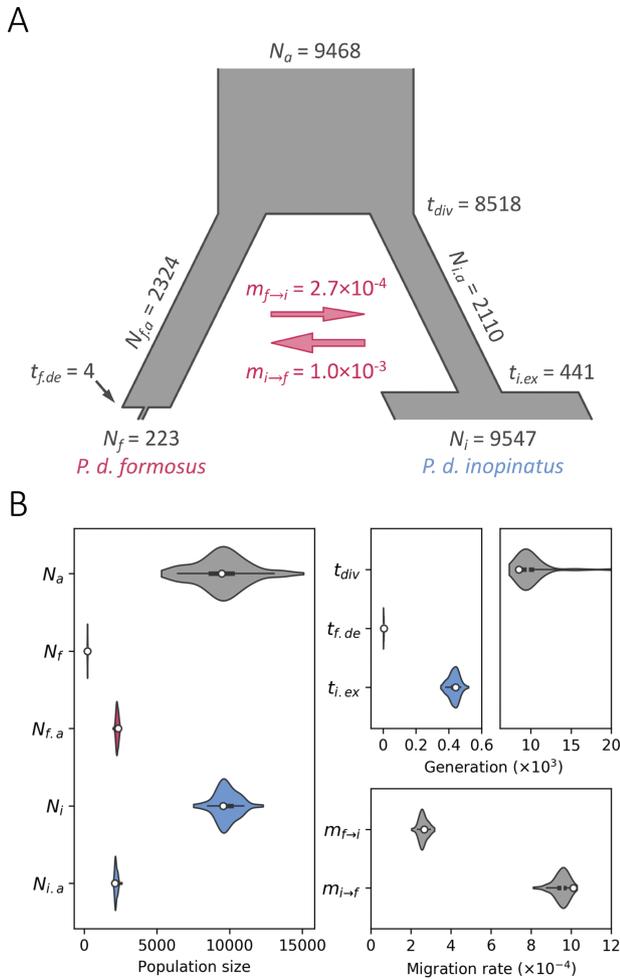
A

B

Figure 5. The best-fit demographic model and estimates of demographic parameters. (**A**) Demographic history and estimates of demographic parameters were inferred based on the observed joint fSFS of *Pteropus dasymallus formosus* and *P. d. inopinatus* by assuming the best-fit Model F (Table 3) using *fastsimcoal2*. The maximum likelihood estimates (MLEs) of population size (N), demographic event times (t with unit of generation), and per-generation migration rate (m) are shown. Note that the times of recent demographic events are not scaled to better demonstrate the changes in the population sizes of *P. d. formosus* and *P. d. inopinatus*. (B) Bootstrap distributions and quantiles for each demographic parameter. Parametric bootstraps were performed 100 times based on Model F. Their distributions and 5%, 25%, 50%, 75%, and 95% quantiles are shown. The white dot represents the MLEs of each parameter. All inferred demographic parameters are labeled as: N_a = population size of the common ancestral population of the *formosus* (*f*) and *inopinatus* (*i*); $N_{f.a}$ = the ancestral population size of the *formosus*; $N_{i.a}$ = the ancestral population size of the *inopinatus*; N_f = the current population size of the *formosus*; N_i = the current population size of the *inopinatus*; t_{div} = the time of the divergence of the 2 population; $t_{f.de}$ = the time of a recent decline in the *formosus*; $t_{i.ex}$ = the time of a recent expansion in the *inopinatus*; $m_{f \to i}$ = migration rate from the *formosus* to *inopinatus*; $m_{i \to f}$ = migration rate from the *inopinatus* to *formosus*.

Lin and Pei (1999) reported that *P. d. formosus* was once abundant on Green Island (the major habitat of *formosus*) but could be barely sighted after the 1970s. However, it is unclear whether the population on Green Island actually declined or this habitat was simply abandoned. In this study, our best-fit demographic model suggested a very recent population decline in the *formosus* population. The MLE for the time of population collapse ($t_{f.de}$) is 4 generations

Table 4. Maximum likelihood estimates of demographic parameters of *Pteropus dasymallus formosus* and *Pteropus dasymallus inopinatus* based on Model F

