東海大學生命科學系博士論文
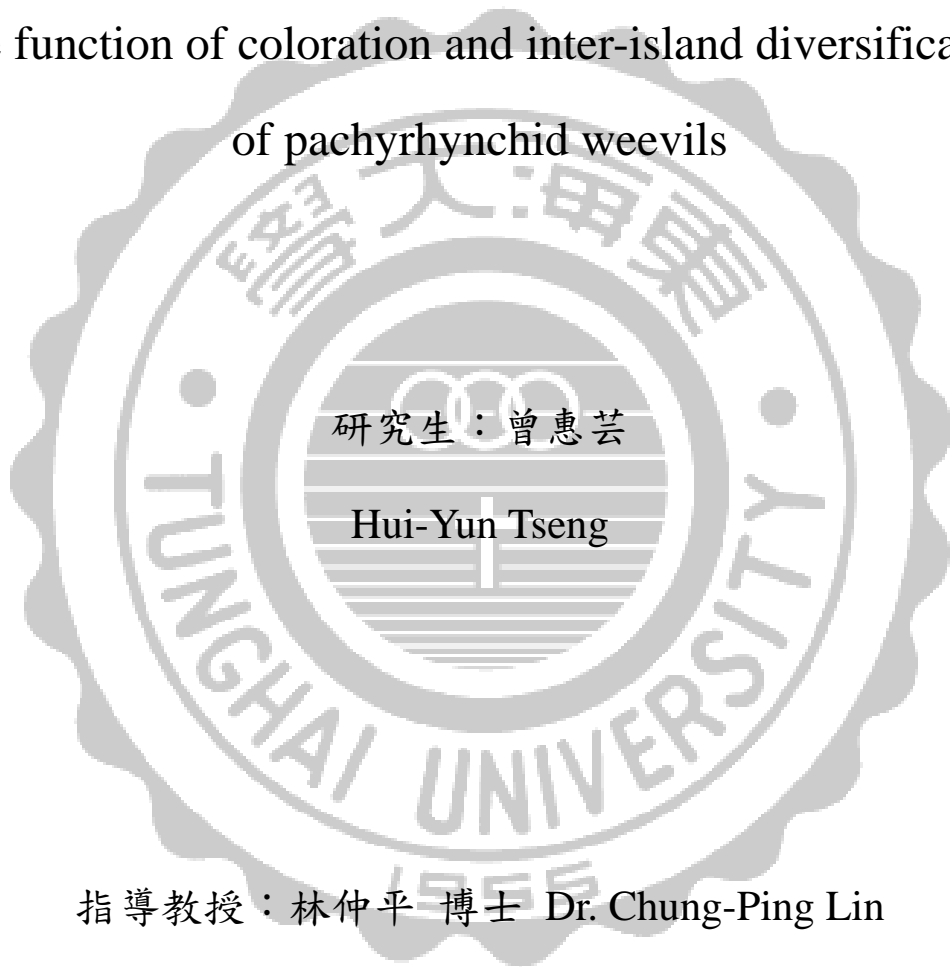
球背象鼻蟲斑紋功能與島嶼間演化歷史

The function of coloration and inter-island diversification

of pachyrhynchid weevils

研究生：曾惠芸

Hui-Yun Tseng

指導教授：林仲平 博士 Dr. Chung-Ping Lin

黃文山 博士 Dr. Wen-San Huang

劉少倫 博士 Dr. Shao-Lun Liu

中華民國 一百零六年一月

January, 2017

東海大學生命科學系
博士論文學位考試審定書

生命科學系博士班研究生<u>曾惠芸</u>君所撰寫之論文

（中文）

球背象鼻蟲斑紋功能與島嶼間演化歷史

（英文）

The function of coloration and inter-island diversification of pachyrhynchid weevils

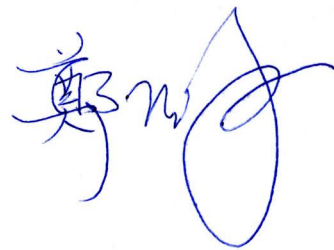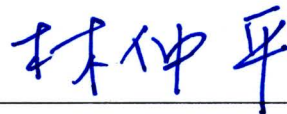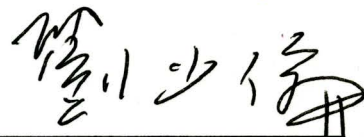經本委員會審定通過，特此證明。

學位考試委員會　　召集人　　　　　　　　（簽名）

　　　　　　　　　委　員

中華民國　　　年　　　月　　　日

謝誌

一切要從五年半前的某一天開始說起…

和林仲平老師在東海校園內的丹堤討論起是否要念博班與可能的研究方向，老師說起他在菲律賓捉角蟬時，常常會看到球背象鼻蟲，問我對這種蟲的了解，當時我只知道蘭嶼有這種蟲，但是完全不了解，只問說”是那種爬一爬之後會飛走的那種嗎???”到現在，我還是不知道老師當時的想法。因為球背象鼻蟲很大的特色就是牠們”不會飛”，而當時的我對球背象鼻蟲的認識比許多喜愛昆蟲的人要少很多很多。好心的老師在這種強況下還是收了我當學生，而且一路走來，不論是我賴皮不想做的事、請求幫忙的事、煩惱的事，老師總是盡力的協助，從來沒有聽到一句責備的話。仲平老師對研究的要求、對人的溫暖，讓我堅信一位研究做得極好的學者也可以同時是給予人力量的人。另一位指導老師黃文山老師同時也是我工作上的上司，從我決定念博班開始，便開始給我無限的支持。在我們對球背象鼻蟲完全不了解的時候，陪著我一起在蘭嶼、綠島、菲律賓走上一整天尋找球背象鼻蟲。對一位兩爬學者而言，收了我一位做昆蟲的學生，還一併支持自己的其他學生參與球背象鼻蟲的研究，讓我們對球背的了解在這短短幾年內有了快速的進展。在工作上能有這樣一位全力支持研究的老闆我想是所有人夢寐以求的。劉少倫老師在仲平老師離開東海後接手了我這個學生，在系上 seminar 的報告中，總是給我最多的建議與鼓勵。對我在學校、系上的事，老師總是沒有第二句話就直接幫忙。東海大學卓逸民老師也是從我進入東海開始，從第一年 seminar 開始總是給予寶貴的意見與鼓勵；中研院蔡怡陞老師對未來在球背象鼻蟲基因體研究上給了很多珍貴的建議，讓我對未來的研究方向有了更多的信心；科博館鄭明倫博士從研究一開始就給了極多的建議與鼓勵，總是在第一時間協助解答我的疑惑，從不吝於分享他的想法與知識。

野外工作團隊

科博館兩爬學門的黃主任、雄哥與嫆雅在球背象鼻蟲研究的一開始就幫忙蘭嶼與

綠島的野外工作，球背象鼻蟲雖然不會飛，但是也沒有輕鬆容易的捕捉方法，只能靠著雙眼尋找，雄哥展顯了極佳的眼力與體力，幫忙我開始了蘭嶼與綠島的研究；熔雅在工作與研究上不停分擔我的工作量，並且極細心的提點許多注意的事項，讓我的研究與工作都能順利進行。一起參與菲律賓野外工作的夥伴，寶島賣螢團的明倫學長在多次菲律賓的野外工作中，總是把球背列為重點工作項目之一，出人出錢出力，帶著我這麼肉腳的夥伴也從來沒有嫌棄過，總是把最好的地方讓給我睡覺，在我捉不到球背非常沮喪的夜晚，熬夜和我分享自己的研究歷程；團長陳燦榮在野外調查途中展現了過人的野外能力並且細心的照顧到每一人，從白天到晚上辛苦的野外工作完後，睡前即使自己已經有拍不完的螢火蟲，還要在我呼呼大睡的時候熬夜幫我拍下美麗的球背照片；張志遠老師在野外的工作中生火幫大家在寒冷的雨後取暖、煮熱熱的東西給大家喝。有這樣的團隊一起在野外工作讓人非常放心與開心。在菲律賓常幫我上廁所把風的黃主任、一邊點燈和我分享易經的宇堂、犧牲自己得登革熱的立偉、喜愛爬樹捉球背的高如柏醫師、一邊展現高超的抓蝴蝶技巧一邊和我分享蟲界歷史的徐堉峰老師、身手矯捷帶我一起找海蛇的杜銘章老師、手受傷不能動時照顧我的品辰，每個人對我的幫忙與照顧讓我在菲律賓的日子留下了滿滿快樂的回憶，也從大家身上學到很多生活的態度。菲律賓的野外研究中最重要的就是 Hilario Cahilog 了，幫我們打點了生活中大大小小的事、還要上山下海幫忙採集，Reagan Joseph T. Villanueva 在採集行程與安排上的建議讓我們的菲律賓野外工作能更順利。沒機會一起出野外但是幫忙菲律賓採集的仲平老師、傅風鳥、黃嘉龍、蔡經甫學長，謝謝大家的幫忙才讓我能安逸地待在台灣還有研究上的收獲與進展。

研究分析

若凡總是無私地跟我說明 DNA 分析的細節、貼心的提醒我實驗室相關的事情；現在人在日本的書書總是幫忙解決我分析遇到的障礙、並在第一時間給我分析上的建議；草魚實驗室的嘉偉在論文最後的關頭幫忙火燒屁股的我解決所有

Contents

Abstract

Pachyrhynchid weevil (Coleoptera: Curculionidae) is a group of insects with astonishing and diverse coloration, but their biology and diversification are poorly known. This dissertation studied the function of these diverse colors and how these weevils and their colors have evolved on the oceanic islands where they inhabited. Chapter 1 aims to test whether these conspicuous colorations function as aposematic signals for the arboreal predator, Swinhoe's tree lizard (*Japalura swinhonis*). *Pachyrhynchus tobafolius* and *Eupyrgops waltonianus* were selected as potential prey, and the predatory responses of allopatric versus sympatric lizard populations were compared by presenting weevils with intact or artificially-removed markings. Sympatric lizards were found to attack masked weevils more frequently than intact ones, whereas allopatric lizards attacked intact weevils with a significantly higher rate due to the novelty of the preys. The results supported the hypothesis that the conspicuous markings of pachyrhynchid weeils function as aposematic signals for lizards. Chapter 2 tested the stepping-stone hypothesis of colonization of the *Pachyrhynchus orbifer* complex in the Taiwan-Luzon volcanic belt using phylogenies reconstructed from a multi-locus data-set. Divergence time estimates of all extant island lineages (<1 Mys of Middle Pleistocene) are much more recent than the geological ages of the islands (22.4-1.7 Mys). Complex and stochastic colonization of *P. orbifer* was inferred to have involved both northward and southward dispersals, which are strongly inconsistent with the strict stepping-stone hypothesis. Speciation of the *P. orbifer* complex may have occurred only between islands, suggesting that peripatric speciation through founders of stochastic dispersals was the major evolutionary driver. Chapter 3 evaluated the effect of filtering parameters in the process of assembling RAD-sequences. The remarkable advances in next-generation sequencing technology provide a powerful tool

for phylogenetics and population genetics, especially in recently divergent populations and shallow phylogenetic relationships, which were problematic using a limited number of loci. However, potential errors from different sources of RAD-seq may cause incorrect inference in downstream analyses. RAD sequences of *P. orbifer* species complex with different filtering parameters combinations in *Populations* of *Stacks* program were used to evaluate the stability of inferred phylogenies by comparing the number of retained loci, number of SNPs, heterozygosity, percentage of polymorphic sites, bootstrap supports and tree topologies. These comparisons showed that even processing the loci in the population level caused different levels of variation in the phylogenetic inferences, and setting higher values of *r*, *p* and *m* might not be always suitable. We recommend choosing filtering parameters by testing different combination of parameters.

中文摘要

球背象鼻蟲（鞘翅目：象鼻蟲科）是一群體色鮮豔且斑紋變異極大的昆蟲，但相關的生物學與物種分化的研究至今仍然非常不足。本論文研究這些多樣性極高的斑紋的功能，並探討球背象鼻蟲與他們身上的斑紋如何在海洋島嶼上演化而來。本論文的第一個研究即為探討球背象鼻蟲身上的斑紋對掠食者而言是否具有警戒色的功能。以兩種球背象鼻蟲（小圓斑球背象鼻蟲與碎斑硬象鼻蟲）為實驗對象，斯文豪氏攀蜥為掠食者，比較和球背象鼻蟲共域與異域的掠食者對有無斑紋的象鼻蟲的捕食反應。結果顯示共域的掠食者對斑紋被移除的球背象鼻蟲有較高的捕食率，而異域的掠食者則對有斑紋的象鼻蟲有較高的捕食率。這些結果驗證了球背象鼻蟲的斑紋對蜥蜴的警戒色功能。第二個研究則利用多基因探討臺灣與呂宋火山島弧間球斑球背象鼻蟲複合群（*Pachyrhynchus orbifer* complex）的演化歷史，檢測象鼻蟲是否在島嶼間發生踏石擴散的現象。球斑球背象鼻蟲在島嶼形成（22.4−1.7 百萬年）後更晚近的時間才在島嶼間分化，分化時間大約是更新世中期之後（一百萬年以內）。球斑球背象鼻蟲在島嶼間的擴散歷史非常複雜，包括了南向與北向的擴散，也因而反駁了踏石擴散的假說。島嶼間隨機擴散造成的邊域種化可能是球斑球背象鼻蟲在島弧上種化的重要機制。第三個章節則探討限制酶位點標定定序後不同篩選參數的影響。次世代定序技術提供強有力的研究工具進行親緣關係與族群遺傳研究，特別是近期分化或用有限基因無法解決的親緣關係議題。然而，限制酶位點標定定序過程中有許多產生誤差的可能來源，而造成後續分析的錯誤推論。本研究利用球斑球背象鼻蟲的限制酶位點標定序列在 *Stacks* 軟體中的 *Populations* 程式進行不同篩選參數的測試，以保留基因座數量、單核苷酸多態型數量、異型結合性、多型性點位比例、統計支持度與樹型來比較不同的篩選參數的結果，結果顯示不同的參數組合會造成的不同程度變異的親緣關係推論，越高的 r、p 與 m 在後續分析上並不一定是最適合的參數。因此我們建議在進行後續分析之前要先測試不同的參數組合，以選取最佳的參數。

Chapter 1. Biology and questions of pachyrhynchid weevils

Pachyrhynchid weevils (Insecta: Coleoptera: Curculionidae) are perhaps one of the most colorful and charismatic groups of insular beetles found throughout the Old World tropics. Even Alfred Russel Wallace was impressed by these colorful insects and once said that, "In these islands also we find the extensive and wonderful genus of weevils (*Pachyrhynchus*), which in their brilliant metallic colouring surpass anything found in the whole eastern hemisphere, if not in the whole world" (p. 385-386, Wallace 1891). The astonishing colorations of these weevils (Fig. 1), along with their peculiar island distribution, have raised great interests of both amateur collectors and biologists. The coloration of these weevils is formed by the reflection of light through the scales covering their body surfaces. The underlying structural reflector of the scale was identified as three-dimensional photonic crystals (Welch et al. 2007, Seago et al. 2009).

The geographical range of pachyrhynchid weevils is restricted and only extends from Southern Ryukyus (Iriomote and Ishigaki Island) to New Guinea, with highest species diversity found in the Philippine archipelago (Schultze 1923). Presumably because of their limited dispersal ability (wingless), pachyrhynchid weevils exhibit a high level of endemism throughout their distribution (Schultze 1923). Restricted geographic distribution and island endemism make pachyrhynchid weevils a prime candidate for studying the ecological and evolutionary mechanisms underlying the biogeographic pattern of insular biota (Starr and Wang 1992). With already more than 100 species of *Pachyrhynchus* being described (Schultze 1923, Yoshitake 2012, Barševskis 2016, Rukmane and Barševskis 2016), the species diversity of these weevils may still be underestimated, particularly in the remote Batanes and Babuyan Islands situated between Luzon and Taiwan and other little explored area such as the

northeastern mountain ranges of Luzon, Mindoro Island, most of Visayan islands and central Mindanao (Schultze 1923).

The completely fused elytra is an unique insular adaptation of pachyrhynchid weevils and likely forms a protective shield against predators (Schultze 1923). Because of their extremely rigid bodies, remarkably similar colors and general body shape between different genera of pachyrhynchid weevils and a few mimic insects groups (e.g. longhorn beetles, crickets) (Fig. 2), these colorful markings were considered as aposematic/warming signals for their predators (Schultze 1923, Welch et al. 2007, Barševskis 2013, 2016). However, at present no scientific evidence supports the above-mentioned aposematic hypothesis.

Chapter 2 used Swinhoe's tree lizard (*Japalura swinhonis*), a potential predator, and two weevil species, *P. tobafolius* and *Eupyrgops waltonianus* (synonym *Kashotonus multipunctatus*), to examine the aposematic hypotheses. The specific questions are: 1) Do these conspicuous colorations function as aposematic signals for the lizard predator? 2) Do sympatric lizards avoid attacking this colorful prey? and 3) Do allopatric, naive lizards attack more on these presumably unpalatable prey than sympatric ones?

In the northern peripheral range of Pachurhynchid weevils, only one species, *P. infernalis*, is distributed in Ishigaki and Iriomote Islands of southern Ryukyus. In Orchid and Green Islands, five (*P. insularis*, *P. sonani*, *P. sarcitis*, *P. tobafolius* and *P. yamianus*) and four (*P. sonani*, *P. sarcitis*, *P. tobafolius* and *P. yamianus*) species of *Pachyrhynchus*, respectively, were recorded (Starr and Wang 1992). Only three (*P. chlorites*, *P. viridans* and *P. orbifer*) (Schultze 1923, Kano 1936) and two species (*P.*

sarcitis and *P. orbifer*) (Schultze 1923) were recorded in Babuyan Islands and Batanes Islands, respectively. However, recent collecting trips to these islands by the author and colleagues revealed that there are more new *Pachyrhynchus* species to be discovered (Fig. 3).

Because of the fused elytra, pachyrhynchid weevils can draw and keep air under their elytra allowing afloat on the surface of water for several hours (Schultze 1923), and the eggs and larvae live and grow in the fruits or woods which may drift or float cross the ocean (Schultze 1923, Kayashima 1940, Oshiro 1991). Although no explicit scientific research ever conducted to evaluate the floating capacity of these weevils, the floating ability was presumed to increase the probability of successful overwater dispersal of adult weevils.

Taiwan-Luzon volcanic chain forms an island chain stretching from northern Luzon to Green Island. In this region, the Kuroshio Current flows from south to north along these islands. A well-sampled *Pachyrhynchus* weevils across island chain in this region will provide a good study system to test stepping-stone dispersal and Kuroshio current-mediated dispersal hypotheses. Chapter 3 reconstructed the phylogeny of the weevils from Yaeyama Islands, Green Island, Orchid Island, Batanes Islands, Babuyan Islands and northern Luzon to examine the following specific questions: 1) Whether the historical colonization pattern of *Pachyrhynchus* weevils follows a northward stepping-stone dispersal along Taiwan-Luzon volcanic chain? and 2) Are the divergence time of species from southern islands older than that of species from northern islands in sequence? 3) Whether species from southern islands exhibit a historical demography of stable or declining population size (ancient colonization),

whereas species from northern islands experienced a demographical expansion following a population bottleneck (recent colonization)?