Parameter	MLE	95% CI
N_a	9468	(6014, 13599)
N_f	223	(201, 266)
$N_{f.a}$	2324	(2107, 2427)
N_i	9547	(8118, 11405)
$N_{i.a}$	2110	(2025, 2381)
t_{div}	8518	(7823, 16480)
$t_{f.de}$	4	(1, 5)
$t_{i.ex}$	441	(369, 493)
$m_{f \to i}$	2.7×10^{-4}	(2.2×10^{-4}, 3.0×10^{-4})
$m_{i \to f}$	1.0×10^{-3}	(8.6×10^{-4}, 1.0×10^{-3})

ago, which would be 28 years ago assuming a generation time of 7 years (Fox et al. 2008; Tidemann and Nelson 2011; Vincenot et al. 2017). The 95% CI spans from 7 years ago to 35 years ago, supporting that the event likely occurred very recently (see Figure 5B and Table 4). The magnitude of the decline was estimated to be ~10-fold, dropping from 2324 ($N_{f.a}$) to 223 (N_f) individuals. Therefore, it is more likely that the Green Island population did suffer from a severe population decline rather than habitat abandonment based on our findings. In contrast, we detected that *P. d. inopinatus* on Okinawa Island may have experienced a recent population expansion from 2,110 (N_i) to 9,547 ($N_{i.a}$) (a ~4.5-fold increase) based on the MLEs of the observed fSFS. The time of this population expansion was estimated to be 441 generations ago (~3 kya).

Our estimates also suggest that these 2 populations diverged around 8,518 generations ago (~60 kya) with a 95% CI spanning from ~55 kya to ~115 kya in the late Pleistocene (Figure 5 and Table 4). During that time, the widening Kerama Gap between Okinawa Island and Yaeyama Islands likely served as a geographic barrier to prevent gene flow between these 2 populations (Kizaki and Oshiro 1977; Ota 1998). Our results are compatible with previous studies of the divergence of amphibians, reptiles, and the elegant scops owl (*Otus elegans*) between the Central and Southern Ryukyus (Ota 1998 2000; Hsu 2005), supporting that deep-sea channels, such as the Kerama Gap, serve as effective geographic barriers to drive the genetic differentiation of island species, even for species with high migration ability like flying foxes. After the divergence of *P. d. formosus* and *P. d. inopinatus*, both populations suffered from population declines and had been maintained at relatively small population sizes for a long period of time, which might have led to their present low levels of nucleotide diversity.

It should be noted that our inferences of demographic model and parameter estimates are based on the joint fSFS of a relatively small number of samples (n) in both populations ($2n = 6$ and 8 for *P. d. formosus* and *P. d. inopinatus*, respectively). Although Beichman et al. (2018) noted that recent changes in population size are relatively difficult to capture by SFS-based methods with small sample sizes, theoretical studies have shown that the configuration of SFS is sensitive to subtle changes in population genetic parameters (e.g., effective population size, N_e) and the power of detecting such changes in SFS depends more on the number of SNPs rather than on the sample size (Akashi 1999). In the study by Akashi (1999), the number of haplotypes used in the simulations can be as small as 5 for successfully detecting changes in SFS with 2500 nucleotide sites. By analyzing ~30 000 SNPs, Robinson et al. (2014) also showed that a small sample size (as low as 3 individuals) is sufficient to detect the

correct model of recent population decline and to accurately estimate the decline time.

Although the small sample sizes in our study may have potentially reduced the statistical power to distinguish between the competing demographic models (causing false negative outcomes), we found that the best model appears to fit significantly better than the remaining models (Table 3). This suggests that the effects of the underlying demographic events (in both populations) on the observed SFS were strong and, therefore, allowed us to determine the best model. Finally, the validity of our inferred model and MLEs could also suffer from biased samplings due to small sample sizes. While our samples in both populations were collected from different locations and time points (Figure 1), future studies will focus on collecting more individuals from each of the 5 *P. dasymallus* subspecies populations.

Since the key parameters assumed during the RADSeq assembly could potentially bias population genetics statistics (Harvey et al. 2015; Mastretta-Yanes et al. 2015; Shafer et al. 2017), we also explored other combinations of assembly parameter sets (*m*, *M*, *n*) in *Stacks*, including the optimal parameters determined by the procedure of Paris et al. (2017). Subsequently, we repeated all coalescent simulations (using *fastsimcoal2*) to infer the demographic history of both populations. As a result, Model F remained the best-fit demographic scenario under the assembly parameters of *m* = 3, *M* = 2, and *n* = 3 (referred to as *m3M2n3* hereafter; Supplementary Table S5). Most inferred MLEs were also comparable between the 2 optimal parameter sets with differences in the range of 0.64–2.29-fold changes (Supplementary Figure S5), except for the time of population decline in *P. d. formosus* ($t_{f,de}$). The difference in $t_{f,de}$ was as large as 11.8-fold ($t_{f,de}$ = 4 and 47 for *m3M1n1* and *m3M2n3*, respectively; see Table 4 and Supplementary Tabe S2). In either case, we consider that the population of *P. d. formosus* experienced a recent population decline. We also tested the sensitivity of the estimates of nucleotide diversity, inbreeding coefficient, and between-populations differentiation (Φ_{ST}) under these 2 assembly parameter sets. These parameters remained robust using different assembly parameters. The detailed results are provided in Supplementary Figures S6 and S7.

Strong Genetic Structure Between *P. d. formosus* and *P. d. inopinatus*

Our results from PCA and *STRUCTURE* analyses provided evidence of genetic structure between the *formosus* and *inopinatus* populations (Figure 3). AMOVA also detected significant between-population variance (Table 3). Particularly, the results of *STRCUTURE* showed distinct patterns of genetic ancestries for individuals between these 2 populations, while little individual variation was observed within each population (Figure 3B). The evidence of genetic structure remained robust regardless of the number of ancestral populations assumed in *STRUCTURE* (Supplementary Figure S4C). In Figure 3B, it is interesting to point out that, while *STRUCTURE* assigned only one major ancestry to the individuals of *P. d. formosus*, 2 major ancestries were assigned to all 4 individuals of *P. d. inopinatus*. However, the observed pattern of mixed ancestry in the *P. d. inopinatus* individuals is not necessarily the outcome of population admixture, but may also result from shared ancestries with other populations that are closely related to *P. d. inopinatus*. Under this scenario, the proportion of mixed ancestries for a given individual genome reflects the phylogenetic distance rather than admixture fraction of populations (Lawson et al. 2018). To unravel the most likely demographic scenario resulting in the observed mixed ancestry, more samples from the other neighboring subspecies populations should be included and tested with methods specific for detecting admixtures (e.g., Patterson

et al. 2012). Nevertheless, our results provide strong evidence of genetic structure between the *formosus* and *inopinatus* populations with limited gene flow between them (Table 2). The migration rate from the *inopinatus* to *formosus* ($m_{i \to f}$ = 1.0 × 10⁻³ per generation) is about 3 times higher than that of the reverse direction ($m_{f \to i}$ = 2.7 × 10⁻⁴; see Figure 5 and Table 4). Similarly, by analyzing microsatellite and mitochondrial loci, Chen et al. (2021) also suggested that the *formosus* is genetically differentiated from the *inopinatus*. Therefore, based on our molecular evidence, *P. d. formosus* and *P. d. inopinatus* perhaps should be treated as distinct ESUs and managed separately in terms of conservation. In contrast, *P. d. formosus* individuals from Green Island and the eastern coast of Taiwan could be managed together since there is no apparent genetic structure within them.

The Yaeyama flying fox (*P. d. yayeyamae*) inhabiting Yaeyama Islands and Miyako Islands between Taiwan and Okinawa Island is the geographically closest *P. dasymallus* subspecies to *P. d. formosus*. Chen et al. (2021) observed strong population substructure between the *yayeyamae* and *inopinatus* but nonsignificant differentiation between the *yayeyamae* and *formosus*. Taki et al. (2020, unpublished data) detected genetic differentiations among the 3 *yayeyamae* populations inhabiting Yaeyama Islands: Miyako, Ishigaki-Taketomi-Kohama-Iriomote, and Yonaguni (from east to west). Among them, the Yonaguni population, which is geographically closest to Taiwan (~110 km apart), showed the highest differentiation from the other 2 *yayeyamae* populations. Therefore, it is possible that the *formosus* population is also genetically distinct from the *yayeyamae* populations inhabiting the islands of Miyako, Ishigaki, Taketomi, Kohana, and Iriomote. Future studies should focus on elucidating the levels of genetic structure between *P. d. formosus* and these *yayeyamae* populations.