Recent advances of using the genome-wide single-nucleotide polymorphism (SNP) to infer colonization events and gene flow among closely related species (Emerson et al. 2010, Catchen et al. 2013, Wagner et al. 2013, Benestan et al. 2015), such as the restriction site-associated DNA sequencing (RAD-seq), increase the precision of detecting-historic population structure and recent colonization history. However, errors are prone to occur in large amount of SNPs and questions arise as sequence filtering process can influence the inferences of downstream analyses. In chapter 4, RAD sequences of *P. orbifer* species complex with different filtering parameter combinations in *Populations* of *Stacks* program were used to evaluate the stability of phylogenetic relationships, by comparing number of retained loci, number of SNPs, heterozygosity, percentage of polymorphic site, bootstrap supports and tree topology, and to find the most suitable data set for future analyses.
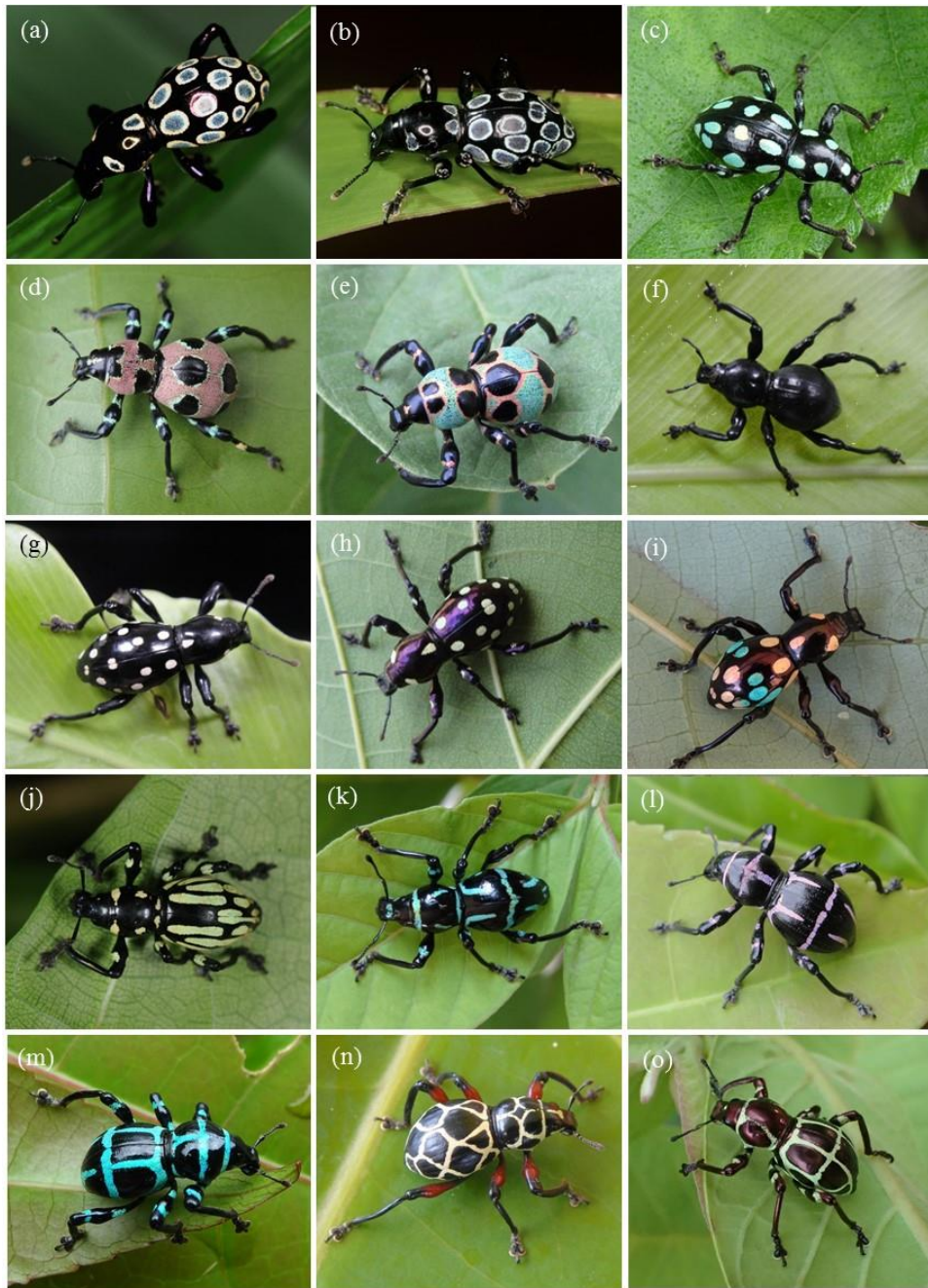
Figure 1. Diverse colouration of *Pachyrhynchus* weevils. (a) *Pachyrhynchus* sp. 1, Luzon; (b) *Pachyrhynchus* sp. 2, Luzon; (c) *P. tobafolius*, Orchid Island; (d) *P. orbifer*, Luzon; (e) *P. orbifer*, Luzon; (f) *P. infernalis*, Iriomote Island; (g) *P. insularis*, Orchid Island; (h) *P. chlorites*, Calayan Island; (i) *Pachyrhynchus* sp. 3, Luzon; (j) *Pachyrhynchus* sp. 4, Luzon; (k) *P. semperi*, Babuyan Island; (l) *P. orbifer*, Camiguin Island; (m) *P. orbifer*, Babuyan Island; (n) *P. reticulatus*, Luzon; (o) *P. gloriosus*, Luzon. (a), (b), and (j) were photographed by T.-L. Chen.
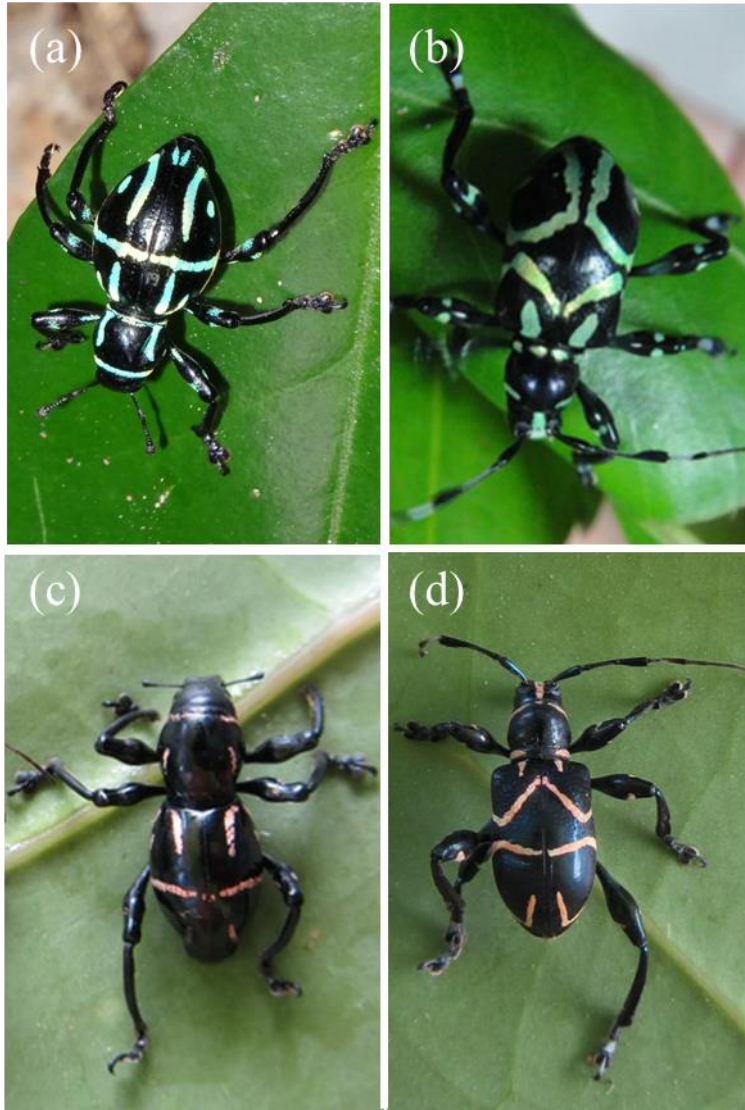
Figure 2. Mimic pairs between pachyrhynchid weevils (left) and longhorn beetles (right). (a) *Pachyrhynchus yamianus* and (b) *Doliops similis* from Orchid Island; and (c) *Pachyrhynchus* sp. and (d) *Doliops* sp. from Camiguin Island.
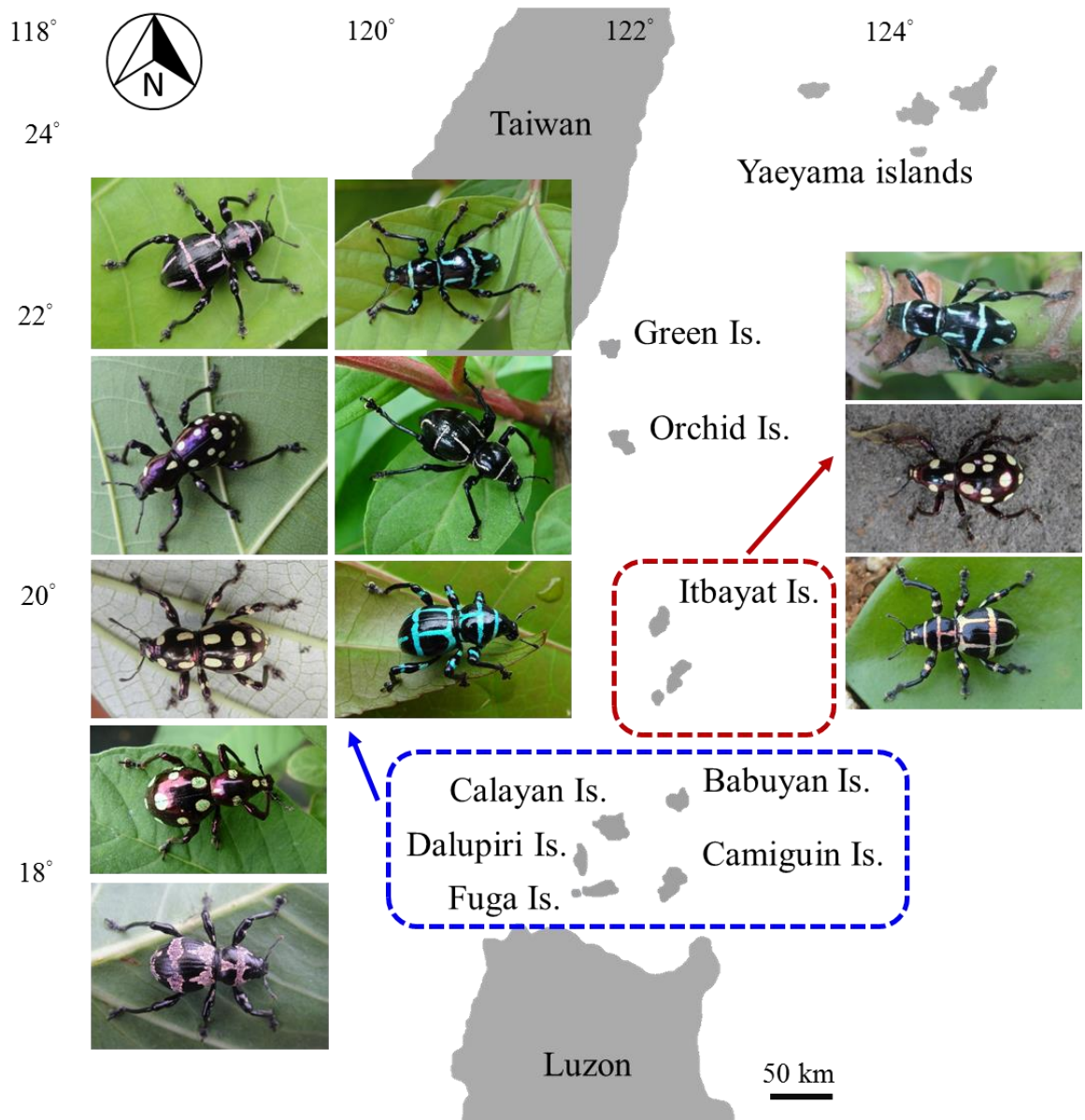
Figure 3. Candidate new *Pachyrhynchus* species discovered in Batanes and Babuyan Islands.

References

Barševskis, A. 2013. Contribution to the knowledge of the genus Doliops Waterhouse, 1841 (Coleoptera: Cerambycidae). Baltic Journal of Coleopterology **13**:73-89.

Barševskis, A. 2016. New species of *Pachyrhynchus* Germar, 1824 and *Macrocyrtus* Heller, 1912 (Coleoptera: Curculionidae) from the Marinduque Island (Philippines) as a new example of mimetic species pair. Acta Biologica Universitatis Daugavpiliensis **16**:1-6.

Benestan, L., T. Gosselin, C. Perrier, B. Sainte-Marie, R. Rochette, and L. Bernatchez. 2015. RAD‑genotyping reveals fine‑scale genetic structuring and provides powerful population assignment in a widely distributed marine species; the American lobster (*Homarus americanus*). Molecular Ecology **24**:3299-3315.

Catchen, J., S. Bassham, T. Wilson, M. Currey, C. O'Brien, Q. Yeates, and W. A. Cresko. 2013. The population structure and recent colonization history of Oregon threespine stickleback determined using restriction‑site associated DNA‑sequencing. Molecular Ecology **22**:2864-2883.

Emerson, K. J., C. R. Merz, J. M. Catchen, P. A. Hohenlohe, W. A. Cresko, W. E. Bradshaw, and C. M. Holzapfel. 2010. Resolving postglacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences **107**:16196-16200.

Faria, C., A. Machado, I. R. Amorim, M. J. Gage, P. A. Borges, and B. C. Emerson. 2016. Evidence for multiple founding lineages and genetic admixture in the evolution of species within an oceanic island weevil (Coleoptera, Curculionidae) super-radiation. Journal of Biogeography **43**:178-191.

Kano, T. 1936. Some problems concerning the biogeography of Kôthshô. Geographical review of Japan **12**:1107-1133.

Kayashima, I. 1940. Notes on a Weevil (*Pachyrrhynchus moniliferus* Germ.) injurious

to Cacao Fruit in the Philippine Islands. Transactions of the Natural History Society of Formosa **30**:126-129.

Oshiro, Y. 1991. Studies on the *Pachyrhynchus infernalis* effecting *Mangifera indica* in Ishigaki Island. Okinawa Agriculture **26**:19-23.

Rukmane, A., and A. Barševskis. 2016. Nine new species of the genus *Pachyrhynchus* Germar, 1824 (Coleoptera: Curculionidae) from the Philippines. Baltic Journal of Coleopterology **16**:77-96.

Schultze, W. 1923. A monograph of the pachyrrhynchid group of the Brachyderinae, Curculionidae: part 1. The Philippine Journal of Science **23**:609-673.

Seago, A., P. Brady, J. Vigneron, and T. Schultz. 2009. Gold bugs and beyond: a review of iridescence and structural colour mechanisms in beetles (Coleoptera). Journal of the Royal Society Interface **6**:S165-S184.

Sequeira, A., A. Lanteri, L. R. Albelo, S. Bhattacharya, and M. Sijapati. 2008. Colonization history, ecological shifts and diversification in the evolution of endemic Galápagos weevils. Molecular Ecology **17**:1089-1107.

Starr, C. K., and H. Y. Wang. 1992. Pachyrhynchine weevils (Coleoptera: Curculionidae) of the islands fringing Taiwan. Journal of Taiwan Museum **45**:5-14.

Tänzler, R., E. F. Toussaint, Y. R. Suhardjono, M. Balke, and A. Riedel. 2014. Multiple transgressions of Wallace's Line explain diversity of flightless Trigonopterus weevils on Bali. Proceedings of the Royal Society of London B: Biological Sciences **281**:20132528.

Wagner, C. E., I. Keller, S. Wittwer, O. M. Selz, S. Mwaiko, L. Greuter, A. Sivasundar, and O. Seehausen. 2013. Genome‑wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular Ecology **22**:787-798.

Wallace, A. R. 1891. Natural selection and tropical nature: Essays on descriptive and

theoretical biology, Macmillan, London.

Welch, V., V. Lousse, O. Deparis, A. Parker, and J. P. Vigneron. 2007. Orange reflection from a three-dimensional photonic crystal in the scales of the weevil *Pachyrrhynchus congestus pavonius* (Curculionidae). Physical Review E **75**:041919.

Yoshitake, H. 2012. Nine new species of the genus Pachyrhynchus Germar (Coleoptera: Curculionidae) from the Philippines. Esakia **52**:17-34.

Chapter 2. The functional significance of aposematic signals on pachyrhynchid weevils

Abstract

Conspicuous colouration can evolve as a primary defence mechanism that discourages predator attacks by advertising unprofitability. Old World pachyrhynchid weevils display an incredible diversity of bright patterns, which Alfred Russel Wallace hypothesized served as aposematic signals. We experimentally tested this hypothesis using two weevil species (*Pachyrhynchus tobafolius* and *Eupyrgops waltonianus*) and a widespread visual predator, Swinhoe's tree lizard (Agamidae, *Japalura swinhonis*). We compared the predatory responses of allopatric and sympatric lizard populations to weevils with intact patterns or those with experimentally-masked patterns. Sympatric lizards attacked weevils without conspicuous patterns at higher rates than weevils with intact conspicuous patterns, whereas allopatric lizards attacked weevils with intact patterns at higher rates. The conspicuous coloration of pachrhynchid weevils is thus attractive to naïve predators, but also can serve as aposematic signals for sympatric lizards. Allopatric lizards ignored (rather than attacked) weevils after a single encounter, which persisted for more than three weeks. The diversity of pachyrhinchid weevils on Old World tropical islands is likely shaped by complex predator-prey interactions because of high prey species turnover (and thus complexity in aposematic signalling) across small geographic distances. Widespread predators can vary their responses to aposematic prey depending upon encounter rates, which can be influenced by geographical and habitat overlap, local abundance, and prior experience. Understanding how predators interact with a diverse range of aposematic prey is fundamental towards a fuller understanding of local trophic dynamics, community composition, and species turnover.

14

Introduction

Many distasteful or toxic organisms have conspicuous colour patterns, which can act as a primary defence mechanism by advertising unprofitability (Smith 1977). Conspicuous colours or patterns can function as aposematic signals, which are easily detected or memorized by predators (Prudic et al. 2007, Kuchta et al. 2008). Individual fitness can increase substantially when predators are able to easily recognize, and subsequently avoid, unprofitable prey displaying obvious visual signals (Sillén-Tullberg 1985, Kuchta et al. 2008). Responsive predators also save time and energy through the early detection of unpalatable prey, which increases prey survival (Sillén-Tullberg 1985, Kuchta et al. 2008) and decreases wasted predation attempts by predators (Ruxton et al. 2004).

Conspicuous colours and/or patterns that act as true aposematic signals can be learned and avoided by vertebrate and invertebrate predators, both in the laboratory (Langham 2004; Ham et al. 2006; Exnerová et al. 2007; Prudic et al. 2007; Svádová et al. 2009) and in the field (Kapan 2001, Borer et al. 2010). Elements of the conspicuous colouration itself can enhance the cognitive ability of predators in response to unpalatable prey (Terrick et al. 1995, Prudic et al. 2007). For example, high chromatic contrast and brightness of aposematic prey not only increase the predator learning speed, but also memory retention of aversive responses (Gittleman and Harvey 1980, Prudic et al. 2007) . Predators, such as the mantis, can retain aversive responses longer for conspicuous unpalatable prey than for cryptic prey (Prudic et al. 2007). Memory retention may be variable across predator taxa, however, and relevant studies of the duration of aposematic prey recognition are scarce.

Pachyrhynchid weevils (Insecta: Coleoptera: Curculionidae) show astonishing

diversity in colouration, and their thoraces, elytra and legs are often decorated with brightly coloured stripes, circles, and/or spots against high-contrasting dark bodies (Fig. 1). However, the colour structure and iridescence mechanisms on these weevils are not well understood (Welch et al. 2007, Seago et al. 2009). Many sympatric insects, comprised of different weevil genera (Schultze 1923), other Coleopterans (e.g., longhorn beetles) (Wallace 1889, Linsley 1959, Starr and Wang 1992), and even Orthopterans (Wallace 1889) mimic the colouration and pattern of some pachyrhynchid weevils. To understand whether these cues are aposematic, we need to understand whether the weevils are in fact the model that other taxa are mimicking. For this to be the case, weevils should be unprofitable prey. Alfred Russel Wallace first hypothesised that the conspicuous colours of pachyrhynchid weevils served as warning advertisements to predators (Wallace 1889, Welch et al. 2007). Weevils could be unpalatable for some predators because of their extremely tough exoskeleton (chemical defences are unknown in this group of insects(Schultze 1923, Pasteels et al. 1983)). After more than 120 years, however, we still do not understand the adaptive significance of bright colouration in pachyrhynchid weevils, despite their conspicuous nature and high diversity throughout the Old World tropics.

We experimentally investigated the biological function and adaptive significance of conspicuous colouration in pachyrhynchid weevils. First, we tested the aposematic function of bright patterns by analysing responses of predators to weevils with intact or experimentally masked conspicuous colouration. Next, we tested for geographic variation in predatory responses to conspicuously coloured weevils by comparing the frequency of predatory attacks between allopatric and sympatric predator populations. We predicted that allopatric predators would show higher predation rates upon weevils because of lack of prior experience with these prey. Finally, we studied the ability of

allopatric predators to learn to avoid weevils after their first encounter, and the duration

of any avoidance response.

Materials and methods

Study species

The weevil *Pachyrhynchus tobafolius* (Fig. 1a) is distributed on Green (22°39'33"N, 121°29'15E) and Orchid Islands (22° 3'18"N, 121°32'41"E), located 30 and 60 kilometres from southeastern Taiwan, respectively. The weevil *Eupyrgops waltonianus* (Fig. 1b) is endemic to Green Island. Neither of these species occurs on Taiwan. These two species have black bodies decorated with metallic green (*P. tobafolius*) or blue (*E. waltonianus*) or spots on dorsal surfaces of the head, thorax, elytra and legs (Fig. 1). Swinhoe's tree lizards (*Japalura swinhonis*, Agamidae) are widespread, semi-arboreal predators of weevils that are often observed in weevil host plants alongside weevils (Huang 2004, Huang 2006a, b, 2007, 2010, 2011).

Study populations

We captured lizards using a noose or by hand from populations on Green and Orchid Islands, Jinshan in northern Taiwan (25°13'18.74"N, 121°38'10.55"E), and Kenting in southern Taiwan (22°1'5.64"N, 120°44'42.36"E; June 2011 - October 2012). Each lizard was assigned a unique identification number and housed individually in a plastic container (34×17×24 cm). Water was available *ad libitum* and mealworms (*Tenebrio molitor*) were provided every three days. To help ensure that lizards were hungry at the time of testing, we did not feed them 24 hours prior to trials.

Manipulating bright colour signals

We randomly assigned individual *P. tobafolius* and *E. waltonianus* into two groups: 1) a control group, comprised of weevils with intact colour markings, and 2) an experimental group, in which we masked all bright weevil colouration using a black marker pen (No. 3102003A, Simbalion). We also applied the mask to the black thoraces

and elytra of control weevils (Fig. 1a, b). To confirm that our mask was biologically meaningful, we measured the reflectance spectra of different components of intact weevil patterns and the background colour of the weevil's body (Fig. 2).

To confirm that our black mask was biologically meaningful, we measured the reflectance spectra of different parts of intact weevil patterns and the background colour of the weevil's body, along with the same physical locations in weevils with the black mask applied (black pen, No. 3102003A, Simbalion). We used a Jaz spectrometer (detection range: 250-800 nm, Ocean Optics, Inc., Dunedin, FL) connected to a reflection probe (ZFQ-13101, Ocean Optics, Inc.), which was lit by a deuterium–tungsten halogen light source (DH-2000-BAL, Ocean Optics, Inc.). For each weevil measurement ($N = 2$ per species/treatment combination), the probe was placed vertically 2 mm above the focal part of the cuticle and adjusted to obtain the maximal reflectance spectra. We calibrated our machine using two reflectance standards (Ocean Optics STAN-SSH-NIST for high reflectance measurements, or WS-1-SL for low reflectance measurements). These calibrations provide a standard basis against which measurements can be compared. STAN-SSH-NIST is a mirror reflectance standard, and is specific to high reflectance measurements, whereas WS-1-SL is used for standard reflectance readings. When using the STAN-SSH-NIST calibration, the readings for the black mask and the weevil background were ~0%, and thus this calibration standard was not very meaningful. When using the standard calibration (WS-1-SL), the bright patterns on the weevils exceeded the 100% standard measurement, which indicates that these are very reflective markings. To provide a more useful comparison, we present data from the WS-1-SL calibration. Data were processed using SpectraSuite software (Ocean Optics, Inc.).

Experimental design

Our experimental design focuses on two weevil species with intact or experimentally-masked patterns, and lizards from three lizard populations: (1) Taiwan lizards are allopatric to both weevil species, and no other pachyrhynchid weevils are found there; (2) Green Island lizards are sympatric with the weevils *P. tobafolius* and *E. waltonianus*; and (3) Orchid Island lizards are sympatric with the weevil *P. tobafolius* and allopatric to the weevil *E. waltonianus*. We predicted that allopatric predators would show higher predation rates on weevils overall, because of the lack of prior experience with these prey. By contrast, we predicted that sympatric predators would be more likely to consume masked weevils than those with bright colouration.

Behavioural trials

For each trial, we placed a lizard into an arena (39×30×20 cm) for one minute prior to introducing a weevil, which was tied to a black cotton thread and positioned approximately 10 cm in front of the lizard. Trials were terminated after two minutes because most lizards attacked prey within this time. The behavioural response of each lizard was recorded as "attack" or "ignore." Attack behaviour was defined as a lizard approaching and biting the weevil, which was either consumed or spat out. Ignore behaviour was defined as a lizard not attacking the weevil during a two-minute trail. Immediately after each weevil trail, we tested whether lizards were hungry by offering each a mealworm tied to a cotton thread. If the lizard ate the mealworm, we classified it as having ignored the weevil as palatable prey. We excluded the individuals from analysis which lizards neither attacked the weevil nor consumed the mealworm, because they may not have been hungry during testing. We used chi-squared tests to determine whether individual lizard populations differed significantly in their behavioural responses towards intact and masked weevils, and whether responses were similar between sympatric (Green and Orchid Islands) and allopatric (Taiwan) predator

populations. In these trials, each lizard was used only once, and thus all data are independent.

Learning and memory retention

We used the lizard populations from Taiwan to test whether allopatric lizards learn to avoid *E. waltonianus* weevils after their first encounter, and whether any avoidance behaviour is retained over time. Allopatric lizards from Taiwan that successfully attacked a weevil were randomly assigned to one of four groups, which were presented with another weevil after 1, 5, 13, or 23 days. A contingency table analysis was used to test whether the frequency of predatory attacks differed with the number of days passing before being offered another weevil.

Results

We collected 156, 338, 178 and 126 lizards from Orchid Island, Green Island, southern Taiwan, and northern Taiwan (n = 798 lizards).