In conclusion, we have revealed critical information for the conservation of the endangered Formosan flying fox and the Orii's flying fox. *Pteropus dasymallus formosus* suffers from low genetic diversity and high inbreeding pressure, likely resulting from its long-term small population history; the situation was exacerbated by a drastic and very recent population decline. Conservation acts for its population recovery are urgently needed. Habitat should be preserved and recovered on Green Island as well as the eastern coast of Taiwan. In contrast, *P. d. inopinatus* is a relatively stable population with recent growth. However, the current population bears low genetic diversity and inbreeding pressure. Long-term monitoring for the potential adverse consequences of these genetic factors should be considered. Outcrossing between *P. d. formosus* and *P. d. inopinatus* should be prevented before conducting any further studies assessing the degree of outbreeding depression between these subspecies.

Conflict of Interest

The authors declare that there is no conflict of interest.

Data Availability

Raw demultiplexed sequences are deposited in NCBI Sequence Read Archive (SRA) with the accession number of PRJNA713511. The assembled ddRADSeq loci and SNPs used in this study are available in Dryad Digital Repository at https://doi.org/10.5061/dryad.msbcc2fxh.

References

Akaike H. 1974. A new look at the statistical model identification. *IEEE Trans Autom Control*. 19:716–723.

Akashi H. 1999. Inferring the fitness effects of DNA mutations from polymorphism and divergence data: statistical power to detect directional selection under stationarity and free recombination. *Genetics*. 151:221–238.

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nat Rev Genet*. 11:697–709.

Arnold B, Corbett-Detig RB, Hartl D, Bomblies K. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Mol Ecol*. 22:3179–3190.

Aziz SA, Clements GR, McConkey KR, Sritongchuay T, Pathil S, Abu Yazid MNH, Campos-Arceiz A, Forget PM, Bumrungsri S. 2017. Pollination by the locally endangered island flying fox (*Pteropus hypomelanus*) enhances fruit production of the economically important durian (*Durio zibethinus*). *Ecol Evol*. 7:8670–8684.

Beichman AC, Huerta-Sanchez E, Lohmueller KE. 2018. Using genomic data to infer historic population dynamics of nonmodel organisms. *Annu Rev Ecol Evol Syst*. 49:433–456.

Brown VA, Brooke A, Fordyce JA, McCracken GF. 2011. Genetic analysis of populations of the threatened bat *Pteropus mariannus*. *Conservation Genetics* 12:933–941.

Cariou M, Duret L, Charlat S. 2016. How and how much does RAD-seq bias genetic diversity estimates? *BMC Evol Biol*. 16:240.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda)*. 1:171–182.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 4:7.

Cheng H-C, Changchien L-W, Lin R-S, Yang C-H, Chang S-W. 2017. *The Red List of Terrestrial Mammals of Taiwan, 2017*. Taiwan: Endemic Species Research Institute and Forestry Bureau, Council of Agriculture, Executive Yuan, Nantou.

Chen S-F, Chen S-L, Li Z-L, Lin H-C, Chang, M.-H. 2009. A survey of terrestrial vertebrates on Green Island. *Journal of National Park* 19:1–22.

Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, *et al*. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience*. 7:1–6.

Chen S-F, Juan C-H, Rossiter S, Kinjo T, Fukui D, Kawai K, Tsang SM, Veluz MJ, Sakurai H, Jang-Liaw N-H, *et al*. 2021. Population genetic structure of the insular Ryukyu flying fox, *Pteropus dasymallus*. *Biotropica*. 00:1–12.

Cox PA, Elmqvist T, Pierson ED, Rainey WE. 1991. Flying foxes as strong interactors in South Pacific island ecosystems: a conservation hypothesis. *Conserv Biol*. 5:448–454.

Ellstrand NC, Elam DR. 1993. Population genetic consequences of small population size: implications for plant conservation. *Annu Rev Ecol Syst*. 24:217–242.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 14:2611–2620.

Ewing G, Hermisson J. 2010. MSMS: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics*. 26:2064–2065.

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genet*. 9:e1003905.

Excoffier L, Foll M. 2011. fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*. 27:1332–1334.

Fox S, Luly J, Mitchell C, Maclean J, Westcott DA. 2008. Demographic indications of decline in the spectacled flying fox (*Pteropus conspicillatus*) on the Atherton Tablelands of northern Queensland. *Wildl Res*. 35:417.

Frankham R. 1996. Relationship of genetic variation to population size in wildlife. *Conserv Biol*. 10:1500–1508.

Frankham R. 1997. Do island populations have less genetic variation than mainland populations? *Heredity (Edinb)*. 78:311–327.

Frankham R. 2005. Genetics and extinction. *Biol Conserv*. 126:131–140.

Frankham R. 2008. Inbreeding and extinction: island populations. *Conserv Biol*. 12:665–675.

Frankham R. 2010. Challenges and opportunities of genetic approaches to biological conservation. *Biol Conserv*. 143:1919–1927.

Frankham R, Briscoe DA, Ballou JD. 2002. *Introduction to conservation genetics*. Cambridge: Cambridge University Press.

Frankham R, Lees K, Montgomery ME, England PR, Lowe EH, Briscoe DA. 1999. Do population size bottlenecks reduce evolutionary potential? *Anim Conserv*. 2:255–260.

Fujita MS and Tuttle MD. 1991. Flying foxes (Chiroptera: Pteropodidae): threatened animals of key ecological and economic importance. *Conserv Biol*. 5:455–463.

Funakoshi K, Osawa Y, Osawa K. 2006. Distribution of the Ryukyu flying-fox *(Pteropus dasymallus inopinatus)* on islands adjacent to Okinawa Island, with special reference to their ecology on Yoron Island. *Honyuyui Kagaku (Mammalian Science)* 46:29–34.

Funakoshi K, Osawa Y, Osawa K. 2012. First record of the Orii's flying-fox, *Pteropus dasymallus inopinatus* on Okinoerabu-jima Island in Kagoshima Prefecture, Japan, with special reference to their inhabitation. *Honyurui Kegaku (Mammalian Science)* 52:179–184.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. 2012. Harnessing genomics for delineating conservation units. *Trends Ecol Evol*. 27:489–496.

Gautier M, Gharbi K, Cezard T, Foucaud J, Kerdelhué C, Pudlo P, Cornuet JM, Estoup A. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Mol Ecol*. 22:3165–3178.

Graffelman, J. 2015. Exploring diallelic genetic markers: the HardyWeinberg package. *J Stat Soft*. 64:1–23.

Graffelman J, Camarena JM. 2008. Graphical tests for Hardy-Weinberg equilibrium based on the ternary plot. *Hum Hered*. 65:77–84.

Harvey MG, Judy CD, Seeholzer GF, Maley JM, Graves GR, Brumfield RT. 2015. Similarity thresholds used in DNA sequence assembly from short reads can reduce the comparability of population histories across species. *PeerJ*. 3:e895.

Hsu Y-C. 2005. *Molecular ecology of elegant scops owl (Otus elegans) (Doctoral dissertation)*. Taipei (Taiwan): National Taiwan University.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 23:1801–1806.

Kano T. 1929. The distribution and behavior of Taiwan mammals (I). *Dobutsugaku Zasshi (Zoological Magazine)* 40:332–340.

Kinjo K, Nakamoto A. 2015. *Pteropus dasymallus*. In: Ohdachi SD, Ishibashi Y, Iwasa MA, Fukui D, and Saitoh T, editors. *The wild mammals of Japan*. 2nd ed. Kyoto, Japan: Shoukadoh Book Sellers. p. 52–53.

Kizaki K, Oshiro I. 1977. Paleogeography of the Ryukyu Islands. *Marine Science Monthly* 9:542–549.

Kullback S, Leibler RA. 1951. On information and sufficiency. *Ann Math Stat*. 22:79–86.

Kuroda N. 1933. A revision of the genus Pteropus found in the islands of the Riu Kiu Chain, Japan. *American Society of Mammalogists* 14:312–316.

Lawson DJ, van Dorp L, Falush D. 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Commun*. 9:3258.

Lin L-K, Pei J-C. 1999. On the current status of field population of Formosan fruit bat (*Pteropus dasymallus formosus*). *Endemic Species Research* 1:12–19.

Lynch M, Walsh B. 1998. *Genetics and analysis of quantitative traits*. Sunderland (MA): Sinauer.

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol Ecol Resour*. 15:28–41.