Predatory responses towards intact and masked weevils

For the lizards from Orchid Island, only 26.0% (20/77) attacked intact *P. tobafolius*, whereas a significantly higher percentage of the lizards 48.0% (36/75) attacked masked weevils ($\chi^2 = 7.92$, df = 1, $P < 0.01$) (Fig. 3a). However, sympatric lizards from Green Island exhibited similar attack rates towards intact (51.8%, 43/83) and masked (52.5%, 42/80) *P. tobafolius* ($\chi^2 = 0.01$, df = 1, $P = 0.92$) (Fig. 3a). The sympatric lizards from Green island had significantly higher attack rates towards masked (57.1%, 44/77), rather than intact (36.1%, 30/83) *E. waltonianus* ($\chi^2 = 7.08$, df = 1, $P < 0.01$), which was similar to the response displayed by sympatric lizards from Orchid Island (Fig. 3a).

Predatory responses of sympatric and allopatric lizards

The allopatric lizards from southern (n = 87) and northern (n = 66) Taiwan and the sympatric lizards from Green Island (n = 83) showed significantly higher attack rates towards *P. tobafolius* than did sympatric lizards from Orchid Island (n = 77) (southern Taiwan vs. Orchid Island, $\chi^2 = 58.29$, df = 1, $P < 0.01$; northern Taiwan vs. Orchid Island, $\chi^2 = 37.41$, df = 1, $P < 0.01$; Green Island vs. Orchid Island, $\chi^2 = 11.17$, df = 1, $P < 0.01$) (Fig. 3b). We found similar predatory responses in the two allopatric lizards from southern (n = 82) and northern Taiwan (n = 52). These allopatric populations also had significantly higher attack rates upon *E. waltonianus* than did sympatric lizards from Green Island (n = 83) (southern Taiwan vs. Green Island, $\chi^2 = 37.43$, df = 1, $P < 0.01$; northern Taiwan vs. Green Island, $\chi^2 = 14.01$, df = 1, $P < 0.01$) (Fig. 3c). Overall, the vast majority of lizards that attacked the weevils after initially biting them (97.1%)

spat them out, with only a few lizards chewing up, crushing, or consuming weevils (2.9%, 13/444 lizards).

Memory retention of prey avoidance

After initially attacking a weevil, allopatric lizards strongly avoided consuming a subsequent weevil for up to 23 days (Fig. 4). We found no significant difference among our treatments in the propensity of individual lizards to re-attack weevils after different durations ($\chi^2 = 1.22$, df = 3, $P = 0.748$). After a single day, only 14.3% of lizards were willing to attack a weevil, and this strong avoidance behaviour was maintained for at least 23 days after initially attacking a weevil (6.7% attack rate; Fig. 4).

Discussion

Our study provides evidence that the conspicuous colouration of pachyrhynchid weevils can function as aposematic signals that reduce attack rates by widespread vertebrate predators. The colour markings of *P. tobafolius* and *E. waltonianus* greatly decreased attack rates of sympatric Swinhoe's tree lizards from Orchid and Green Islands, respectively. This is consistent with the responses of other predators towards conspicuously-coloured organisms (Hasegawa and Taniguchi 1994, Dolenská et al. 2009, Chouteau and Angers 2011, Veselý et al. 2013). Surprisingly, however, the sympatric lizards from Green Island attacked both intact and masked *P. tobafolius* to similar degrees. Prey recognition of a predator can be shaped by the relative frequency with which aposematic prey species are encountered (Pfennig et al. 2001, Hunter 2009). On Green Island, *P. tobafolius* is less abundant than *E. waltonianus*, likely because of low host plant abundance (Starr and Wang 1992). When prey abundance decreases, predators such as these lizards may be less likely to have prior experience with some aposematic patterns, and consequently may mistakenly attack these prey.

Overall, the attack rates of sympatric lizards were significantly higher upon masked weevils than attack rates upon intact weevils, but these rates were lower than those of allopatric Taiwanese lizards attacking intact weevils (Fig. 3). These results suggest that lizards can recognise the unprofitability of masked weevils using cues other than bright colouration. Colour is one component of an aposematic signal, which can include complex patterns (Kikuchi and Pfennig 2010, Aronsson and Gamberale-Stille 2012), shapes (Dolenská et al. 2009, Hunter 2009, Veselý et al. 2013), sizes (Pekár and Jarab 2011), and even odours (Siddall and Marples 2008). In our study, almost half of all sympatric lizards avoided attacking masked weevils (42.9 – 52.0%, depending on the population), suggesting that physical shape, size, behaviour, or scent can provide

24

additional aposematic cues.

The results of our study provide the first demonstration of geographic variation in prey recognition by a squamate reptile. Similar responses have been reported in other taxa, especially avian predators (Noonan & Comeault, 2009; Chouteau & Angers, 2011). Geographic variation of prey recognition may result from different local prey communities, which can impose diverse selection pressures upon predators even in adjacent geographic regions (Chouteau & Angers, 2011). Allopatric lizards were much more likely to attack aposematic weevils, probably because they had no prior experience with the patterns from these prey species. This is a common phenomena in many predator-prey systems (Noonan and Comeault 2009, Chouteau and Angers 2011). However, in other instances, predators may show avoidance to novel aposematic prey because of neophobia or dietary conservatism (Marples et al. 1998; Exnerová et el. 2007).

A substantial proportion of the sympatric lizards we tested consumed their respective local weevil species, suggesting that learning plays a role in the ability of individual predators to recognize aposematic prey. Incomplete avoidance behaviour by predators represents an ongoing learning/continued testing process (Noonan and Comeault 2009), poor learning/forgetting (Lynn 2005), or predator naivety (Saporito et al. 2007). Many organisms, especially birds, can learn to avoid to unprofitable prey based on specific, general signals (Gamberale-Stille and Tullberg 1999, Ham et al. 2006, Svádová et al. 2009). However, studies focused on avoidance learning in reptiles remain rare (Terrick et al. 1995). Our preliminary results of memory retention suggest that Swinhoe's tree lizards can learn to avoid weevils immediately after an encounter, and retain their avoidance behaviour for up to 23 days. In many studies of avian predators,

the experimental duration of memory retention tests is less than 7 days after the first treatment (Gamberale-Stille and Tullberg 1999, Ham et al. 2006). Even in memory tests of garter snakes (*Thamnophis radix*) that continued for 22 days, attack latencies towards aposematic prey decreased over time (Terrick et al. 1995). By contrast, Swinhoe's tree lizards maintained high and stable rates of continued prey avoidance behaviour (Fig. 4). Because the ability to recognize these unpalatable prey did not decline by the end of our 23-day trial, the aposematic signals of these weevils not only influence predatory responses, but also form strong and long-lasting memory associations in predators.

We provide the first empirical support for the long-standing hypothesis that the colour markings of pachyrhynchid weevils function as aposematic signals. Animals displaying aposematic signals often emit or possess chemical compounds, such as alkaloids (Holloway et al. 1991), cardenolides (Reichstein et al. 1968), or formic acid (Kanehisa and Murase 1977) to aid in predator deterrence. Earlier studies have been unable to find chemical defences or secretory organs in pachyrhynchid weevils (Schultze 1923, Pasteels et al. 1983). The primary defensive mechanism of these weevils is likely to be their tough exoskeleton (Wallace 1889, Welch et al. 2007). Of the hundreds of lizards tested in our study, only a few individuals were willing to consume these prey once they attacked; instead, most lizards spat the weevils out almost immediately. The ability of predators to recognise these prey depends upon encounter rates, which can be influenced by geographic and habitat overlap, local abundance, and prior experience (Endler and Rojas 2009).

The conspicuous colouration of pachyrhynchid weevils can serve as warning signals, and thus these insects provide a geographically and phylogenetically complex group that can be used to expand insights into predator-prey relationships, as well as

whether these weevils potentially are the models in their mimetic relationships with other insects in the region. Understanding how predators interact with the diverse range of aposematic models and mimics found throughout the Old World tropics is fundamental towards a comprehensive understanding of the ecological basis for local trophic dynamics, community composition, and species turnover.

Figure 1 Pachyrhynchid weevils used in our experiments. (a) *Pachyrhynchus tobafolius*; (b) *Eupyrgops waltonianus*; (c) and (d) show *P. tobafolius* and *E. waltonianus*, respectively, with their aposematic markings experimentally masked with a black marker pen.
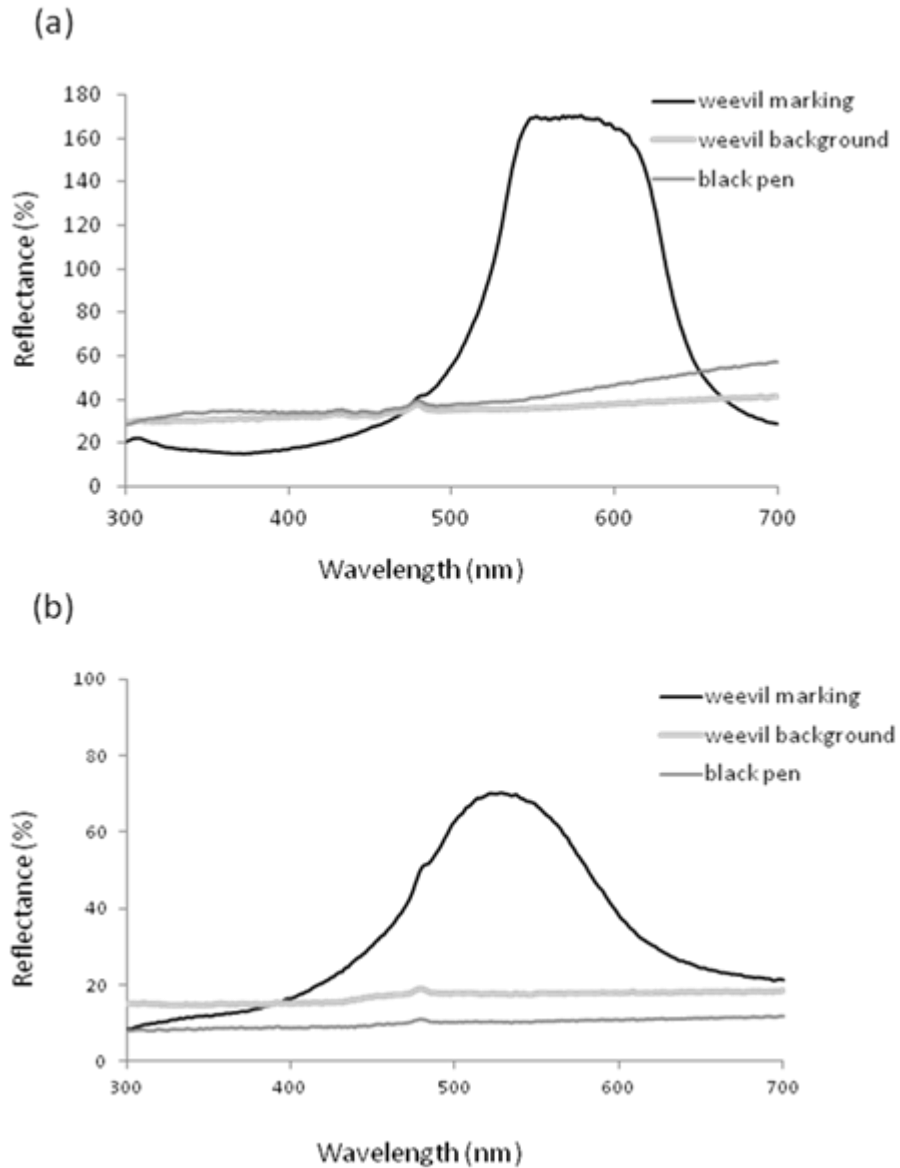
Figure 2 Reflectance spectra of the dark background colour of weevils, the bright patterning of weevils, and the black marker used to mask the colourful patterning for (a) *Pachyrhynchus tobafolius* and (b) *Eupyrgops waltonianus*. Note that the black marker more closely matches the background colouration of both weevil species than the bright patterns that we masked.

29

(a) *P. tobafolius*   *E. waltonianus*

sympatric   sympatric   sympatric

(b) *P. tobafolius*

(c) *E. waltonianus*
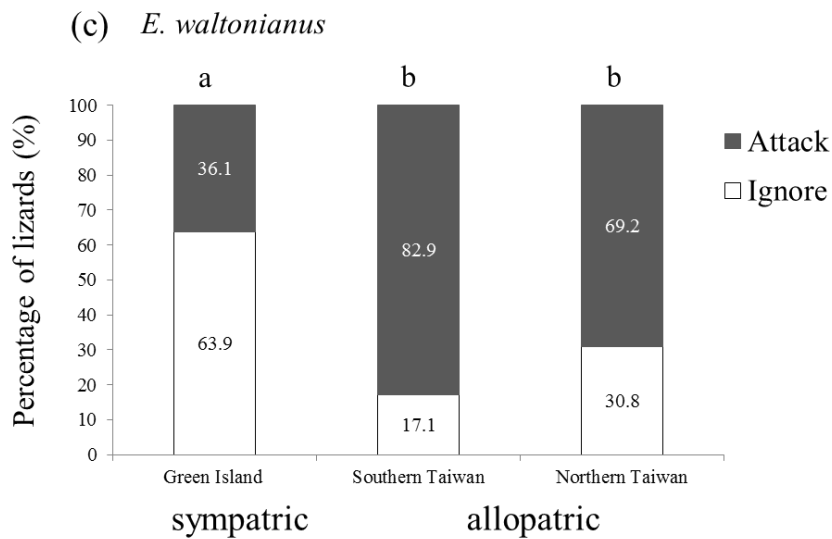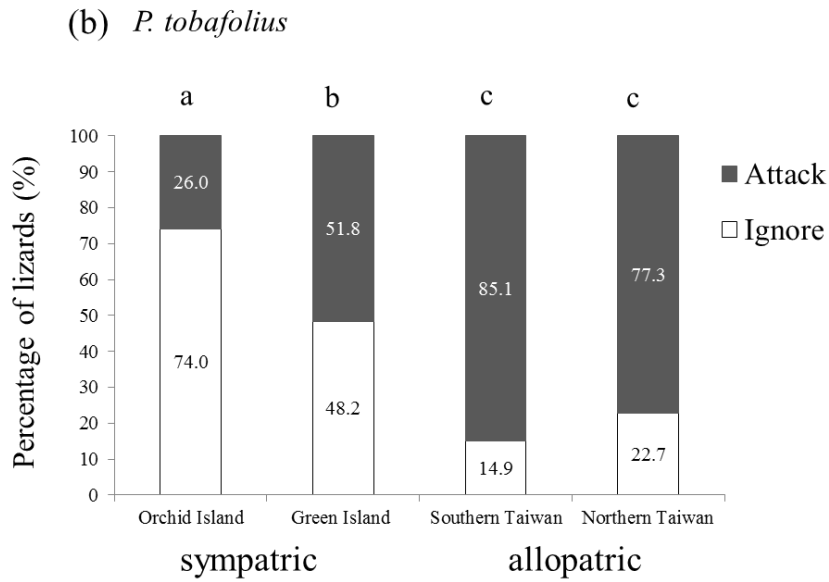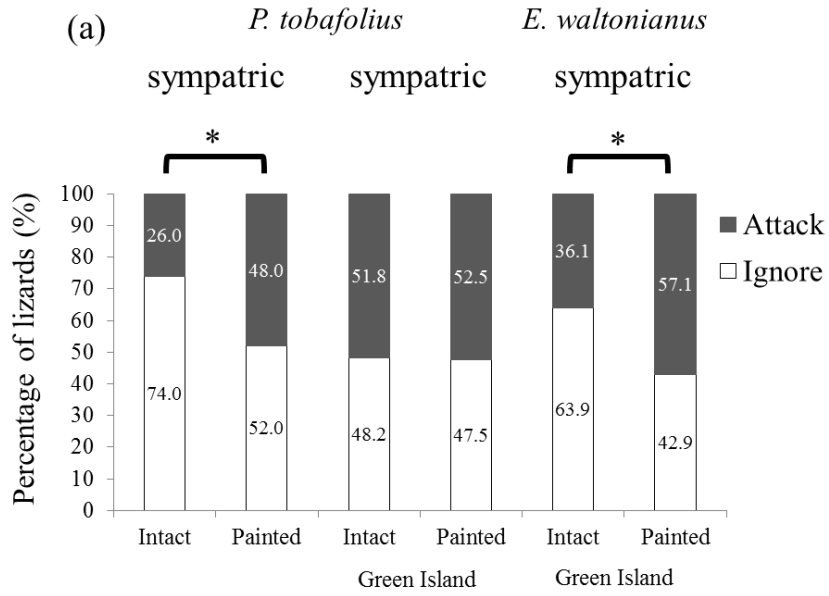
Figure 3 Proportion of lizards from different populations exhibiting different predatory behaviour. (a) The first four bars represent the response of sympatric lizards from Orchid Island and Green Island to *P. tobafolius*. The rightmost bars show the response of sympatric lizards from Green Islands to *E. waltonianus*; (b) Response of lizards from different localities to *P. tobafolius* with intact markings; (c) Response of lizards from different localities to *E. waltonianus* with intact markings. Different letters indicate statistically significant differences.
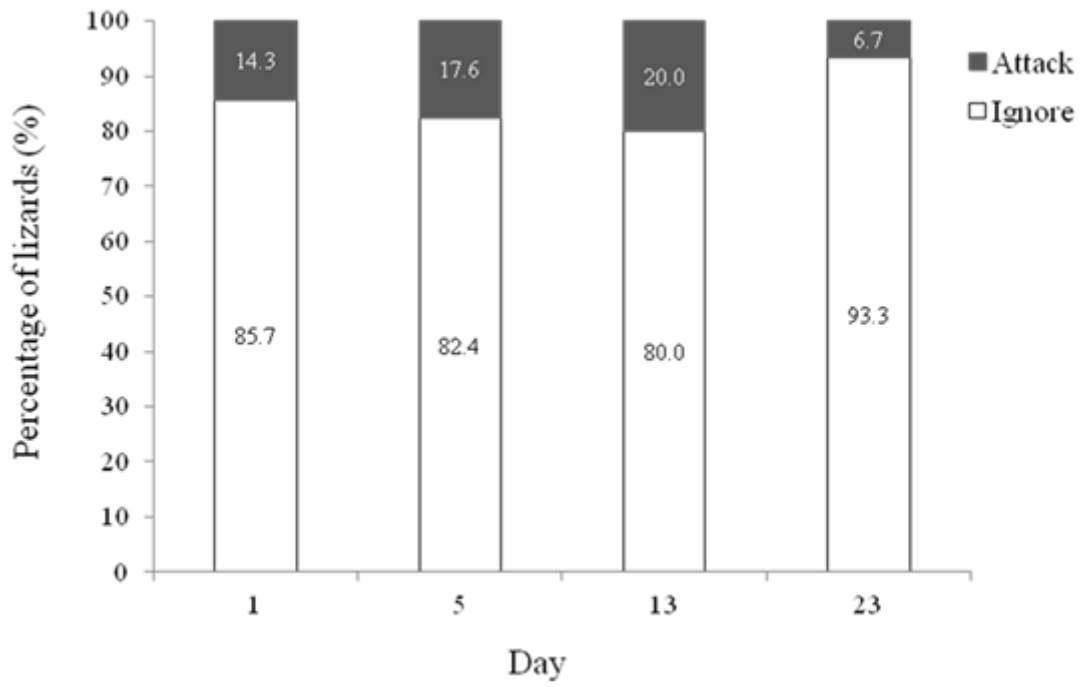
Figure 4 Memory retention tests of the predatory responses of allopatric lizards to colorful weevil prey (*Eupyrgops waltonianus*).

References

Aronsson, M., and G. Gamberale-Stille. 2012. Colour and pattern similarity in mimicry: evidence for a hierarchical discriminative learning of different components. Animal Behaviour **84**:881-887.

Borer, M., T. Van Noort, M. Rahier, and R. E. Naisbit. 2010. Positive frequency-dependent selection on warning color in Alpine leaf beetles. Evolution **64**:3629-3633.

Chouteau, M., and B. Angers. 2011. The role of predators in maintaining the geographic organization of aposematic signals. The American naturalist **178**:810-817.

Dolenská, M., O. Nedved, P. Veselý, M. Tesarová, and R. Fuchs. 2009. What constitutes optical warning signals of ladybirds (Coleoptera: Coccinellidae) towards bird predators: colour, pattern or general look? Biological Journal of the Linnean Society **98**:234-242.

Gamberale-Stille, G., and B. S. Tullberg. 1999. Experienced chicks show biased avoidance of stronger signals: an experiment with natural colour variation in live aposematic prey. Evolutionary Ecology **13**:579-589.

Gittleman, J. L., and P. H. Harvey. 1980. Why are distasteful prey not cryptic? Nature **286**:149-150.

Ham, A., E. Ihalainen, L. Lindström, and J. Mappes. 2006. Does colour matter? The importance of colour in avoidance learning, memorability and generalisation. Behavioral Ecology and Sociobiology **60**:482-491.

Hasegawa, M., and Y. Taniguchi. 1994. Visual avoidance of a conspicuously colored carabid beetle *Dischissus mirandus* by the Lizard *Eumeces okadae*. Journal of Ethology **12**:9-14.

Holloway, G. J., P. W. de Jong, P. M. Brakefield, and H. de Vos. 1991. Chemical defence in ladybird beetles (Coccinellidae). I. Distribution of coccinelline and individual

variation in defence in 7-spot ladybirds (*Coccinella septempunctata*). Chemoecology **2**:7-14.

Huang, W.-S. 2004. Reptile ecology and the evolution of parental care on a tropical Asian island. Cornell University, Ithaca, New York.

Huang, W.-S. 2006a. Ecological characteristics of the skink, *Mabuya longicaudata*, on a tropical East Asian island. Copeia **2006**:293-300.

Huang, W.-S. 2006b. Ecology and reproductive patterns of the grass lizard, *Takydromus sauteri*, in a tropical rain forest of an East Asian island. Journal of Herpetology **40**:267-273.

Huang, W.-S. 2007. Ecology and reproductive patterns of the agamid lizard *Japalura swinhonis* on an east Asian island, with comments on the small clutch sizes of island lizards. Zoological Science **24**:181-188.

Huang, W.-S. 2010. Ecology and reproductive characteristics of the skink *Sphenomorphus incognitus* on an East Asian island, with comments on variations in clutch size with reproductive modes in *Sphenomorphus*. Zoological Studies **49**:779-788.

Huang, W.-S. 2011. Ecology and reproductive patterns of the skink, *Emoia atrocostata*, on an East Asian tropical rainforest island. Zoological Studies **50**:506-512.

Hunter, J. S. 2009. Familiarity breeds contempt: effects of striped skunk color, shape, and abundance on wild carnivore behavior. Behavioral Ecology **20**:1315-1322.

Kanehisa, K., and M. Murase. 1977. Comparative study of the pygidial defensive systems of carabid beetles. Applied entomology and zoology **12**:225-235.

Kapan, D. D. 2001. Three-butterfly system provides a field test of Mullerian mimicry. Nature **409**:338-340.

Kikuchi, D. W., and D. W. Pfennig. 2010. Predator cognition permits imperfect coral snake mimicry. The American naturalist **176**:830-834.

Kuchta, S. R., A. H. Krakauer, and B. Sinervo. 2008. Why does the yellow-eyed *Ensatina* have yellow eyes? Batesian mimicry of pacific newts (genus *Taricha*) by the salamander *Ensatina eschscholtzii xanthoptica*. Evolution **62**:984-990.

Linsley, E. G. 1959. Ecology of Cerambycidae. Annual Review of Entomology **4**:99-138.

Lynn, S. K. 2005. Learning to avoid aposematic prey. Animal Behaviour **70**:1221-1226.

Noonan, B. P., and A. A. Comeault. 2009. The role of predator selection on polymorphic aposematic poison frogs. Biology Letters **5**:51-54.

Pasteels, J. M., J. Gregoire, and M. Rowell-Rahier. 1983. The chemical ecology of defense in arthropods. Annual Review of Entomology **28**:263-289.

Pekár, S., and M. Jarab. 2011. Assessment of color and behavioral resemblance to models by inaccurate myrmecomorphic spiders (Araneae). Invertebrate Biology **130**:83-90.

Pfennig, D. W., W. R. Harcombe, and K. S. Pfennig. 2001. Frequency-dependent Batesian mimicry. Nature **410**:323.

Prudic, K. L., A. K. Skemp, and D. R. Papaj. 2007. Aposematic coloration, luminance contrast, and the benefits of conspicuousness. Behavioral Ecology **18**:41-46.

Reichstein, T., J. v. Euw, J. A. Parsons, and M. Rothschild. 1968. Heart poisons in the monarch butterfly. Science **161**:861-866.

Ruxton, G., T. Sherratt, and M. Speed. 2004. Avoiding attack: the evolutionary ecology of crypsis, aposematism, and mimicry. Oxford University Press, Oxford, New York.

Saporito, R. A., R. Zuercher, M. Roberts, K. G. Gerow, and M. A. Donnelly. 2007. Experimental evidence for aposematism in the dendrobatid poison frog *Oophaga pumilio*. Copeia **2007**:1006-1011.

Schultze, W. 1923. A monograph of the pachyrrhynchid group of the Brachyderinae,

Curculionidae: part 1. The Philippine Journal of Science **23**:609-673.

Seago, A., P. Brady, J. Vigneron, and T. Schultz. 2009. Gold bugs and beyond: a review of iridescence and structural colour mechanisms in beetles (Coleoptera). Journal of the Royal Society Interface **6**:S165-S184.

Siddall, E. C., and N. M. Marples. 2008. Better to be bimodal: the interaction of color and odor on learning and memory. Behavioral Ecology **19**:425-432.

Sillén-Tullberg, B. 1985. Higher survival of an aposematic than of a cryptic form of a distasteful bug. Oecologia **67**:411-415.

Smith, S. M. 1977. Coral-snake pattern recognition and stimulus generalisation by naive great kiskadees (Aves: Tyrannidae). Nature **265**:535-536.

Starr, C. K., and H. Y. Wang. 1992. Pachyrhynchine weevils (Coleoptera: Curculionidae) of the islands fringing Taiwan. Journal of Taiwan Museum **45**:5-14.

Svádová, K., A. Exnerová, P. Stys, E. Landová, J. Valenta, A. Fucíková, and R. Socha. 2009. Role of different colours of aposematic insects in learning, memory and generalization of naïve bird predators. Animal Behaviour **77**:327-336.

Terrick, T. D., R. L. Mumme, and G. M. Burghardt. 1995. Aposematic coloration enhances chemosensory recognition of noxious prey in the garter snake *Thamnophis radix*. Animal Behaviour **49**:857-866.

Veselý, P., D. Luhanová, M. Prášková, and R. Fuchs. 2013. Generalization of mimics imperfect in colour patterns: the point of view of wild avian predators. Ethology **119**:138-145.

Wallace, A. R. 1889. Darwinism: an exposition of the theory of natural selection with some of its applications, MacMillan, London.

Welch, V., V. Lousse, O. Deparis, A. Parker, and J. P. Vigneron. 2007. Orange reflection from a three-dimensional photonic crystal in the scales of the weevil *Pachyrrhynchus congestus pavonius* (Curculionidae). Physical Review E

75:041919.

Chapter 3. Complex inter-island colonization and peripatric founder speciation promote diversification of flightless weevils (*Pachyrhynchus orbifer* complex) in the Taiwan-Luzon volcanic belt

Abstract

Oceanic islands are excellent models for investigating colonization history and speciation processes. Diversification of species from the Taiwan-Luzon volcanic belt at the Asian-Philippine faunal boundary might have originated through a northward "stepping-stone model" of asymmetric colonization. Our study specifically tested this stepping-stone hypothesis in a group of colorful and flightless weevils - the *Pachyrhynchus orbifer* complex - using their colonization histories reconstructed from a multi-locus data-set. The temporal order of weevil origins did not follow the linear sequence of spatial arrangement of the islands. Divergence time estimates of all extant island lineages (<1 Mys of Middle Pleistocene) are much more recent than the geological ages of the islands (22.4-1.7 Mys). Complex and stochastic colonization of *P. orbifer* was inferred to have involved both northward and southward movement, which are strongly inconsistent with the strict stepping-stone hypothesis. Speciation of the *P. orbifer* complex may have occurred only between islands, suggesting that peripatric speciation through founders of stochastic dispersals was the major evolutionary driver. Evident long-distance dispersal between the northern and southern edges of weevil distribution reinforces the notion that extremely long-distance and multiple shorter dispersal are equally probable in oceanic islands. Coloration of *P. orbifer* can diverge rapidly, which indicates little phylogenetic conservatism. The statistically delimited seven cryptic *Pachyrhynchus* species and substantial genetic divergence between weevils of extremely close neighboring islands (Barit and Fuga) imply that the diversity of *Pachyrhynchus* from these small peripheral islands continues

to be largely under-estimated.

Introduction

Islands are excellent natural systems to study the diversification of organisms because of their small sizes, large numbers, young ages and geographical isolation (Carson 1983; Losos & Ricklefs 2009). For remote oceanic islands and archipelagos of volcanic origin, effective overwater dispersal and subsequent successful establishment in novel habitats are fundamental processes in generating endemic island biodiversity (MacArthur & Wilson 1967; Whittaker *et al.* 2008). Colonization of remote oceanic islands by continental species is mainly an interaction between the relative isolation of islands and the vagility of the organisms (Gillespie & Roderick 2002). Therefore, more remote islands usually have fewer colonizers than nearer islands, and vagile organisms often colonize more islands than their less mobile counterparts. In addition to the importance of distance on the dispersal organisms, colonizing remote oceanic island also depends mostly on the available ecological space, with most successful colonization occurring in the youngest islands within an isolated archipelago (Whittaker & Fernández-Palacios 2007).

Inter-island dispersal in an isolated oceanic archipelago is capable of producing a highly concordant colonization pattern consistent with the chronological sequence of the origin of its islands (Funk & Wagner 1995; Juan *et al.* 2000). The "stepping-stone model" of asymmetric colonization from geologically older to younger islands is the predominant mode in several well-documented archipelagos, showing a clear linear age sequence of the origin of islands, such as the Hawaiian and Galapagos Islands in the Pacific Ocean (Funk & Wagner 1995; Gillespie 2004; Parent *et al.* 2008), and the Canary Islands in the Atlantic (Juan *et al.* 2000; Planas & Ribera 2014). The stepping-stone hypothesis has been frequently proposed to explain the colonization and origin of endemic species on these hotspot archipelagos (Emerson 2002; Harbaugh *et al.* 2009).

However, some observations of stochastic dispersal suggested that the unpredictability of inter-island colonization is linked to the geological history of the islands in a more complex manner (Gillespie *et al.* 2012). The pattern of long-distance dispersal across oceanic barriers, previously perceived as being rare and almost impossible events, may be as equally probable as stepwise dispersal across islands (Crisp *et al.* 2011).

Situated in an unique biogeographical junction between the Asian and Philippine faunal regions known as Huxley's line (Lomolino *et al.* 2010), the Taiwan-Luzon volcanic belt (Fig. 1a) is currently one of the least groups of oceanic islands in Southeast Asia to be studied for its biogeographical history of island endemics. Endemic species of the Taiwan-Luzon volcanic belt were hypothesized to have originated through a northward stepping-stone dispersal from observations of decreasing number of shared Luzon species when moving from southern to northern islands (Oliveros *et al.* 2011; Ota 1991; Ota & Huang 2000). This northward stepping-stone model of colonization may have been facilitated by the regionally dominant surface current of Kuroshio, which originates at the equator and flows northerly toward the Philippines, Taiwan-Luzon volcanic belt and Japanese islands. The Kuroshio Current is thought to be a strong oceanic mechanism in shaping the distribution of tropical sea grasses (Kuo *et al.* 2006) and population genetic structures of skinks in the Ryukyu Archipelago (Kurita & Hikida 2014). It also served as a vehicle for northward larval transport of damselfishes between Taiwan and Okinawa (Liu *et al.* 2008). However, the stepping-stone model of colonization in the Taiwan-Luzon volcanic belt has never being rigorously tested using a molecular phylogeographical approach with the exception of some phylogenetic studies of shrews (Esselstyn & Oliveros 2010), birds (Oliveros & Moyle, 2010) and geckos (Siler *et al.*, 2014), which indicated the direction of colonization from either northern Taiwan (shrews) and Batanes (birds), or southern Luzon (geckos).

Nevertheless, the distribution of the species and taxon sampling of these earlier studies were limited with representatives from only a few islands across Taiwan-Luzon volcanic belt to fully address the direction and stepwise pattern of the stepping-stone hypothesis.

The Taiwan-Luzon volcanic belt consists of a group of remote volcanic islands adjacent to the continental shelf of mainland Asia, including southern Babuyan Islands, central Batanes Islands, northern Orchid Island (also known as "Lanyu" or "Botel Tobago") and Green Island (Lauto) of southeastern Taiwan (Fig. 1a). This Archipelago contains two parallel island chains (Eastern and Western) in a north-south direction, which were all created *de novo*, without any connections since their emergence to nearby larger islands (Luzon & Taiwan) or the Asian mainland (Voris 2000). The deep ocean channels surrounding these islands separate them from neighboring Luzon by >25 km to the south (Babuyan islands), and >20 km from the continental island of Taiwan to the west (Green Island). The Blintang Channel between Babuyan and Batanes Islands is ~ 75 km wide. North of Batanes Islands, approximately 150 km across Bashi Channel, lies Orchid Island. Although the two island chains are separated by only 50 km just north of Luzon and merge into a single island chain near Batan, most islands in the Western chain (Orchid, Itbayat, Sabtang, Calayan, Dalupiri and Fuga) are much older and mainly composed of volcanic rocks of Miocene to Pliocene age (>3.5 Mys). In contrast, all currently active volcanoes are in the Eastern chain (Green, Babuyan, Batan, Camiguin), where the majority of islands are younger and have origins in the Quaternary (<2.8 Mys) (Yang *et al.* 1996). Most of the islands of the Taiwan-Luzon volcanic belt were not connected to each other or Asian mainland during the Pleistocene glacial period (Voris 2000), which minimizes the effect of recent climatic changes on regional phylogeographical pattern (e.g., Nakamura *et al.* 2014). Within

each island chain, there is a general chronological order of island origins beginning from the older southern islands towards the younger northern islands (Yang *et al.* 1996). The complex spatial arrangement and temporal settings of the double arc island systems in the Taiwan-Luzon volcanic belt provides a challenge to study the colonization history and speciation processes of its endemic biota.

We set out specifically to test the stepping-stone model of colonization in a group of endemic weevils, *Pachyrhynchus orbifer* complex (Insecta: Coleoptera: Curculionidae), from the Taiwan-Luzon volcanic belt. *Pachyrhynchus* is a group of plant-feeding insects with a large number of endemic species in the Philippine Archipelago (Schultze 1923). Most pachyrhynchids are decorated with strikingly diverse color patterns, which function as aposematic visual signals to deter predation by lizards (Tseng *et al.* 2014). The *P. orbifer* complex represents an excellent set of organisms to test the stepping-stone hypothesis, not only because of the existence of closely related endemic species on nearly every sizable islands across the archipelago, but also due to three intriguing adaptations that may facilitate their cross-sea dispersal through rafting of floating vegetation by oceanic currents: (1) *Pachyrhynchus* adults are wingless and have completely fused elytra; (2) they have the ability to draw and keep air under their elytra, allowing them to float on water for at least several hours (Schultze 1923); (3) their eggs are inserted into plant tissues, and the larvae live and feed inside the stems of the host plants.

We first reconstructed the phylogeny, colonization history, divergence times, and historical demography of the *P. orbifer* complex using a multi-locus data-set to test the following specific predictions from the stepping-stone hypothesis: (1) the topology of phylogeny and inferred colonization history are consistent with asymmetric northward

colonization, initiating near the southern islands (Fuga & Camiguin) and ending at the far northern islands (Yaeyama) (Fig. 1b, hypothesis of stepping-stone by distance); Because the Taiwan-Luzon volcanic belt is a double arc island chain, the weevils were expected to disperse northwards initially from northern Luzon to Fuga of the western arc and Camiguin of the eastern arc, and then either colonize Dalupiri and Calayan from Fuga of the western arc, or colonize Babuyan, Itbayat, Orchid, Green and Yaeyama Islands in sequence; (2) the divergence time of species from the southern islands are greater than that of the species from northern islands in this sequence; (3) species from the southern islands have a history of a stable or declining population size (ancient colonization), whereas species from the northern islands showed a demographic expansion following a population bottleneck (recent colonization). Collecting trips to the Batanes and Babuyan Islands found that island populations of *P. orbifer* displayed similar stripe patterns, but different colors, in extant species. For example, *P. orbifer* from various islands of the Taiwan-Luzon volcanic belt all had narrow stripes similar to that of *P. sonani* from Green and Orchid Island, but the coloration of their stripes differ from each other (Fig. 1a, blue to green in Green, Orchid and Babuyan Islands; yellow in Dalupiri and Itbayat Islands; pink in Calayan and Camiguin Islands). In contrast, *P. orbifer* from Fuga Island have wider pink stripes resembling that of *P. orbifer* from northern Luzon. The extensive and characteristic stripe and color variation of endemic island populations suggest that the widespread *P. orbifer* (Schultze 1923) might comprise a group of cryptic species, which calls for closer investigation of its species boundary. Therefore, we used coalescent-based species delimitation methods to test species boundaries within the *P. orbifer* complex.

Materials and Methods

Taxon sampling

 Pachyrhynchus weevils were collected from 11 locations between northern Luzon and southern Ryukyus, including Cagayan Province (Claveria), Babuyan Islands (Camiguin, Fuga, Dalupiri, Calayan and Babuyan), Batanes Islands (Itbayat), Orchid Island, Green Island, and Yaeyama Islands (Ishigaki and Iriomote) (Fig. 1a, Appendix 1 in Supporting Information). An individual (NMNS7618-27) collected from Barit, a small island located ~1km from Fuga (Fig. 1a), was included into the population of Fuga. Ishigaki and Iriomote were connected during the glacial period (Ota, 1998), so we grouped individuals from these two islands into a single population. This sample collection included three known (Northern Luzon, Batanes, and Fuga Island) (Kano 1936; Schultze 1923) and four newly recorded populations of *P. orbifer* and closely related species (*P. sonani* Kono, 1930 and *P. infernalis* Kano, 1897). The *Pachyrhynchus orbifer* complex is a species group mainly distributed in northern Luzon (Ilocos Norte and Cagayan Province) and Babuyan Islands (Schultze 1923). Although we did not sample all historically recorded populations in northern Luzon, our sampling from Taiwan-Luzon volcanic belt was adequate to infer phylogenetic relationships of the selected island populations. Specimens of closely related *P. speciosus* Waterhouse, 1841 (Samar Island), *P. tobafolius* Kano, 1929 (Orchid Island), and *P. nobilis yamianus* Kano, 1929 (Orchid Island) were used as outgroups for the phylogenetic analyses. A preliminary species-level phylogeny of *Pachyrhynchus* suggested that they are the closest relatives of the *P. orbifer* complex (Tseng *et al*., unpublished data). Field-collected specimens were preserved in 95% ethanol, and later stored in a −20 °C freezer at the insect collection of the National Museum of Natural Science, Taichung, Taiwan.

DNA extraction and sequencing

Genomic DNA of the specimens was extracted from the legs or thoracic muscles using Genomic DNA Extraction Tissue DNA Mini Kit (Favorgen Biotech, Taiwan). The DNA fragments of two mitochondrial (cytochrome c oxidase subunit 1, *CO1*; NADH dehydrogenase subunit 2, *ND2*) and two nuclear genes (elongation factor 1-α, *EF1-α*; ribosomal internal transcribed spacer, *ITS*) were amplified using a combination of newly designed and published primers (Appendix 2). PCR protocols followed the profiles described by Huang and Lin (2010), except for the annealing temperatures of 52, 48, 64, and 56°C for *CO1*, *ND2*, *EF1-α*, and *ITS* gene, respectively. The amplified PCR products were sequenced directly from both directions using an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA). The DNA sequences (GenBank accessions KU953405-KU953778) were edited and assembled manually, and aligned using ClustalW (Tamura *et al.* 2013). The aligned sequences of protein-coding genes were translated into amino acid sequences using a standard genetic code (exon of *EF1-α*) and invertebrate mitochondrial genetic codes (*CO1* and *ND2*) in Mega6 (Tamura *et al.* 2013) to confirm codon assignments and possible stop codons.

Phylogenetic analyses and topology test

Phylogenetic relationships were reconstructed using maximum likelihood (ML) in RAxML-VI-HPC (Stamatakis 2006) and Bayesian inference (BI) in MrBayes v3.2.2 (Ronquist *et al.* 2012). The best-fit models of nucleotide substitution in ML and BI analyses were selected for each gene and codon partitions with linked branch lengths in PartitionFinder v1.1.0 (Lanfear *et al.* 2012) using Bayesian Information Criterion (BIC). The GTRGAMMA model was used in ML analyses for all partitions, with the branch support assessed by 1,000 bootstrap replicates of the rapid hill-climbing algorithm. The tree searches of BI were analyzed using 1×10^{8} generations of Markov chain Monte Carlo (MCMC) processes, with trees sampled for every 1,000 generations.

The best-fitted models of nucleotide substitution for each codon partition (Appendix 3) were used in the BI analysis. The first 25,000 trees (25%) were discarded as burn-in, and the remaining trees were used to compute a 50% majority rule consensus tree, with the percentage of trees recovering the node representing the node's Bayesian posterior probability (BPP). Convergence of the MCMC processes was assessed based on the values of the Effective Sample Size (ESS>200) of all model parameters in Tracer v1.6 (Rambaut *et al.* 2014), and by examining the average split frequencies between runs (<0.01). Phylogenetic data was deposited in Dryad Digital Repository (doi:10.5061/dryad.25fc4). We evaluated the topological prediction (Fig. 1b) using likelihood-based methods in CONSEL v0.1i (Shimodaira & Hasegawa 2001). Per site log likelihoods of alternative tree topologies (constrained topology of the stepping-stone hypothesis, Fig. 1b vs. the topology of our best tree, Fig. 2) were estimated from ML heuristic searches of 1,000 generations using RAxML-VI-HPC (Stamatakis 2006). The P-value of the approximately unbiased (AU) test, Kishino-Hasegawa (KH) test, Shimodaira- Hasegawa (SH) test, weighted Kishino-Hasegawa (WKH) test and weighted Shimodaira-Hasegawa (WSH) test, were calculated with 10,000 bootstrap replicates. The topological hypotheses with a p-value <0.05 were rejected.

Divergence time estimation

The divergent time was estimated in BEAST v1.8.2 (Drummond *et al.* 2012), based on an uncorrelated lognormal relaxed clock model. The substitution models were unlinked among gene partitions. The upper bound and lower bounds of the mutation rate in mitochondrial DNA (mtDNA) were applied, based on the standard (0.0115 mutations site^{-1} Myr^{-1}; Brower 1994) and a revised molecular clock for insects (0.0177 mutations site^{-1} Myr^{-1}; Papadopoulou *et al.* 2010). The substitution rate of *EF1-α* and *ITS* followed the mutation rate of synonymous substitution for nuclear genes (0.016

mutations site^{-1} Myr^{-1}; Moriyama & Gojobori 1992). All priors were set to default values, except for the priors of mutation rates. The analysis was run for 1 x 10^8 generations with a sampling frequency of every 1,000 generations. The values of ESS of the parameters were assessed to confirm the convergence of MCMC runs in Tracer (Rambaut *et al.* 2014). The first 25% of the resulting trees were discarded and the consensus tree was generated in a maximum clade credibility (MCC) tree obtained in TreeAnnotator v1.8.2 (Drummond & Rambaut 2007).

Ancestral range and historical dispersal

The ancestral range and biogeographical history were reconstructed using Bayesian Binary MCMC (BBM) (Ronquist & Huelsenbeck 2003) methods implemented in RASP 3.2 (Yu *et al.* 2015) and a R package, BioGeoBEARS (Matzke 2013a). BBM analysis used the best trees generated from BI analyses (Fig. 2), with the range of the specimen coded as the island of origin. The MCMC chains of five independent runs of 5x10^5 generations were conducted, with a sampling frequency of 1,000 iterations and a burn-in of 100 samples. The fix state frequency of Jukes-Cantor (JC) model was used and the among-site rate variation was set at equal. BioGeoBEARS implements likelihood version of various biogeographic models: Dispersal-Vicariance (DIVA)(Ronquist 1997), Dispersal-Extinction Cladogenesis (DEC) (Ree *et al.* 2005; Ree & Smith 2008) and BayArea (Landis *et al.* 2013). BioGeoBEARS implements the "founder-event" parameter (+J) to each of the above models, which is a particularly important parameter for modeling speciation on remote oceanic islands (Matzke 2013b; 2014). A total of six models were used (DEC-LIKE, DEC-LIKE+J, DIVA-LIKE, DIVA-LIKE+J, BAYAREA-LIKE, BAYAREA-LIKE+J). The islands of origin for the ten sampling localities were coded as discrete ranges. The phylogenetic trees resulting from Bayesian inference of BEAST were used as an input file to estimate the

probabilities of ancestral ranges. The fit of the six biogeographical models were compared using likelihood values and Akaike Information Criterion (AIC).

Population demographic history

The demographic histories of weevils on 10 localities (Claveria, Camiguin, Fuga, Dalupiri, Calayan, Babuyan, Itbayat, Orchid, Green and Yaeyama Islands) were inferred from multi-locus data, using coalescent-based Extended Bayesian Skyline Plot (EBSP) (Heled & Drummond 2008) implemented in BEAST v1.8.2 (Drummond *et al.* 2012). The substitution models were set according to the best model selected for each island population. A strict clock model was used, as recommended for intra-species data sets (Drummond & Bouckaert 2015). The MCMC runs of 10^8 generations with a sampling frequency of 1,000 generations were conducted, with values of ESS of all parameters accessed for confirmation of convergence (ESS>200). The sample size of three populations (Babuyan, N=4; Calayan, N=3; Claveria, N=6) was near the minimum number for EBSP. Nevertheless, we conducted EBSP of these populations as preliminary estimates of their demographic histories because simulation studies showed sample size to be less sensitive to estimation error than the number of loci (Heled & Drummond 2008). The number and 95% highest posterior distribution (HPD) of change in population demographic size were inferred. The population was inferred to be demographically stable during the time period when the number of population size change events was insignificantly different from zero (Lim & Sheldon 2011).