McConkey KR, Drake DR. 2007. Indirect evidence that flying foxes track food resources among islands in a Pacific Archipelago. *Biotropica* 39:436–440.

Mickleburgh SP, Hutson AM, Racey PA. 1992. *Old World fruit bats. An action plan for their conservation*. Gland, Switzerland: IUCN, 263.

Mickleburgh S, Waylen K, and Racey P. 2009. Bats as bushmeat: A global review. *Oryx* 43:217–234.

Nachman MW, Crowell SL. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics*. 156:297–304.

Nakamoto A, Itabe S, Sato A, Kinjo K, Izawa M. 2011. Geographical distribution pattern and interisland movements of Orii's flying fox in Okinawa Islands, the Ryukyu Archipelago, Japan. *Popul Ecol*. 53:241–252.

Ota H. 1998. Geographic patterns of endemism and speciation in amphibians and reptiles of the Ryukyu Archipelago, Japan, with special reference to their paleogeographical implications. *Res Popul Ecol*. 40:189–204.

Ota H. 2000. The current geographic faunal pattern of reptiles and amphibians of the Ryukyu Archipelago and adjacent regions. *Tropics*. 10:51–62.

Paradis E. 2010. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*. 26:419–420.

Paris JR, Stevens JR, Catchen JM. 2017. Lost in parameter space: a road map for stacks. *Methods Ecol Evol*. 8:1360–1373.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics*. 192:1065–1093.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 7:e37135.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155:945–959.

Purcell S, Chang C. 2020. PLINK 1.9. Available from: www.cog-genomics.org/plink/1.9/

QGIS Development Team 2019. *QGIS Geographic Information System*. Open Source Geospatial Foundation Project. Available from: http://qgis.org

Ray DA, Feschotte C, Pagan HJ, Smith JD, Pritham EJ, Arensburger P, Atkinson PW, Craig NL. 2008. Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Res*. 18:717–728.

Raymond M, Rousset F. 1995. GENEPOP (version 1.2) : population genetics software for exact tests and ecumenicism. *J Hered*. 86, 248–249.

Roberts BJ, Catterall CP, Eby P, Kanowski J. 2012. Long-distance and frequent movements of the flying-fox *Pteropus poliocephalus*: implications for management. *PLoS One*. 7:e42532.

Robinson JD, Coffman AJ, Hickerson MJ, Gutenkunst RN. 2014. Sampling strategies for frequency spectrum-based population genomic inference. *BMC Evol Biol*. 14:254.

Rousset F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour*. 8:103–106.

Shafer ABA, Peart CR, Tusso S, Maayan I, Brelsford A, Wheat CW, Wolf JBW. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods Ecol Evol*. 8:907–917.

Sovic MG, Carstens BC, Gibbs, HL. 2016. Genetic diversity in migratory bats: Results from RADseq data for three tree bat species at an Ohio windfarm. *PeerJ*. 2016:1–17.

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

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

Taki Y, Vincenot CE, Sato Y, Inoue-Murayama M. 2020. Genetic diversity and population structure in the Yaeyama flying fox. Manuscipt submitted for publication. *bioRxiv* doi:10.1101/2020.03.23.004622

Tidemann CR and Nelson JE. 2011. Life expectancy, causes of death and movements of the grey-headed flying-fox (*Pteropus poliocephalus*) inferred from banding. *Acta Chiropterologica* 13:419–429.

Vincenot CE. 2017. *Pteropus dasymallus*. The IUCN Red List of Threatened Species 2017: e.T18722A22080614. Available from: https://dx.doi.org/10.2305/IUCN.UK.2017-2.RLTS.T18722A22080614.en.

Vincenot CE, Florens FB, Kingston T. 2017. Can we protect island flying foxes? *Science*. 355:1368–1370.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol*. 7:256–276.

Webb NJ, Tidemann CR. 1996. Mobility of Australian flying-foxes, Pteropus spp. (Megachiroptera): evidence from genetic variation. *Proc Biol Sci*. 263:497–502.

Wu H-W. 2010. *Preliminary ecological study on the Formosan flying foxes (Pteropus dasymallus formosus) (Master's thesis)*. Taipei (Taiwan): National Taiwan University.

Yoshiyuki M. 1989. A systematic study of the Japanese Chiroptera. *Natl Sci Mus Monogr*. 7:1–242.