Species delimitation

The species boundaries of these weevils were estimated using the multi-species coalescent model (MSC) and reversible-jump Markov chain Monte Carlo (rjMCMC) algorithms in BP&P 3.1 (Yang 2015). The priors of the two model parameters, θ_s (the

ancestral population size) and τ_0 (the root age of the tree), were assigned a Gamma G (α, β) distribution. Three different combinations of τ_0 and θ_s were applied: (1) $\theta_s \sim$ G (1, 10) and $\tau_0 \sim$ G (1, 10), assuming relatively large ancestral population sizes and deep divergences; (2) $\theta_s \sim$ G (2, 2000) and $\tau_0 \sim$ G (2, 2000), representing relatively small population size and shallow divergences; and (3) $\theta_s \sim$ G (2, 2000) and $\tau_0 \sim$ G (1, 10), specifying relatively small population sizes and deep divergences. The species tree was estimated from *BEAST (Heled & Drummond 2010) implemented in BEAST v1.8.2 (Appendix 4). Tree prior of the Yule process and uncorrelated lognormal relax clock were selected in the analysis. The best fitted substitution models and substitution rates for each gene were derived from the setting of the divergence time estimation. This version of BP&P used nearest-neighbor interchange (NNI) branch swapping of the specified guide tree to incorporate phylogenetic uncertainty in species delimitation (Yang & Rannala 2014). The mutation rate prior among loci was specified as a Dirichlet distribution D (α=2.0) because of combining mitochondrial and nuclear genes. The rjMCMC analysis was run for 1 x 10^6, with a sampling frequency of every five generations and a burn-in of the first 1 x 10^5 generations. All analyses were run twice to confirm the consistency between runs.

Results

Phylogeny and topological test of stepping-stone hypothesis

A sequence matrix of 4050 base pairs were obtained from assembling *CO1* (675 bp), *ND2* (983 bp), *EF1-α* (810 bp), and *ITS* fragments [1582 bp, including *18S* (226 bp), *ITS1* (798 bp), *5.8S* (163 bp), and *ITS2* (390 bp), and *28S* (5 bp)] of 93 ingroup and three outgroup taxa. PCR amplification and sequencing of *ITS* from 10 specimens were unsuccessful (Appendix 1, 9 ingroup and 1 outgroup samples). The *CO1*, *ND2* and *ITS* gene trees all resulted in well-resolved phylogenetic relationships, with the majority of lineages clustered by island origins (Appendix 5). In contrast, the *EF1-α* gene tree showed only a few clusterings by island origins (Green+Itbayat, Dalupiri and Fuga), but limited phylogenetic structure among island lineages. The topologies of each of the four gene trees showed little consistency, except for the monophyly of the lineages from Green and Itbayat Islands, and a close relationship between the lineages of Fuga and Yaeyama Islands (*CO1*, *ND2*, and *ITS*). The mitochondrial *CO1* and *ND2* gene trees consistently formed a monophyletic group of Babuyan and Orchid Islands, whereas the nuclear *EF1-α* and *ITS* gene trees suggested a close relationship between the lineages of Babuyan and Camiguin (*EF1-α*)/Calayan (*ITS*) Islands, and a monophyletic lineage of Orchid and Green+Itbayat Islands. The incongruence between mitochondrial and nuclear gene trees may be associated with the gene tree/species tree problem, possible introgression, sex-biased dispersal, or incomplete lineage sorting of ancestral polymorphism. Nevertheless, the topological incongruence between mitochondrial and nuclear loci was weakly supported, suggesting little discrepancy. The phylogeny of the combined data indicated a well-resolved tree with strong branch supports (LB/BPP > 99%/1.0) for seven reciprocally monophyletic lineages grouped by current island boundaries (Camiguin, Fuga, Dalupiri, Calayan, Babuyan, Orchid and Yaeyama Islands), except for a sister Green+Itbayat lineage (Fig. 2, node 1). The lineages of

Green, Itbayat and Orchid Islands (Fig. 2, node 10), Yaeyama and Fuga Islands (Fig. 2, node 12), and Babuyan and Camiguin Islands (Fig. 2, node 14) each formed a monophyletic group, with high to moderate branch support (LB/BPP = 79-100%/0.94-1.00). The lineages of Green, Itbayat, Orchid, Yaeyama and Fuga Islands formed a monophyletic group with high branch support (LB/BPP = 81%/1.00) (Fig. 2, node 11), but the relationship of Babuyan+Camiguin, Calayan and northern Luzon were only strongly supported by Bayesian inference (LB/BPP = 57-59%/1.00) (Fig. 2, node 15, 16). One of the basal nodes was not well supported (Fig. 2, node 13). Because the island populations were reciprocally monophyletic, we also reconstructed the species tree using coalescence-based analyses in *BEAST (Appendix 4). The majority of sister relationships of the species tree are similar to that of the tree inferred from concatenated data, with the inconsistency of tree topologies not well supported. One notable exception was the most basal lineage of Dalupiri Island in the species tree. Maximum likelihood topological tests (AU, KH, SH, WKH and WSH) of phylogenetic hypotheses all significantly supported the topology of the best tree (Table 1; Fig. 2), and rejected the tree topology derived from the stepping-stone hypothesis (p<0.05; Fig. 1b).

Divergence time and colonization history

The divergence time estimation provided a mean estimate of ~ 0.91 Mya [95% confidence interval (CI): 0.61-1.31 Mya] for the origin of the *P. orbifer* complex (Fig. 3, node 17). The most recent common ancestors (MRCA) of all nine extant island lineages of the *P. orbifer* complex fell within the time-frame of the late Ionian to Tarantian (in the late Pleistocene) between 0.27 and 0.05 Mya (Fig. 3, node 7, 95% CI: 0.12-0.48 Mya; node 18, 95% CI: 0.01-0.12 Mya), with the majority of island lineages originating ~ 0.2 Mya. The youngest island lineage was found only ~ 50,000 years ago in Itbayat (Fig. 3, node 18). The majority of the southern island lineages (Calayan,

Camiguin, Babuyan and Dalupiri, but not Fuga) were older than that of northern

counterparts (Itbayat, Green and Yaeyama Islands), except for Orchid Island being

among the older island lineages (Fig. 3). BioGeoBEARS analyses showed that the best-

fit model was DIVA-LIKE+J (LnL= -37.6159, AIC=81.2317) (Table 2). This model

inferred a process of combining dispersal, extinction and peripatric events for

diversification among extant islands, and highlighted the essential process of founder

speciation for generating endemic diversity in these oceanic islands. Ancestral range

reconstruction of BBM and DIVA-LIKE+J gave comparable results in recently

diverged nodes (Fig. 3). Among deeper nodes, only BBM distinguished higher probable

ancestral areas, whereas the DIVA-LIKE+J model showed ambiguous reconstruction

(Fig. 3). This could be a result of a high number of ancestral area combinations or higher

phylogenetic uncertainty at basal nodes. We therefore interpreted the colonization

history based on the results of BBM analyses. Two major historical colonization routes

were identified for *P. orbifer* complex (Fig. 4, solid vs. open arrows); the first route

(solid arrows) suggested that the ancestral *P. orbifer* probably originated from northern

Luzon [given equally probable ancestral range of Luzon (32.4%) and Dalupiri (32.4%)],

then first split and moved to Dalupiri Island at ~ 0.84 Mya (Fig. 3, node 13, 95% CI:

0.56-1.22 Mya; Fig. 4a, dispersal 1). The descendent lineage of Dalupiri Island

dispersed remotely to northern Yaeyama Island (Fig. 3, node 11, Yaeyama: 20.19%) at

~ 0.76 Mya (95% CI: 0.49-1.10 Mya; Fig. 4a, dispersal 3), then spread southward to

Green Island (Fig. 3, node 10, Green: 47.9%, Orchid: 40.21%) at ~ 0.53 Mya (95% CI:

0.34-0.80 Mya; Fig. 4b, dispersal 5), and later traveled a long distance from Yaeyama

southward to Fuga Island (Fig. 3, node 12, Fuga: 45.75%) at ~ 0.47 Mya (Fig. 3, node

12, 95% CI: 0.23-0.78 Mya; Fig. 4c, dispersal 7). The ancestral lineage of Orchid Island

originated from Green Island (Fig. 3, node 2, Green: 99.7%) ~ 0.23 Mya (Fig. 3, node

2, 95% CI: 0.08-0.46 Mya; Fig. 4d, dispersal 8), and later the ancestral Green Island

lineage dispersed southward toward Itbayat Island just 50,000 years ago (Fig. 3, node 18; Fig. 4d, dispersal 9). The second colonization route (open arrows) indicated that the ancestral *P. orbifer* dispersed from northern Luzon to Calayan Island (Fig. 3, node 16, northern Luzon: 71.64%; node 15, Calayan: 67.09%) at ~ 0.76 Mya (Fig. 3, node 16, 95% CI: 0.49-1.14 Mya; Fig. 4a, dispersal 2), subsequently colonized nearby Camiguin Island (Fig. 3, node 15, Camiguin 41.68% and Babuyan 41.55%) at ~ 0.64 Mya (Fig. 3, node 15, 95% CI: 0.39-1.00 Mya; Fig. 4b, dispersal 4), and later spread to neighboring Babuyan Island nearly 0.50 Mya (Fig. 3, node 14, 95% CI: 0.27-0.80 Mya; Fig. 4c, dispersal 6). Because a few equally probable ancestral areas were reconstructed at node 10, 11, 17 and 14 (Fig.3), we inferred alternative colonization histories to assess the robustness of our interpretation of the data (Appendix 6). All eight colonization scenarios contained a few long distance dispersals, and the numbers of southward and northward events were similar.

Species status and demographic history

The BPP species delimitation suggested a total of ten species corresponding to current island boundaries (Fig. 3, nodes marked with an asterisk) and supported by high posterior probability (0.99-1.0; Appendix 7). The results indicate that previously recognized *P. sonani* might consist of two cryptic species originated from Green and Orchid Island (Fig. 3, nodes 1 and 2). Other lineages with similar color stripes from neighboring (N. Luzon vs. Fuga; Camiguin vs. Calayan) and distant islands (Babuyan vs. Green/Orchid) were each assigned to putative species (Fig. 3; Appendix 7). EBSP analyses suggested that the 95% HPD of the number of demographic population size change included zero in nine island populations (except Claveria of N. Luzon; Appendix 8), indicating that historical sizes of these populations largely remained constant during the time period. The population of *P. orbifer* from northern Luzon

undergone recent population growth starting ~ 0.1 Mya (Appendix 8). The populations of Green, Itbayat, Babuyan, Camiguin and Dalupiri Islands had the highest frequency of one population size change, suggesting only a minor trend of increase in recent population size in these populations.

Discussion

Rejecting the strict stepping-stone hypothesis of island colonization

The *P. orbifer* complex of Taiwan-Luzon volcanic belt had a complex colonization history of combining both northward and southward directions with short and long distance dispersal events. By consideration of equally probable colonization scenarios, the inferred colonization history is therefore strongly inconsistent with the strict stepping-stone hypothesis, which assumes an asymmetric northward colonization starting from near the southern islands toward the far northern islands in sequence. Several lines of additional evidence further suggest the rejection of the strict stepping-stone hypothesis. First, the divergence time of all extant island lineages of the *P. orbifer* complex fall within the last 1 Mys of Middle Pleistocene. This time-frame of diversification is much more recent than the geological age of the islands in the Taiwan-Luzon volcanic belt (average 22.4-1.7 Mys). Second, although most island lineages of the southern Taiwan-Luzon volcanic belt are older than that of their northern counterparts, the temporal order of origin of these island lineages did not follow a linear sequence of spatial arrangement in the archipelago. Third, historical demography suggests that all island lineages show no recognizable population fluctuation, and there was no apparent pattern of southern stable vs. northern expansion among island lineages. Together these findings establish that the direction, sequence, timing and demographic history of colonization of the *P. orbifer* complex were not correlated with the geo-historical organization of the islands in a south-north sequence, as predicted by a strict stepping-stone model.

Many terrestrial organisms on hot-spot archipelago have followed a strict stepping-stone colonization model, and diversified in parallel to the geological formation of the islands (Gillespie 2004; Parent *et al.* 2008; Planas & Ribera 2014).

However, the history of inter-island colonization of a few species, for example, Darwin's finches, *Galagete* moths and *Galapaganus* weevils of Galápago Islands (reviewed in Parent *et al*., 2008), are deviations from such a model. These exceptions were likely due to considerable inter-island dispersal of highly vagile species (Darwin's finches, Petren *et al.* 2005), a more complex colonization history combined with possible ancient colonization of now submerged islands (*Galapaganus* weevils, Sequeira *et al.* 2008), or back-colonization expected from a stochastic dispersal pattern (*Galagete* moths, Schmitz *et al.* 2007). *Pachyrhynchus* weevils are wingless, sedentary insects with a considerable degree of host plant specialization that precludes the mobility of organisms as a possible explanation for the inferred complex colonization history. Insular *P. orbifer* is a recently diverged lineage of the Pleistocene, when all extant islands of the archipelago had emerged above sea level. Thus, non-linear, more complex spatial settings of the Taiwan-Luzon volcanic belt and stochastic dispersal are probable key factors shaping the colonization history of the *P. orbifer* complex. Geomorphologic evidence indicates that most islands of the Taiwan-Luzon volcanic belt were submarine, and only uplifted above the sea level ~ 3 Mya (Yang *et al.* 1996). Therefore, the chronological sequence of island emergence may not necessarily follow a south-north direction of geographical arrangement in a stepping-stone fashion.

Complex inter-island colonization and peripatric founder speciation as major drivers for diversification

One of the most striking patterns of *Pachyrhynchus*'s diversification is the apparent stochastic nature of inter-island colonization. Nevertheless, diversification of the *P. orbifer* complex in the Taiwan-Luzon volcanic belt appears to have occurred only between islands. The inferred colonization history and species delimitation show that colonization of new islands always leads to speciation among *Pachyrhynchus* lineages.

These findings therefore strongly suggest that peripatric speciation through founders of stochastic dispersals was the major evolutionary driver of diversification. Peripatric speciation by founder events can be a predominant mode of speciation on oceanic islands, given prolonged periods of isolation between ancestral and founder populations, and through the effect of genetic drift and environmental heterogeneity among islands (Barton & Charlesworth 1984; Mayr 1954; Templeton 1980). Our finding of no closely related *Pachyrhynchus* species inhabiting the same island indicates that the process of *in situ* speciation within individual islands was less important than inter-island diversification. Ecological differentiation within an island does not seem to have happened among all lineages of the *P. orbifer* complex in the Taiwan-Luzon volcanic belt.

The second unique feature of *Pachyrhynchus*'s diversification is the evident long-distance dispersal between the northern and southern edges of their distribution. At least three of the nine evolutionarily successful colonization events accomplished were caused by long-distance dispersal (i.e., dispersal between non-neighboring extant islands in the archipelago; Fig. 4, dispersal events 3, 7 and 9) (alternative scenarios, Appendix 6). This finding reinforces the notion that a single extreme long-distance dispersal event could be equally, or more probable than multiple shorter dispersal events (i.e., the stepping-stone model), given that probabilities multiply in a chain of independent consecutive rare events (Hardy 2009; Nathan 2006). Within the Taiwan-Luzon volcanic belt, long-distance dispersal by skipping the stepping-stone islands in between has been inferred for the Philippine bulbuls (*Microscelis amaurotis*) from Japan to Babuyan and Batanes Islands (Oliveros & Moyle 2010), and for Kikuchi's geckos (*Gekko kikuchii*) from Luzon to Orchid Island (Siler *et al.* 2014). However, the prevalence of long-distance dispersal in the Taiwan-Luzon volcanic belt shown not only

here, but in earlier studies requires further examination by comparative analyses of colonization histories across taxonomic groups.

The enormous diversity of weevil species is often attributed to adaptation and co-evolution to the complexity and diversity in phenotypes, life histories, and habitats of flowering plants (McKenna *et al.* 2009). Available host plant records of the *P. orbifer* complex (Appendix 9) suggest that inter-island diversification of these weevils is almost invariably associated with shifts of host plant range or to distantly related host plants. For example, the host plants of *P. sonani* from Green Island are Ceylon ardisia (*Ardisia elliptica*, Myrsinaceae), Elephant's ear (*Macaranga tanarius*, Euphorbiaceae), Iron-wood (*Casuarina equisetifolia*, Casuarinaceae) and Melastoma (*Melastoma* prob. *Affine*, Melastomataceae) (Starr & Wang 1992). When ancestral *P. sonani* of Green Island colonized Orchid Island, its host plant range shifted to include Fish poison tree (*Barringtonia asiatica*, Lecythidaceae) and Iron-wood. Later, ancestral *P. sonani* of Green Island colonized Itbayat Island and restricted its host plant range to Ceylon ardisia. Historical host shifts also occurred between ancestral *P. orbifer* of Dalupiri (Elephant's ear) and derived *P. infernalis* (Small-leaved barringtonia, *B. racemosa*, Lecythidaceae ; *Glochidion*, Euphorbiaceae; Mango, *Mangifera indica*, Anacardiaceae; Carambole, *Averrhoa carambola*, Oxalidaceae) (Oshiro 1991), and between closely related *P. orbifer* of northern Luzon (Guava, *Psidium guajava*, Myrtaceae) and Camiguin (Fish poison tree). We hypothesize that, in addition to founders of stochastic dispersals, ecological adaptation to specialized host plants may also have played an important role in differentiation among insular *Pachyrhynchus* species in the Taiwan-Luzon volcanic belt. However, at present the incomplete host records and occasional host records of cultivated plants limit our interpretation of co-evolution under a phylogenetic framework between the *P. orbifer* complex and their host ranges.

Mechanisms of overwater dispersal of remote oceanic islands

Our analyses suggested that flightless *Pachyrhynchus* weevils were capable of both short- and relatively long-distance dispersal across the ocean, despite being able to stay afloat on water for just several hours (Schultze, 1923). However, the mechanism and transporting vectors of effective long-distance dispersal across oceanic barriers is currently unknown for these weevils. One potential mechanism of cross-oceanic dispersal is drifting on floating vegetation, wood and debris due to oceanic currents. Rafting of terrestrial organisms to remote oceanic islands has been reported from a range of floating substrates and across all major oceans of the world (Peck 1994; Thiel & Gutow 2005). In particular, live adult weevils of *Sphenophorus* sp. and *Macrancylus linearis* were found on drifting debris offshore of Puerto Rico (Heatwole & Levins 1972). Therefore, it is probable that *Pachyrhynchus* weevils may have colonized the Taiwan-Luzon volcanic belt through rafting as adults on floating substrates, or as eggs, larva and pupae inside the stems and fruits of their host-plants. Larvae of wood-boring beetles like *Pachyrhynchus* weevils are among the most frequent insects found on floating items on the open sea (Heatwole & Levins 1972). This wood-boring habit may facilitate their isolation from saltwater, thereby increasing their survival at sea. One of the host-plants of *Pachyrhynchus* weevils, Fish poison tree (*B. asiatica*), has rather large box-shaped fruit with thick spongy fibrous layers. The fruit is extremely water-resistant and buoyant, and possibly represents adaptations for sea dispersal by ocean current (Tsou & Mori 2002). These specialized fruits of *B. asiatica* could serve as floating vehicles, and at the same time as a food source during oceanic journey for larval *Pachyrhynchus*.

The regional dominant Kuroshio Current may also have been important in

facilitating northward rafting of *Pachyrhynchus* weevils across the Taiwan-Luzon volcanic belt, especially the relatively long-distance dispersal event from Dalupiri to Yaeyama Islands with a straight-line distance of >600 km. Nevertheless, only four of nine historical colonization events in *Pachyrhynchus* weevils have been northward dispersals (Fig. 4, blue arrows), and three of the these four events were relatively short-distance dispersals between neighboring islands at the southern end of the Taiwan-Luzon volcanic belt (Fig. 4, dispersal 1, 2, and 6). The remaining historical colonization events were all southbound dispersals (Fig. 4, red arrows), including two long-distance dispersal events from Yaeyama Islands to Fuga (ca. 650 km) and Green to Itbayat Island (ca. 215 km) (Fig. 4, dispersal 7 and 9, respectively). These findings indicate other unidentified mechanisms besides Kuroshio Current, such as prevailing northeastern monsoons, strong typhoon winds or stochastic movements (largely moving north-south direction) of surface flows in the region (Stommel & Yoshida 1972), could have directed the historical southward colonization routes of *Pachyrhynchus* weevils.

In addition to rafting, although less likely, *Pachyrhynchus* weevils could disperse across oceans by means of animal migration. Migratory birds are important biological agents for efficient transportation of animals and plants because they are excellent flyers with regular pathways (Holzapfel & Harrell 1968; Gillespie *et al.* 2012; Incagnone *et al.* 2015). Plants and invertebrates can be transported by migratory birds by attaching to the birds' feathers or digestive tracts (Brochet *et al.* 2009, 2010a, 2010b; Rogers 2014). One of the important migratory routes in East Asia for large birds, such as gray-faced buzzards (*Butastur indicus*), is between Japan and the Philippines by way of the Taiwan-Luzon volcanic belt (Nuytemans 1998; Shiu *et al.* 2006). *Pachyrhynchus* weevils might accidentally use their strong grasping ability to hold on to bird's bodies when the migratory birds stop in flocks on the forest canopy of these islands. However,

there is currently no empirical data to support this dispersal mechanism.

Prehistoric human expansion can shape the distribution of insular biota (Austin 1999; Chang *et al.* 2015). In one case, the Austronesians expanded from Taiwan into Oceania ~ 6000 years ago, and during the process facilitated southern dispersal of paper mulberry (Chang *et al.* 2015). However, we can rule out the possibility of human-mediated dispersal of *Pachyrhynchus* weevils across the Taiwan-Luzon volcanic belt based on the time-frame of weevil's colonization (0.25-1 Mya), much earlier than that of recent Austronesian expansion.

Cryptic species diversity and conservation of island endemics

The Philippine archipelago is well known as one of global biodiversity hotspots (Myers *et al.* 2000). A growing body of evidence predominantly focusing on vertebrates has characterized the role of Pleistocene island connectivity of the larger Philippine islands (e.g., PAIC, Pleistocene Aggregate Island Complex), and of other ecological and evolutionary factors on patterns and processes of Philippine diversification (reviewed in Heaney 1985; Brown & Diesmos 2009; Brown *et al.* 2013). However, the diversification of invertebrate fauna in the Philippines was poorly studied until recently (Su *et al.* 2014), especially for species colonizing the smaller islands at the fringe of the archipelago. Our study provides the first compelling evidence of a recent diversification of invertebrates across small oceanic islands lying at the northern periphery of the Philippine archipelago between Taiwan and Luzon. In contrast to earlier taxonomic arrangements of *Pachyrhynchus* weevils based mainly on color pattern (Kano 1929; Schultze 1923; but see Yoshitake 2013), we have evidence of seven cryptic, but divergent, monophyletic species of *P. orbifer*, as delimitated through the statistical inference of genetic data. Species of *P. orbifer* complex with similar coloration were

sometimes not the most closely related sister taxa (e.g., *P. sonani* of Orchid Island and *P. orbifer* of Babuyan Island; *P. orbifer* of Fuga and northern Luzon). These findings indicate that color markings of *P. orbifer* complex can evolve rapidly and show little phylogenetic conservatism, such as the loss of coloration in *P. infernalis* on Yayeyama Island and frequent evolutionary transitions of stripe colors between sister *Pachyrhynchus* species. We propose that all seven identified *P. orbifer* should be recognized as candidate species on the basis of monophyletic lineages, the statistical species limits, and in combination with their isolated distributions, host plant ranges and color variation. A formal taxonomic revision by integrating morphological, ecological and molecular traits in this complex is needed, with the possibility of finding further species by exploring smaller islands across the archipelago.

The most remarkable scenario of cryptic diversification in this study comes from the unpredictably large genetic divergence between morphologically identical *P. orbifer* of Barit Island (Fig. 2, P040) and Fuga Island (average genetic distances of four genes, $0.3 \pm 0.1\%$ vs. $0.1 \pm 0.04\%$ between Itbayat and Green Island), these islands being only 1.2 km apart (Fig. 1a). The substantial genetic divergence between weevils of these two extremely close neighboring islands demonstrates a strong isolation effect of the oceanic barrier, and highlights the fact that diversity of cryptic species of *Pachyrhynchus* weevils on these smaller islands may still largely be underestimated. However, Fuga and Barit (and Dalupiri) Island were unified into a single larger island during the Pleistocene glacial cycles. Other unidentified mechanisms may need to be considered in addition to the simple scenario of isolation-based divergence. Since 2009, the Wild Life Conservation Act has officially listed all Taiwanese *Pachyrhynchus* species as protective species of the Category II (i.e., rare and valuable species; Chao *et al.*, 2009). Therefore, identifying cryptic species and distinct evolutionary lineages is

especially important in setting conservation priority and strategies for endangered weevils in these remote oceanic islands. Conservation of insular species and endangered populations in their peripheral ranges is also crucial for preserving rare alleles and unique phenotypes. These rare alleles and localized phenotypes are of particular interest because they may be responsible for adaptation of endemic island species to local habitats.

References

Austin CC (1999) Lizards took express train to Polynesia. Nature **397**, 113-114.

Barton NH, Charlesworth B (1984) Genetic revolutions, founder effects, and speciation. Annual Review of Ecology and Systematics **15**, 133-164.

Brochet AL, Gauthier-Clerc M, Guillemain M*, et al.* (2010a) Field evidence of dispersal of branchiopods, ostracods and bryozoans by teal (*Anas crecca*) in the Camargue (southern France). Hydrobiologia **637**, 255-261.

Brochet AL, Guillemain M, Fritz H, Gauthier-Clerc M, Green AJ (2009) The role of migratory ducks in the long-distance dispersal of native plants and the spread of exotic plants in Europe. Ecography **32**, 919-928.

Brochet AL, Guillemain M, Fritz H, Gauthier-Clerc M, Green AJ (2010b) Plant dispersal by teal (*Anas crecca*) in the Camargue: duck guts are more important than their feet. Freshwater Biology **55**, 1262-1273.

Brower A (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences **91**, 6491-6495.

Brown RM, Diesmos AC (2009) Philippines, biology. In: *Encyclopedia of Islands* (eds. Gillespie RG, Clague DA), pp. 723-732. University of California Press, Berkeley, CA.

Brown RM, Siler CD, Oliveros CH*, et al.* (2013) Evolutionary processes of diversification in a model island archipelago. Annual Review of Ecology, Evolution, and Systematics **44**, 411-435.

Carson HL (1983) Chromosomal sequences and interisland colonizations in Hawaiian Drosophila. Genetics **103**, 465-482.

Chang C-S, Liu H-L, Moncada X*, et al.* (2015) A holistic picture of Austronesian migrations revealed by phylogeography of Pacific paper mulberry. Proceedings

of the National Academy of Sciences **112**, 13537-13542.

Chao, J-T, Yang M-M, Wu M-H (2009) Setting assessing criteria for the conservation of Taiwan's threatened insects. Conservation Research Series No. 97-04. Council of Agriculture, Executive Yuan, Taiwan. (in Chinese)

Crisp MD, Trewick SA, Cook LG (2011) Hypothesis testing in biogeography. Trends in Ecology & Evolution **26**, 66-72.

Drummond AJ, Bouckaert RR (2015) *Bayesian evolutionary analysis with BEAST* Cambridge University Press, Cambridge.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology **7**, 214.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution **29**, 1969-1973.

Emerson B (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. Molecular Ecology **11**, 951-966.

Esselstyn JA, Oliveros CH (2010) Colonization of the Philippines from Taiwan: a multi-locus test of the biogeographic and phylogenetic relationships of isolated populations of shrews. Journal of Biogeography **37**, 1504-1514.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology **3**, 294-299.

Funk VA, Wagner WL (1995) Biogeographic patterns in the Hawaiian Islands. In: *Hawaiian biogeography: evolution on a hot spot archipelago* (eds. Funk VA, Wagner WL), pp. 379-419. Smithsonian Institution, Washington.

Gillespie RG (2004) Community assembly through adaptive radiation in Hawaiian spiders. Science **303**, 356-359.

Gillespie RG, Baldwin BG, Waters JM*, et al.* (2012) Long-distance dispersal: a framework for hypothesis testing. Trends in Ecology & Evolution **27**, 47-56.

Gillespie RG, Roderick GK (2002) Arthropods on islands: colonization, speciation, and conservation. Annual Review of Entomology **47**, 595-632.

Harbaugh DT, Wagner WL, Allan GJ, Zimmer EA (2009) The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). Journal of Biogeography **36**, 230-241.

Hardy OJ (2009) How fat is the tail? Heredity **103**, 437-438.

Heaney L (1985) Zoogeographic evidence for middle and late Pleistocene land bridges to the Philippine Islands. Modern Quaternary Research in Southeast Asia **9**, 127-144.

Heatwole H, Levins R (1972) Biogeography of the Puerto Rican Bank: flotsam transport of terrestrial animals. Ecology, 112-117.

Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. BMC Evolutionary Biology **8**, 289.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution **27**, 570-580.

Hernández-Vera G, Mitrović M, Jović J*, et al.* (2010) Host‑associated genetic differentiation in a seed parasitic weevil *Rhinusa antirrhini* (Coleptera: Curculionidae) revealed by mitochondrial and nuclear sequence data. Molecular Ecology **19**, 2286-2300.

Holzapfel E, Harrell J (1968) Transoceanic dispersal studies of insects. Pacific Insects **10**, 115-153.

Huang JP, Lin CP (2010) Diversification in subtropical mountains: Phylogeography, Pleistocene demographic expansion, and evolution of polyphenic mandibles in Taiwanese stag beetle, *Lucanus formosanus*. Molecular Phylogenetics and

Evolution **57**, 1149-1161.

Incagnone G, Marrone F, Barone R, Robba L, Naselli-Flores L (2015) How do freshwater organisms cross the "dry ocean"? A review on passive dispersal and colonization processes with a special focus on temporary ponds. Hydrobiologia **750**, 103-123.

Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends in Ecology & Evolution **15**, 104-109.

Kano T (1929) Description of three new species of Curculionidae of the genus *Pachyrrhynchus* Germar from the island of Botel-Tobago. Entomological Society of Japan **3**, 237-238.

Kano T (1936) Some problems concerning the biogeography of Kôthshô (5). Geographical review of Japan **10**, 59-83.

Kuo J, Kanamoto Z, Iizumi H, Aioi K, Mukai H (2006) Seagrasses from the Nansei Islands, Southern Japanese Archipelago: species composition, distribution and biogeography. Marine Ecology **27**, 290-298.

Kurita K, Hikida T (2014) Divergence and long-distance overseas dispersals of island populations of the Ryukyu five-lined skink, *Plestiodon marginatus* (Scincidae: Squamata), in the Ryukyu Archipelago, Japan, as revealed by mitochondrial DNA phylogeography. Zoological Science **31**, 187-194.

Landis MJ, Matzke NJ, Moore BR, Huelsenbeck JP (2013) Bayesian analysis of biogeography when the number of areas is large. Systematic Biology **62**, 789-804.

Lanfear R, Calcott B, Ho SY, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution **29**, 1695-1701.

Lim HC, Sheldon FH (2011) Multilocus analysis of the evolutionary dynamics of rainforest bird populations in Southeast Asia. *Molecular Ecology* **20**, 3414-3438.

Liu S-YV, Kokita T, Dai C-F (2008) Population genetic structure of the neon damselfish (*Pomacentrus coelestis*) in the northwestern Pacific Ocean. Marine Biology **154**, 745-753.

Lomolino M, Riddle B, Whittaker R, Brown J (2010) Biogeography. Sinauer Associates, Inc, Sunderland, Massachusetts.

Losos JB, Ricklefs RE (2009) Adaptation and diversification on islands. Nature **457**, 830-836.

MacArthur RH, Wilson EO (1967) *The theory of island biogeography* Princeton University Press, Princeton, NJ.

Matzke NJ (2013a) BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts, CRAN: The Comprehensive R Archive Network, Vienna, Austria. http://cran.r-project.org/package=BioGeoBEARS.

Matzke NJ (2013b) Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Frontiers of Biogeography **5**, 242-248.

Matzke NJ (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Systematic Biology **63**, 951-970.

Mayr E (1954) Change of genetic environment and evolution. In: *Evolution as a process* (eds. Huxley J, Hardy AC, Ford EB), pp. 157-180. Allen and Unwin, London.

McKenna DD, Sequeira AS, Marvaldi AE, Farrell BD (2009) Temporal lags and overlap in the diversification of weevils and flowering plants. Proceedings of the National Academy of Sciences **106**, 7083-7088.

Moriyama EN, Gojobori T (1992) Rates of synonymous substitution and base

composition of nuclear genes in Drosophila. Genetics **130**, 855-864.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature **403**, 853-858.

Nakamura K, Kokubugata G, Rubite RR, Huang Jr C, Kono Y, Yang HA, Lopez-Feliciano A, Labuguen ML, Yokota M, Peng CI (2014) In situ glacial survival at the northern limit of tropical insular Asia by a lowland herb *Begonia fenicis* (Begoniaceae). Botanical Journal of the Linnean Society **174**, 305-25.

Nathan R (2006) Long-distance dispersal of plants. Science **313**, 786-788.

Nuytemans H (1998) Notes on Philippine birds: interesting records from northern Luzon and Batan Island. Forktail **14**, 37-40.

Oliveros C, Ota H, Crombie R, Brown R (2011) The Herpetofauna of the Babuyan group of islands, northern Philippines. Scientific Papers of the Natural History Museum of the University of Kansas **43**, 1-20.

Oliveros CH, Moyle RG (2010) Origin and diversification of Philippine bulbuls. Molecular Phylogenetics and Evolution **54**, 822-832.

Oshiro Y (1991) Studies on the *Pachyrhynchus infernalis* effecting *Mangifera indica* in Ishigaki Island. Okinawa Agriculture **26**, 19-23. (in Japanese)

Ota H (1991) Taxonomic status of *Mabuya multicarinata* (Gray, 1845)(Scincidae: Squamata: Reptilia) from Taiwan, with comments on the herpetofauna of Lanyu Island. Bulletin of the College of Sciences, University of Ryukyus **51**, 11-18.

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. Researches on Population Ecology **40**, 189-204.

Ota H, Huang W-S (2000) *Mabuya cumingi* (Reptilia: Scincidae): an addition to the herpetofauna of Lanyu Island, Taiwan. Current Herpetology **19**, 57-61.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial

molecular clock: the mid-Aegean trench calibration. Molecular Biology and Evolution **27**, 1659-1672.

Parent CE, Caccone A, Petren K (2008) Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. Philosophical Transactions of the Royal Society B: Biological Sciences **363**, 3347-3361.

Peck SB (1994) Sea-surface (pleuston) transport of insects between islands in the Galápagos Archipelago, Ecuador. Annals of the Entomological Society of America **87**, 576-582.

Petren K, Grant P, Grant B, Keller L (2005) Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. Molecular Ecology **14**, 2943-2957.

Planas E, Ribera C (2014) Uncovering overlooked island diversity: colonization and diversification of the medically important spider genus *Loxosceles* (Arachnida: Sicariidae) on the Canary Islands. Journal of Biogeography **41**, 1255-1266.

Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6. *Available from http://beast.bio.ed.ac.uk/Tracer*.

Ree RH, Moore BR, Webb CO, Donoghue MJ (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution **59**, 2299-2311.

Ree RH, Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Systematic Biology **57**, 4-14.

Rogers DC (2014) Larger hatching fractions in avian dispersed anostracan eggs (Branchiopoda). Journal of Crustacean Biology **34**, 135-143.

Ronquist F (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Systematic Biology **46**, 195-203.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**, 1572-1574.

Ronquist F, Teslenko M, van der Mark P*, et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology **61**, 539-542.

Schmitz P, Cibois A, Landry B (2007) Molecular phylogeny and dating of an insular endemic moth radiation inferred from mitochondrial and nuclear genes: The genus *Galagete* (Lepidoptera: Autostichidae) of the Galapagos Islands. Molecular Phylogenetics and Evolution **45**, 180-192.

Schultze W (1923) A monograph of the pachyrrhynchid group of the Brachyderinae, Curculionidae: part 1. The Philippine Journal of Science **23**, 609-673.

Sequeira A, Lanteri A, Albelo LR, Bhattacharya S, Sijapati M (2008) Colonization history, ecological shifts and diversification in the evolution of endemic Galápagos weevils. Molecular Ecology **17**, 1089-1107.

Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**, 1246-1247.

Shiu H-J, Tokita K-i, Morishita E*, et al.* (2006) Route and site fidelity of two migratory raptors: Grey-faced Buzzards *Butastur indicus* and Honey-buzzards *Pernis apivorus*. Ornithological Science **5**, 151-156.

Siler CD, Oaks JR, Cobb K, Ota H, Brown RM (2014) Critically endangered island endemic or peripheral population of a widespread species? Conservation genetics of Kikuchi's gecko and the global challenge of protecting peripheral oceanic island endemic vertebrates. Diversity and distributions **20**, 756-772.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA.

Annual Review of Ecology, Evolution, and Systematics, 545-579.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22**, 2688-2690.

Starr CK, Wang HY (1992) Pachyrhynchine weevils (Coleoptera: Curculionidae) of the islands fringing Taiwan. Journal of Taiwan Museum **45**, 5-14.

Stommel, HM, Yoshida, K (1972) *Kuroshio; Its Physical Aspects*. University of Tokyo Press, Tokyo.

Su YC, Wang JF, Villanueva RJT, Nuñeza OM, Lin CP (2014) Hopping out of Mindanao: Miocene-Pliocene geological processes and cross-island dispersal as major drivers of diversity for Philippine treehoppers. Journal of Biogeography **41**, 1277-1290.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution **30**, 2725-2729.

Templeton AR (1980) The theory of speciation via the founder principle. Genetics **94**, 1011-1038.

Thiel M, Gutow L (2005) The ecology of rafting in the marine environment. II. The rafting organisms and community. Oceanography and Marine Biology: an annual review **43**, 279-418.

Tseng H-Y, Lin C-P, Hsu J-Y, Pike DA, Huang W-S (2014) The functional significance of aposematic signals: geographic variation in the responses of widespread lizard predators to colourful invertebrate prey. PloS one **9**, e91777.

Tsou C-H, Mori SA (2002) Seed coat anatomy and its relationship to seed dispersal in subfamily Lecythidoideae of the Lecythidaceae (the Brazil nut family). Botanical Bulletin of Academia Sinica **43**, 37-56.

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. Journal of Biogeography **27**, 1153-1167.

Weekers PHH, De Jonckheere JF, Dumont HJ (2001) Phylogenetic relationships inferred from ribosomal ITS sequences and biogeographic patterns in representatives of the genus *Calopteryx* (Insecta: Odonata) of the West Mediterranean and adjacent West European zone. Molecular Phylogenetics and Evolution **20**, 89-99.

Whittaker RJ, Fernández-Palacios JM (2007) *Island biogeography: ecology, evolution, and conservation* Oxford University Press, New York.

Whittaker RJ, Triantis KA, Ladle RJ (2008) A general dynamic theory of oceanic island biogeography. Journal of Biogeography **35**, 977-994.

Yang TF, Lee T, Chen CH*, et al.* (1996) A double island arc between Taiwan and Luzon: consequence of ridge subduction. Tectonophysics **258**, 85-101.

Yang Z (2015) The BPP program for species tree estimation and species delimitation. Current Zoology **61**, 854-865.

Yang Z, Rannala B (2014) Unguided species delimitation using DNA sequence data from multiple loci. Molecular Biology and Evolution **31**, 3125-3135.

Yoshitake H (2013) A New Genus and Two New Species of the Tribe Pachyrhynchini (Coleoptera: Curculionidae) from Palawan Island, the Philippines. Esakia **53**, 1-8.

Yu Y, Harris AJ, Blair C, He X (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. Molecular Phylogenetics and Evolution **87**, 46-49.

Table 1 Maximum likelihood values and statistics calculated in RAxML-VI-HPC and
CONSEL for alternative tree topologies of the stepping-stone hypothesis and the best
phylogenetic tree (Fig. 2).

Hypotheses	Log likelihood	AU	KH	SH	WKH	WSH
Stepping-stone	-12341.086	0	0	0	0	0
Best tree	-11815.909	1	1	1	1	1

AU=approximately unbiased test; KH= Kishino-Hasegawa test; SH= Shimodaira-
Hasegawa test; WKH= weighted Kishino-Hasegawa test; and WSH= weighted
Shimodaira-Hasegawa test.

Table 2. Model selection of ancestral area reconstruction in BioGeoBEARS.

	Log likelihood	Number of parameters	d	e	j	AIC
DEC	-45.863	2	0.053081	0.173657	0	95.72594
DEC+J	-42.7932	3	1.00E-12	0.145533	0.004753	91.58633
DIVALIKE	-37.6794	2	0.052299	1.00E-12	0	79.35875
DIVALIKE+J	-37.6159	3	0.042183	1.00E-12	0.000869	81.2317
BAYAREALIKE	-60.8371	2	0.061572	1.503288	0	125.6743
BAYAREALIKE+J	-56.7928	3	0.012976	1.439617	0.003612	119.5857

Figure 1. (a) Map of the Taiwan-Luzon volcanic belt and the representative species of *Pachyrhynchus orbifer* complex. (b) Predicted tree topology derived from the stepping-stone hypothesis by island distance. (c) Geological ages of islands of Taiwan-Luzon volcanic belt.
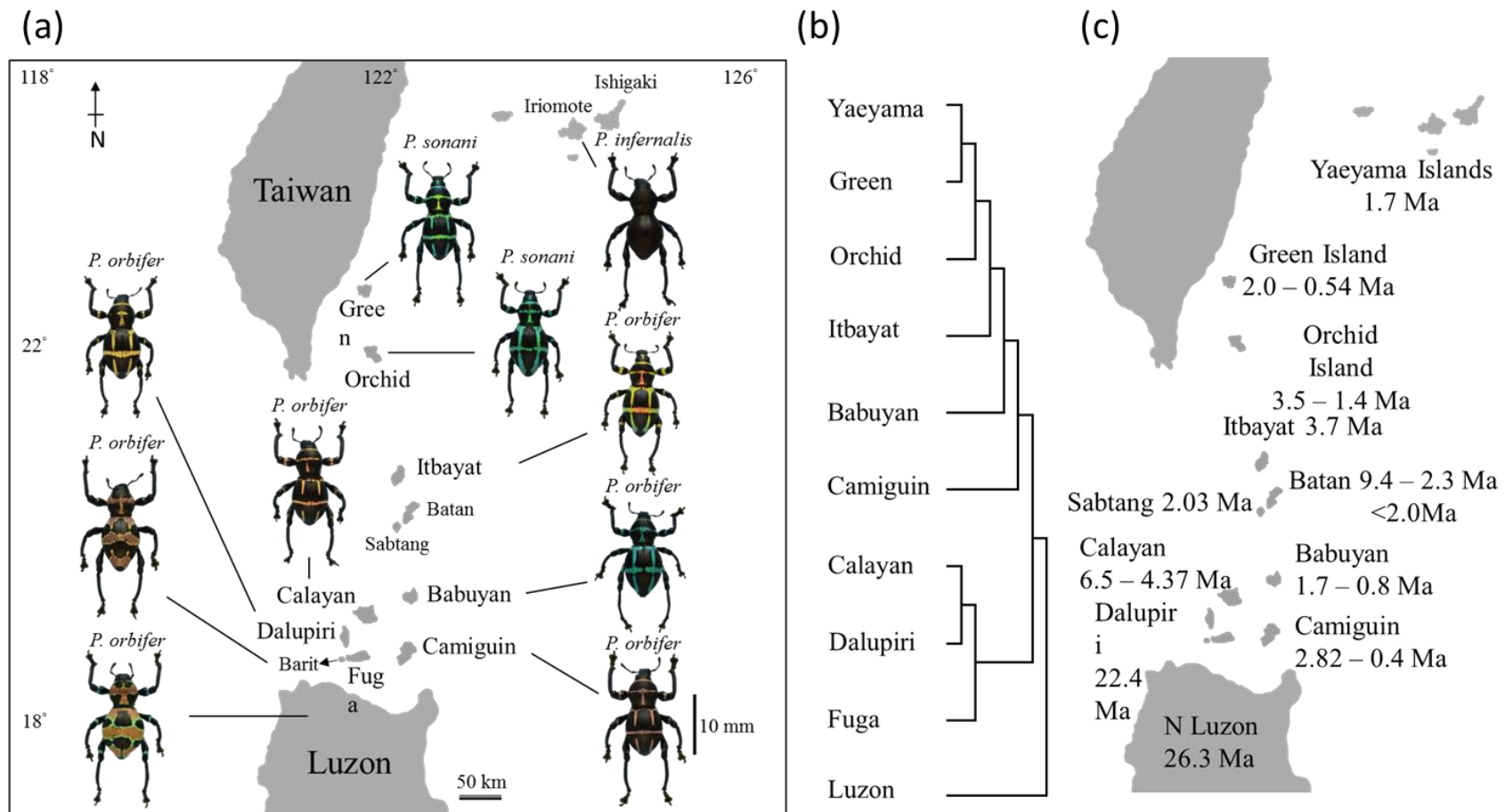
Figure 2. Phylogeny of *Pachyrhynchus orbifer* complex reconstructed from the combined data, using maximum likelihood (ML) analyses of 1,000 bootstrap replicates of the rapid hill-climbing algorithm in RAxML-VI-HPC, and Bayesian inference (BI) of 1×10^8 generations of Markov chain Monte Carlo (MCMC) processes in MrBayes v3.2.2. Numbers near the nodes are branch support values of likelihood bootstrap (LB)/ Bayesian posterior probability (BPP).
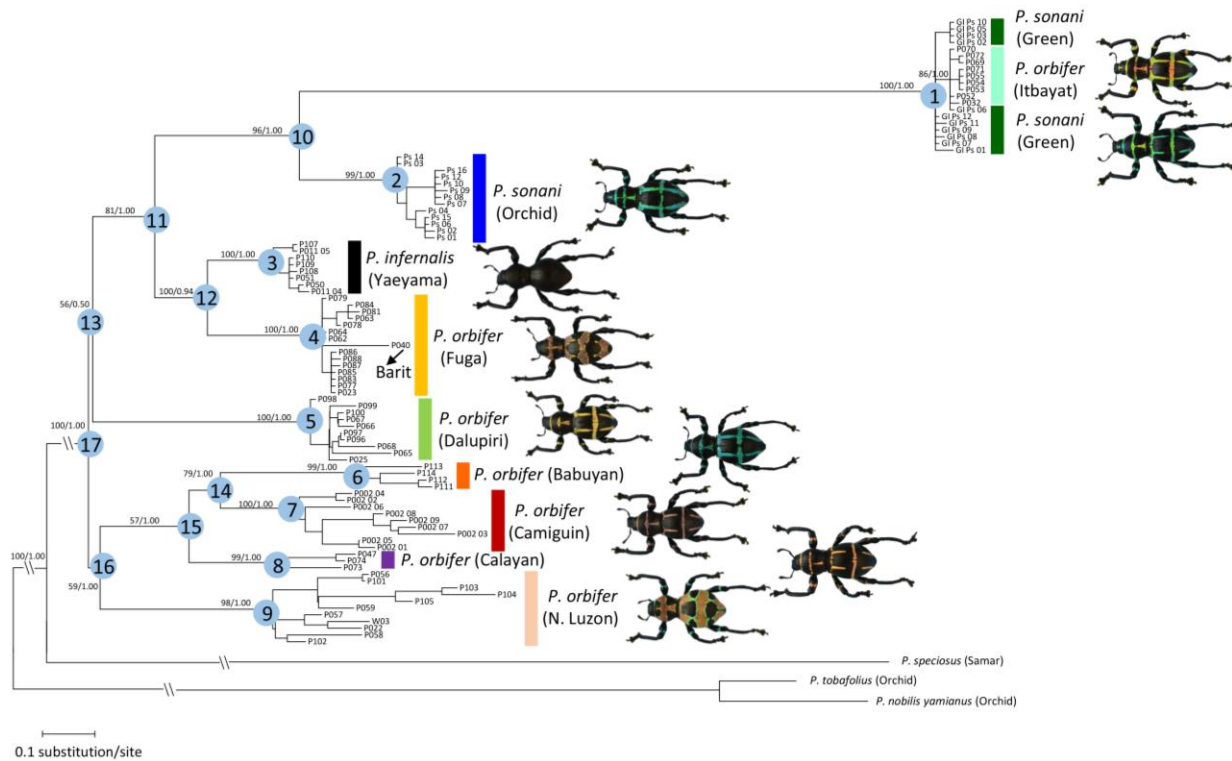
Figure 3. Divergence time, species delimitation and ancestral area of *Pachyrhynchus orbifer* complex. *Pachyrhynchus speciosus* and *Pachyrhynchus tobafolius* were used as outgroups. The pie charts specify the relative probability of ancestral area of nodes based on BBM model, with the arrows indicating the highest posterior probability of the reconstructed area. The color squares represent the most probable ancestral area estimated from the best-fit model DIVA-LIKE+J in BioGeoBEARS. An asterisk indicates the Bayesian posterior probability calculated in BP&P, where all members within a lineage belonging to one species is ≥ 0.99 (Appendix 7).
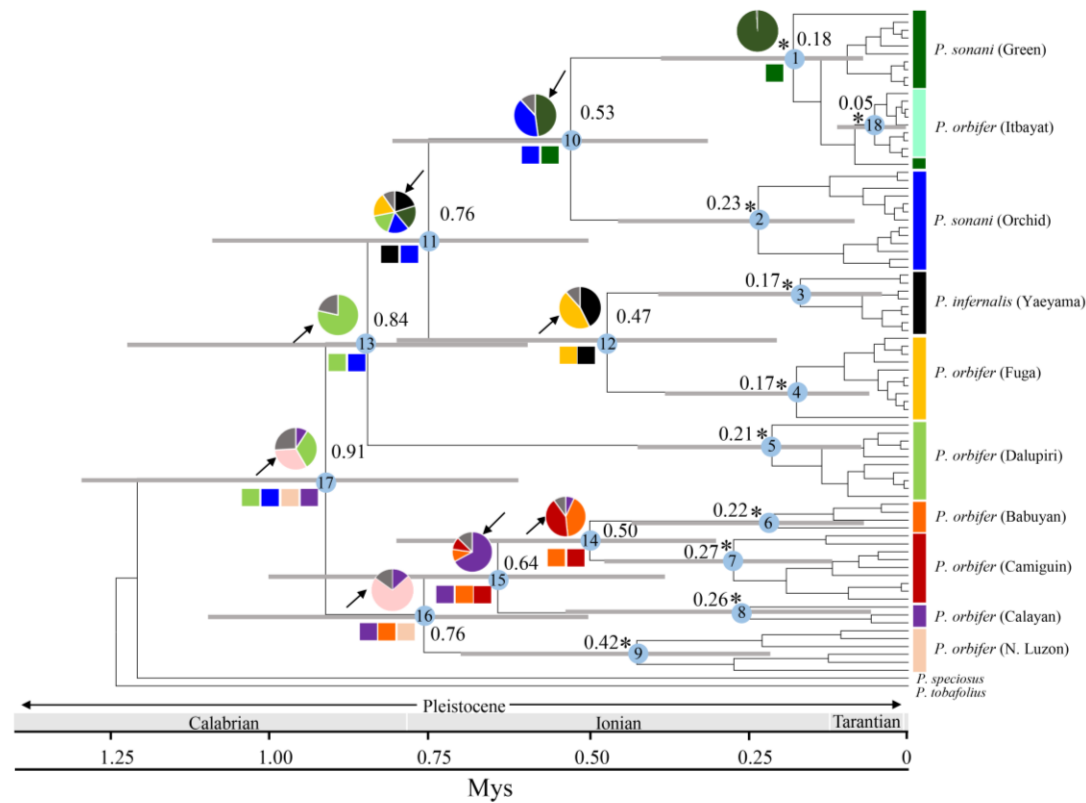
Figure 4. Inferred dispersal history of *Pachyrhynchus orbifer* complex across the Taiwan-Luzon volcanic belt. (a) 0.75-1.00 Mya: the first two colonization events from Luzon to Dalupiri (0.84 Mya) and Calayan (0.76 Mya), and then a long-distance dispersal from Dalupiri to Yaeyama (0.76 Mya); (b) 0.50-0.75 Mya: one southward dispersal from Calayan to the neighbouring island of Camiguin (0.64 Mya), and the other dispersal event from Yaeyama to Green Island (0.53 Mya); (c) 0.25-0.50 Mya: one northward dispersal from Camiguin to adjacent Babuyan Island (0.50 Mya), and one southward long-distance dispersal from Yaeyama to Fuga (0.47 Mya); and (d) 0-0.25 Mya: two southward dispersals including one dispersal from Green Island to Orchid Island (0.23 Mya), and the other from Green Island to Itbayat (0.05 Mya). Blue and red arrows represent northward and southward dispersal, respectively. Solid and open arrows indicate the two major dispersal routes.
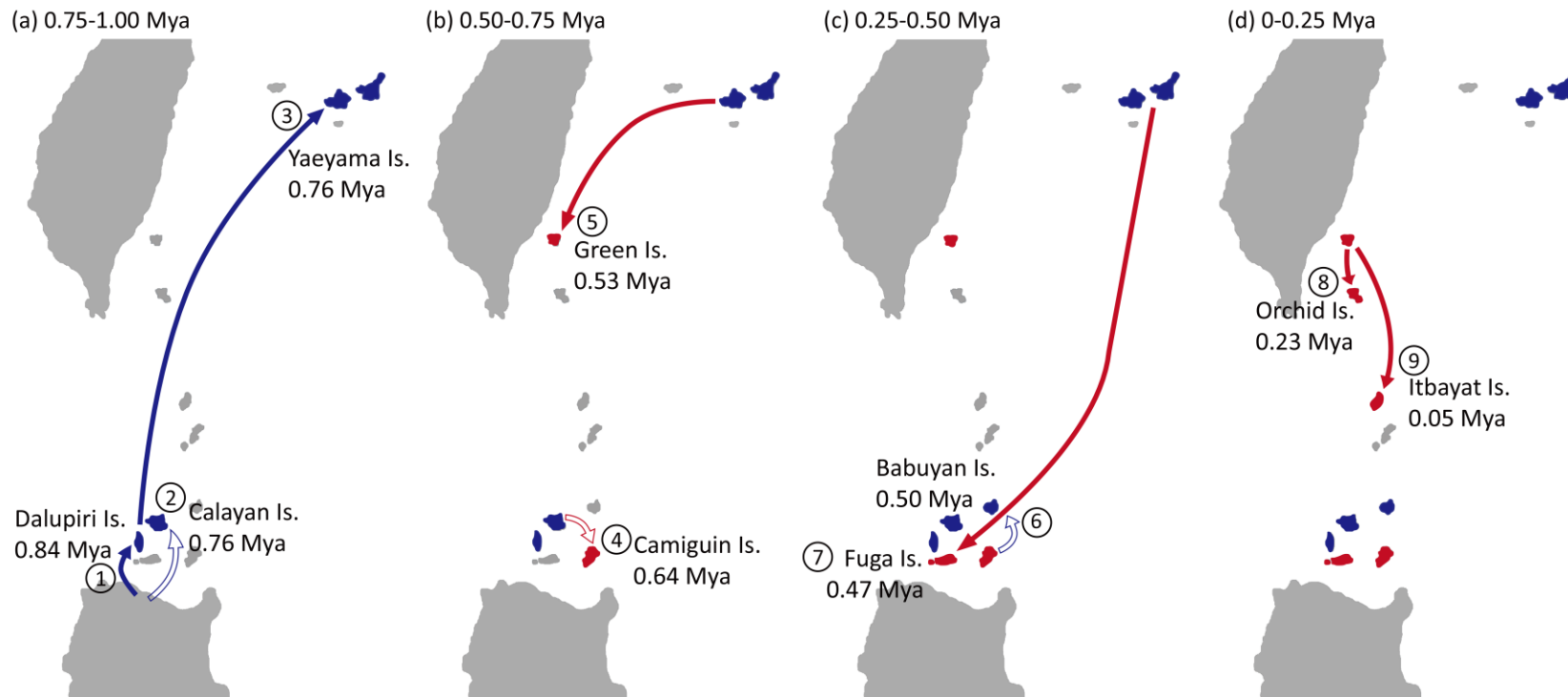
Supporting Information

Appendix 1. List of sample information.

Appendix 2. Primer sequences used in this study.

Appendix 3. Best-fit substitution model of each gene codons.

Appendix 4. Species tree inferred from four genes in Star BEAST (*BEAST).

Appendix 5. Phylogenetic trees of individual genes.

Appendix 6. Alternative colonization history of the *Pachyrhynchus orbifer* complex.

Appendix 7. Posterior probabilities of putative species delimited in BP&P.

Appendix 8. Historical demography of island lineages of the *Pachyrhynchus orbifer* complex.

Appendix 9. Habitat and host pant records of the *Pachyrhynchus orbifer* complex.

Appendix 1. List of sample collection sites, voucher numbers and GenBank accession numbers included in this study. Specimens were deposited in the National Museum of Natural Science in Taiwan.

Taxon	Country	Location	Sequence ID	Voucher	GenBank Accession Number			
					CO1	ND2	EF1-α	ITS
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_01	NMNS7618-1	KU953405	KU953501	KU953597	KU953693
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_02	NMNS7618-2	KU953406	KU953502	KU953598	KU953694
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_03	NMNS7618-3	KU953407	KU953503	KU953599	KU953695
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_05	NMNS7618-4	KU953408	KU953504	KU953600	KU953696
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_06	NMNS7618-5	KU953409	KU953505	KU953601	KU953697
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_07	NMNS7618-6	KU953410	KU953506	KU953602	KU953698
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_08	NMNS7618-7	KU953411	KU953507	KU953603	KU953699
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_09	NMNS7618-8	KU953412	KU953508	KU953604	KU953700
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_10	NMNS7618-9	KU953413	KU953509	KU953605	KU953701
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_11	NMNS7618-10	KU953414	KU953510	KU953606	KU953702
Pachyrhynchus sonani	Taiwan	Green Is.	GI_Ps_12	NMNS7618-11	KU953415	KU953511	KU953607	KU953703
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_01	NMNS7618-12	KU953416	KU953512	KU953608	KU953704
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_02	NMNS7618-13	KU953417	KU953513	KU953609	KU953705
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_03	NMNS7618-14	KU953418	KU953514	KU953610	KU953706
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_04	NMNS7618-15	KU953419	KU953515	KU953611	KU953707
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_05	NMNS7618-16	KU953420	KU953516	KU953612	KU953708
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_06	NMNS7618-17	KU953421	KU953517	KU953613	KU953709
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_07	NMNS7618-18	KU953422	KU953518	KU953614	KU953710

Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_08	NMNS7618-19	KU953423	KU953519	KU953615	KU953711
Pachyrhynchus orbifer	Philippines	Camiguin Is.	P002_09	NMNS7618-20	KU953424	KU953520	KU953616	KU953712
Pachyrhynchus infernalis	Japan	Ishigaki Is.	P011_04	NMNS7618-21	KU953425	KU953521	KU953617	KU953713
Pachyrhynchus infernalis	Japan	Iriomote Is.	P011_05	NMNS7618-22	KU953426	KU953522	KU953618	KU953714
Pachyrhynchus orbifer	Philippines	Claveria	P022	NMNS7618-23	KU953427	KU953523	KU953619	-
Pachyrhynchus orbifer	Philippines	Fuga Is.	P023	NMNS7618-24	KU953428	KU953524	KU953620	KU953715
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P025	NMNS7618-25	KU953429	KU953525	KU953621	KU953716
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P032	NMNS7618-26	KU953430	KU953526	KU953622	KU953717
Pachyrhynchus orbifer	Philippines	Barit Is.	P040	NMNS7618-27	KU953431	KU953527	KU953623	KU953718
Pachyrhynchus orbifer	Philippines	Calayan Is.	P047	NMNS7618-28	KU953432	KU953528	KU953624	KU953719
Pachyrhynchus infernalis	Japan	Ishigaki Is.	P050	NMNS7618-29	KU953433	KU953529	KU953625	KU953720
Pachyrhynchus infernalis	Japan	Ishigaki Is.	P051	NMNS7618-30	KU953434	KU953530	KU953626	KU953721
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P052	NMNS7618-31	KU953435	KU953531	KU953627	KU953722
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P053	NMNS7618-32	KU953436	KU953532	KU953628	KU953723
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P054	NMNS7618-33	KU953437	KU953533	KU953629	KU953724
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P055	NMNS7618-34	KU953438	KU953534	KU953630	KU953725
Pachyrhynchus orbifer	Philippines	Claveria	P056	NMNS7618-35	KU953439	KU953535	KU953631	KU953726
Pachyrhynchus orbifer	Philippines	Claveria	P057	NMNS7618-36	KU953440	KU953536	KU953632	KU953727
Pachyrhynchus orbifer	Philippines	Claveria	P058	NMNS7618-37	KU953441	KU953537	KU953633	KU953728
Pachyrhynchus orbifer	Philippines	Claveria	P059	NMNS7618-38	KU953442	KU953538	KU953634	-
Pachyrhynchus orbifer	Philippines	Fuga Is.	P062	NMNS7618-39	KU953443	KU953539	KU953635	KU953729
Pachyrhynchus orbifer	Philippines	Fuga Is.	P063	NMNS7618-40	KU953444	KU953540	KU953636	-
Pachyrhynchus orbifer	Philippines	Fuga Is.	P064	NMNS7618-41	KU953445	KU953541	KU953637	KU953730

Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P065	NMNS7618-42	KU953446	KU953542	KU953638	KU953731
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P066	NMNS7618-43	KU953447	KU953543	KU953639	KU953732
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P067	NMNS7618-44	KU953448	KU953544	KU953640	KU953733
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P068	NMNS7618-45	KU953449	KU953545	KU953641	KU953734
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P069	NMNS7618-46	KU953450	KU953546	KU953642	KU953735
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P070	NMNS7618-47	KU953451	KU953547	KU953643	KU953736
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P071	NMNS7618-48	KU953452	KU953548	KU953644	KU953737
Pachyrhynchus orbifer	Philippines	Itbayat Is.	P072	NMNS7618-49	KU953453	KU953549	KU953645	KU953738
Pachyrhynchus orbifer	Philippines	Calayan Is.	P073	NMNS7618-50	KU953454	KU953550	KU953646	KU953739
Pachyrhynchus orbifer	Philippines	Calayan Is.	P074	NMNS7618-51	KU953455	KU953551	KU953647	KU953740
Pachyrhynchus orbifer	Philippines	Fuga Is.	P077	NMNS7618-52	KU953456	KU953552	KU953648	KU953741
Pachyrhynchus orbifer	Philippines	Fuga Is.	P078	NMNS7618-53	KU953457	KU953553	KU953649	KU953742
Pachyrhynchus orbifer	Philippines	Fuga Is.	P079	NMNS7618-54	KU953458	KU953554	KU953650	KU953743
Pachyrhynchus orbifer	Philippines	Fuga Is.	P081	NMNS7618-55	KU953459	KU953555	KU953651	-
Pachyrhynchus orbifer	Philippines	Fuga Is.	P083	NMNS7618-56	KU953460	KU953556	KU953652	KU953744
Pachyrhynchus orbifer	Philippines	Fuga Is.	P084	NMNS7618-57	KU953461	KU953557	KU953653	-
Pachyrhynchus orbifer	Philippines	Fuga Is.	P085	NMNS7618-58	KU953462	KU953558	KU953654	KU953745
Pachyrhynchus orbifer	Philippines	Fuga Is.	P086	NMNS7618-59	KU953463	KU953559	KU953655	-
Pachyrhynchus orbifer	Philippines	Fuga Is.	P087	NMNS7618-60	KU953464	KU953560	KU953656	KU953746
Pachyrhynchus orbifer	Philippines	Fuga Is.	P088	NMNS7618-61	KU953465	KU953561	KU953657	KU953747
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P096	NMNS7618-62	KU953466	KU953562	KU953658	KU953748
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P097	NMNS7618-63	KU953467	KU953563	KU953659	KU953749
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P098	NMNS7618-64	KU953468	KU953564	KU953660	KU953750

Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P099	NMNS7618-65	KU953469	KU953565	KU953661	KU953751
Pachyrhynchus orbifer	Philippines	Dalupiri Is.	P100	NMNS7618-66	KU953470	KU953566	KU953662	KU953752
Pachyrhynchus orbifer	Philippines	Claveria	P101	NMNS7618-67	KU953471	KU953567	KU953663	-
Pachyrhynchus orbifer	Philippines	Claveria	P102	NMNS7618-68	KU953472	KU953568	KU953664	-
Pachyrhynchus orbifer	Philippines	Claveria	P103	NMNS7618-69	KU953473	KU953569	KU953665	KU953753
Pachyrhynchus orbifer	Philippines	Claveria	P104	NMNS7618-70	KU953474	KU953570	KU953666	-
Pachyrhynchus orbifer	Philippines	Claveria	P105	NMNS7618-71	KU953475	KU953571	KU953667	KU953754
Pachyrhynchus infernalis	Japan	Iriomote Is.	P107	NMNS7618-72	KU953476	KU953572	KU953668	KU953755
Pachyrhynchus infernalis	Japan	Iriomote Is.	P108	NMNS7618-73	KU953477	KU953573	KU953669	KU953756
Pachyrhynchus infernalis	Japan	Iriomote Is.	P109	NMNS7618-74	KU953478	KU953574	KU953670	KU953757
Pachyrhynchus infernalis	Japan	Iriomote Is.	P110	NMNS7618-75	KU953479	KU953575	KU953671	KU953758
Pachyrhynchus orbifer	Philippines	Babuyan Is.	P111	NMNS7618-76	KU953480	KU953576	KU953672	KU953773
Pachyrhynchus orbifer	Philippines	Babuyan Is.	P112	NMNS7618-77	KU953481	KU953577	KU953673	KU953774
Pachyrhynchus orbifer	Philippines	Babuyan Is.	P113	NMNS7618-78	KU953482	KU953578	KU953674	KU953775
Pachyrhynchus orbifer	Philippines	Babuyan Is.	P114	NMNS7618-79	KU953483	KU953579	KU953675	KU953776
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_01	NMNS7618-80	KU953484	KU953580	KU953676	KU953759
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_02	NMNS7618-81	KU953485	KU953581	KU953677	KU953760
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_03	NMNS7618-82	KU953486	KU953582	KU953678	KU953761
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_04	NMNS7618-83	KU953487	KU953583	KU953679	KU953762
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_06	NMNS7618-84	KU953488	KU953584	KU953680	KU953763
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_07	NMNS7618-85	KU953489	KU953585	KU953681	KU953764
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_08	NMNS7618-86	KU953490	KU953586	KU953682	KU953765
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_09	NMNS7618-87	KU953491	KU953587	KU953683	KU953766

Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_10	NMNS7618-88	KU953492	KU953588	KU953684	KU953767
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_12	NMNS7618-89	KU953493	KU953589	KU953685	KU953768
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_14	NMNS7618-90	KU953494	KU953590	KU953686	KU953769
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_15	NMNS7618-91	KU953495	KU953591	KU953687	KU953770
Pachyrhynchus sonani	Taiwan	Orchid Is.	Ps_16	NMNS7618-92	KU953496	KU953592	KU953688	KU953771
Pachyrhynchus orbifer	Philippines	Claveria	W03	NMNS7618-93	KU953497	KU953593	KU953689	KU953772
Pachyrhynchus speciosus	Philippines	Samar	P018	NMNS7618-94	KU953498	KU953594	KU953690	KU953777
Pachyrhynchus tobafolius	Taiwan	Orchid Is.	Pt_02	NMNS7618-95	KU953499	KU953595	KU953691	KU953778
Pachyrhynchus nobilis yamianus	Taiwan	Orchid Is.	Py_02	NMNS7618-96	KU953500	KU953596	KU953692	-

Appendix 2. Primer sequences used in this study

Gene	Name	Sequence	References
CO1	PI-CO1-F	5'-ACA TGA AGA GGT ATA GTA GGG ACT-3'	This study
CO1	PM-CO1-F	5'-AAA GAT ATT GGA ACY YTA TAT TTT-3'	This study
CO1	PI-CO1-R	5'-AAA TAG GGT CCC CAC CTC CTG CA-3'	This study
CO1	PM-CO1-R	5'-AAA TCA GAA TAA GTG TTG RTA TAA-3'	This study
CO1	LCO1490	5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3'	Folmer *et al.* 1994
CO1	HCO2198	5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3'	Folmer *et al.* 1994
ND2	TW-N1284	5'-ACA RCT TTG AAG GYT AWT AGT TT-3'	Simon *et al.* 2006
ND2	TM-J210	5'-AAT TAA GCT ACT RGG TTC ATA CCC-3'	Simon *et al.* 2006
ND2	P-ND2-F	5'-GAT TTA CTG CTT GAA TAG GAT TAG-3'	This study
ND2	P-ND2-R	5'-CAT AAT GAA AAT AGA TTT GTC ATG-3'	This study
EF1-a	EF1-Bf	5'-AGA ACG TGA ACG TGG TAT CA-3'	Hernández-Vera *et al.* 2010
EF1-a	EF1-Br	5'-CTT GGA GTC ACC AGC TAC ATA ACC-3'	Hernández-Vera *et al.* 2010
ITS	18Sr	5'-TAG AGG AAG TAA AAG TCG-3'	Weekers *et al.* 2001
ITS	28Sr	5'-GCT TAA ATT CAG CGG-3'	Weekers *et al.* 2001

Appendix 3. The best-fit substitution model of each gene codons.

gene		position	model
CO1		pos1	HKY+I+G
		pos2	HKY+I
		pos3	GTR+G
ND2		pos1	HKY+I
		pos2	HKY+I
		pos3	GTR+G
EF1a	exon	pos1	JC+I
		pos2	JC+I
		pos3	HKY+I
	intron		HKY+I+G
ITS	18S		HKY+I+G
	ITS1		HKY+I+G
	5.8S		JC+I
	ITS2		HKY+I+G
	28S		HKY+I+G

Appendix 4. Species tree of *Pachyrhynchus orbifer* complex inferred from two mitochondrial genes and two nuclear genes in Star BEAST (*BEAST). Numbers on the node are Bayesian posterior probability.
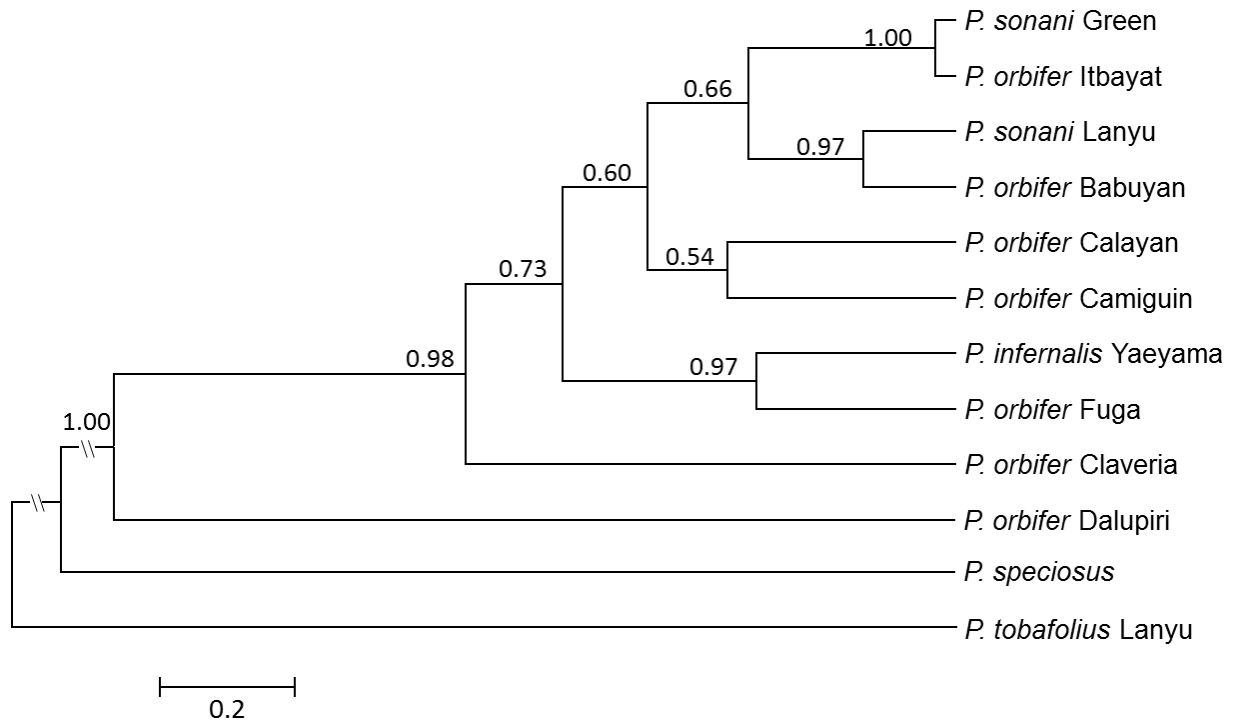
Appendix 5. Phylogenetic trees of the *Pachyrhynchus orbifer* complex reconstructed from (a) cytochrome c oxidase subunit 1, *CO1*; (b) NADH dehydrogenase subunit 2, *ND2*; (c) elongation factor 1-α, *EF1-α*; and (d) ribosomal internal transcribed spacer, *ITS*. Color bars represent different island populations.

(a) *CO1*

(b) *ND2*

(c) *EF1-α*
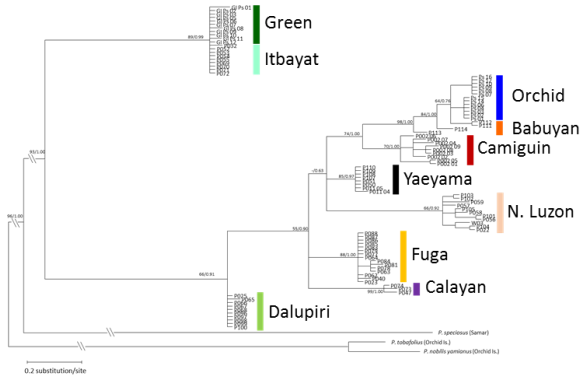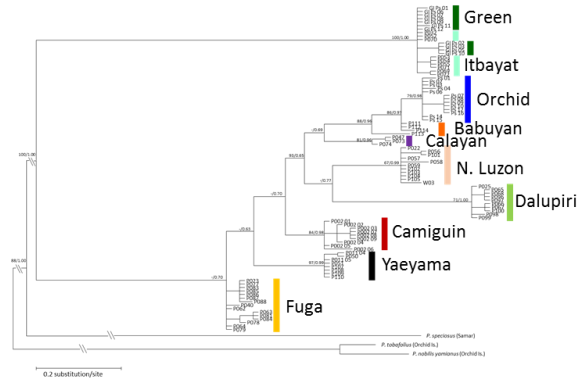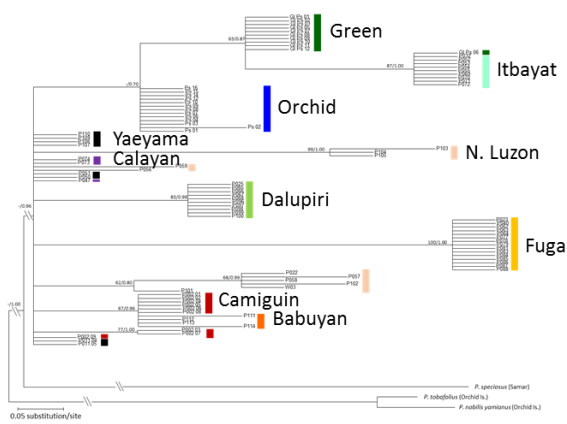
(d) *ITS*
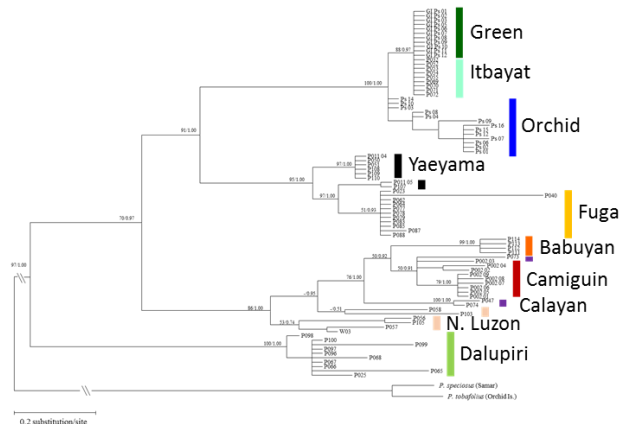
Appendix 6. Alternative colonization history (a-h) of the *Pachyrhynchus orbifer* complex in the Taiwan-Luzon volcanic belt. Numbers are the sequences of colonization events and blue and red arrows are northward and southward dispersals, separately.
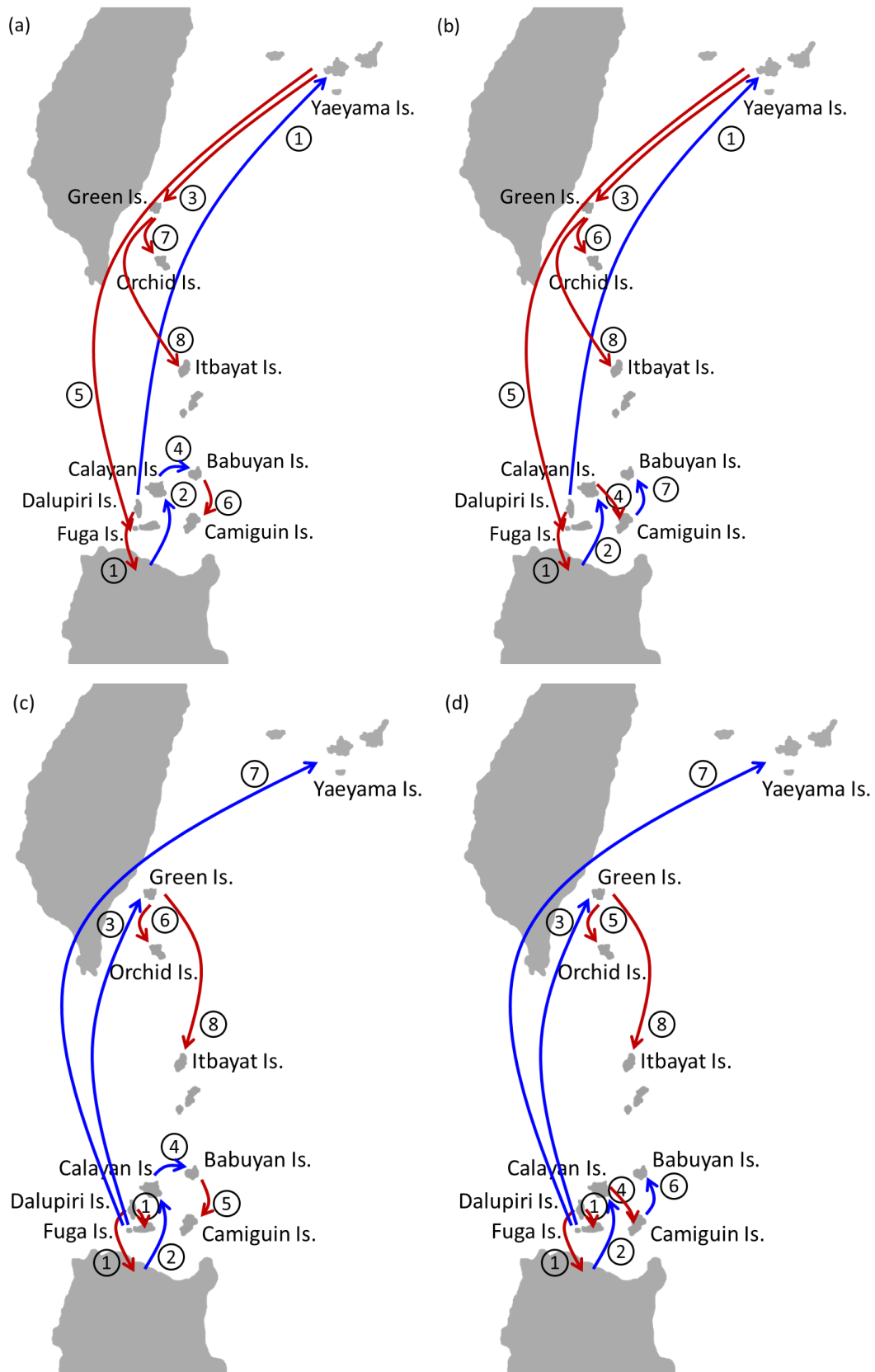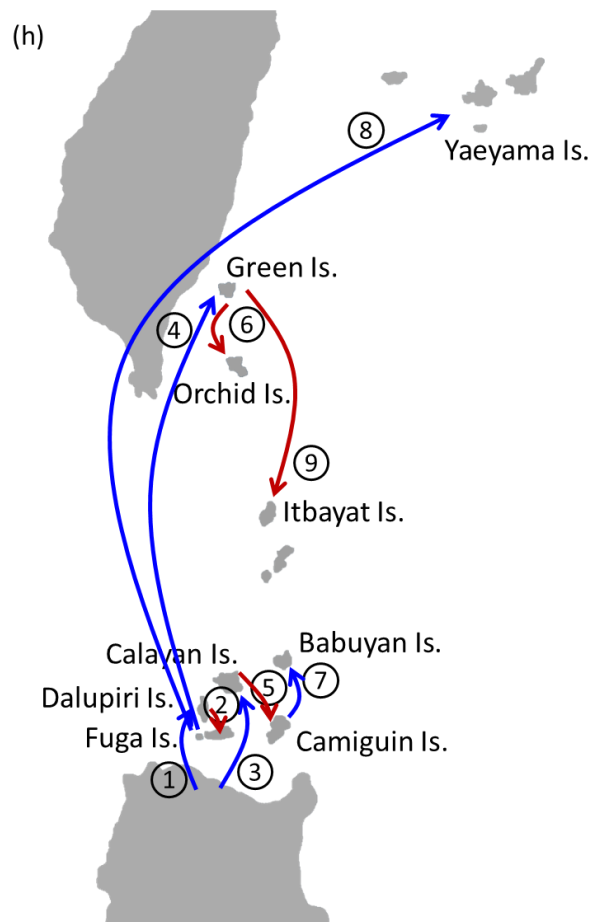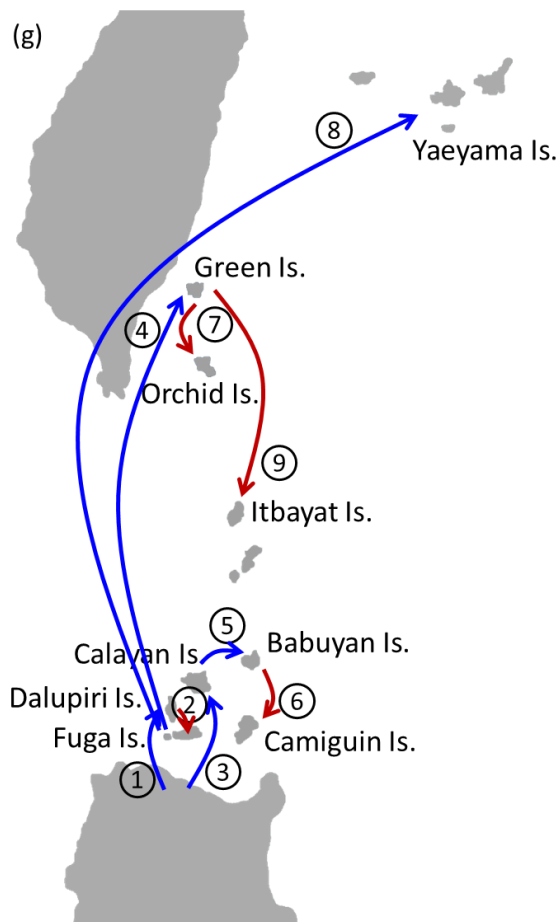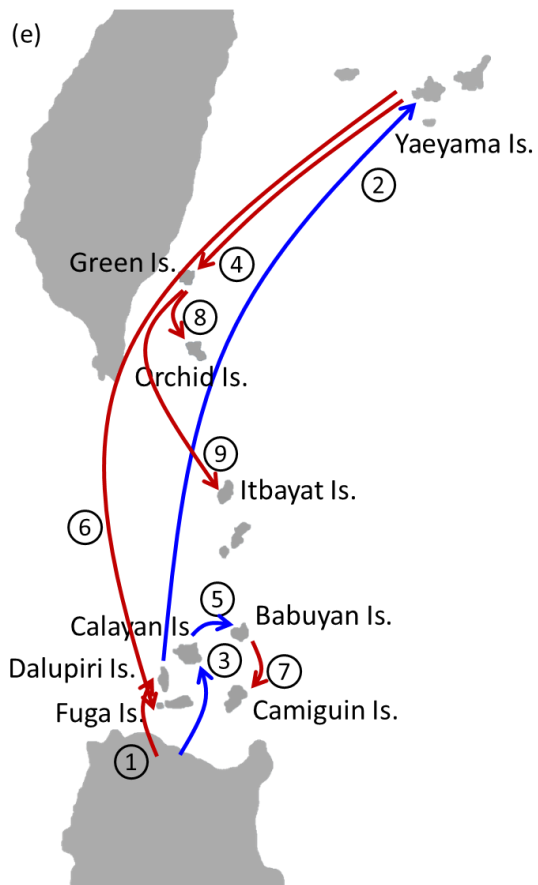
Appendix 7. Posterior probabilities of putative species delimited in BP&P.

Prior distribution	$\theta \sim G(1, 10)$ $\tau_0 \sim G(1, 10)$	$\theta \sim G(2, 2000)$ $\tau_0 \sim G(2, 2000)$	$\theta \sim G(2, 2000)$ $\tau_0 \sim G(1, 10)$
P. sonani Green	1.0	1.0	0.99
P. orbifer Itbayat	1.0	1.0	0.99
P. sonani Orchid	1.0	1.0	1.0
P. orbifer Babuyan	1.0	1.0	1.0
P. orbifer Calayan	1.0	1.0	1.0
P. orbifer Camiguin	1.0	1.0	1.0
P. infernalis Yaeyama	1.0	1.0	1.0
P. orbifer Fuga	1.0	1.0	1.0
P. orbifer Claveria	1.0	1.0	1.0
P. orbifer Dalupiri	1.0	1.0	1.0

Appendix 8. Historical demography of island lineages of *Pachyrhynchus orbifer* complex. (a) The number of estimated population size change. Blue bars represent the range of 95% HPD; (b) Extended Bayesian Skyline Plots illustrate the population size change through time. X-axis represents time in million years and y-axis is population size (generation time*10^6).

(a) Estimated population size change

(b) Population demography through time

Yaeyama
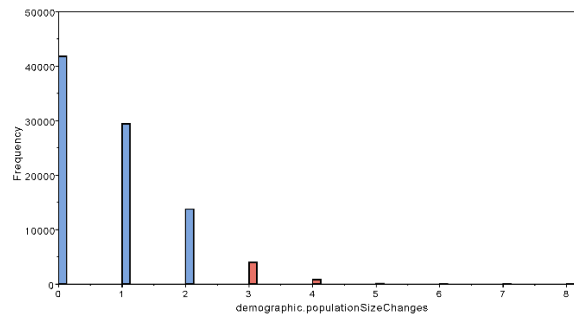
Yaeyama
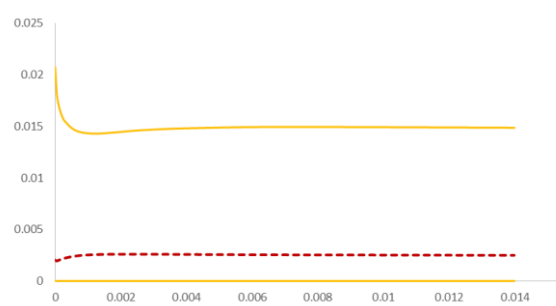
Green Island
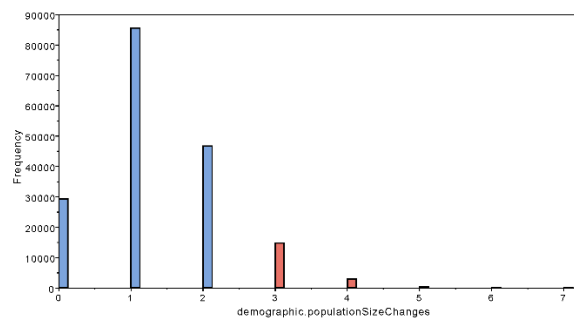
Green Island
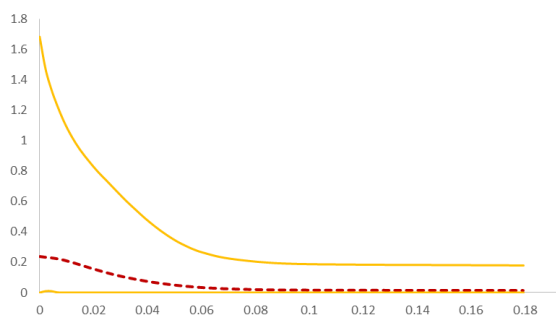
Orchid Island
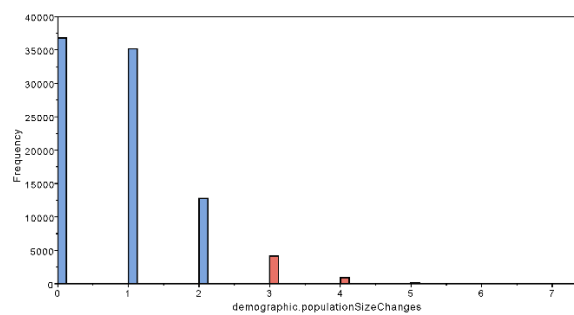
Orchid Island
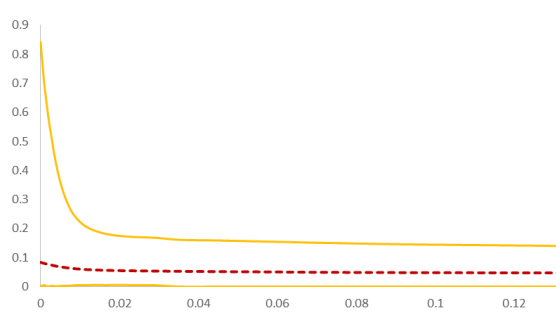
Itbayat
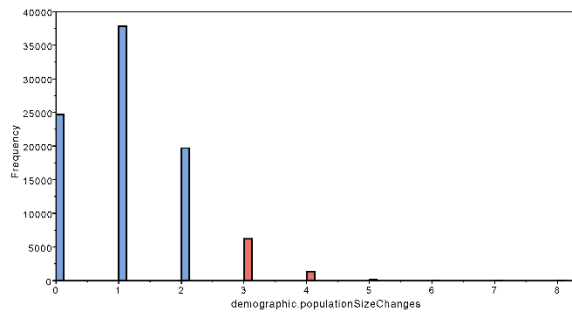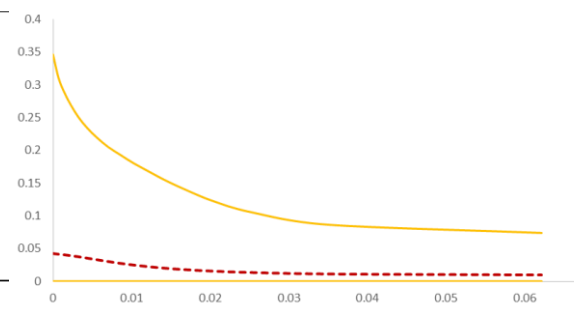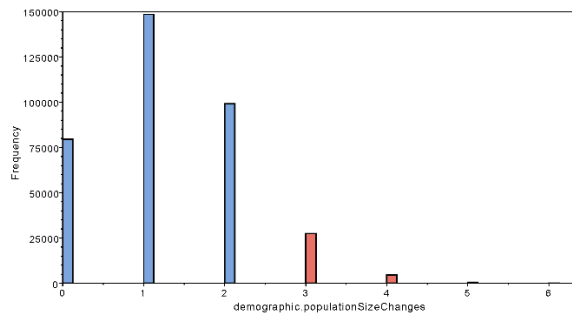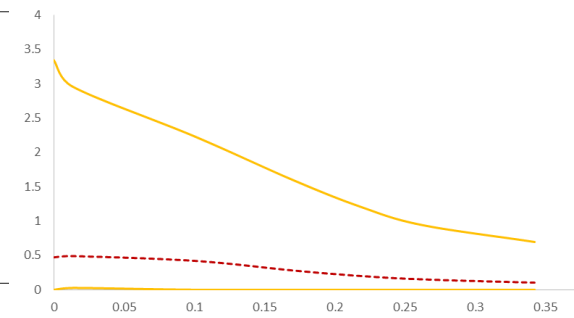
Babuyan

Calayan

Dalupiri

96
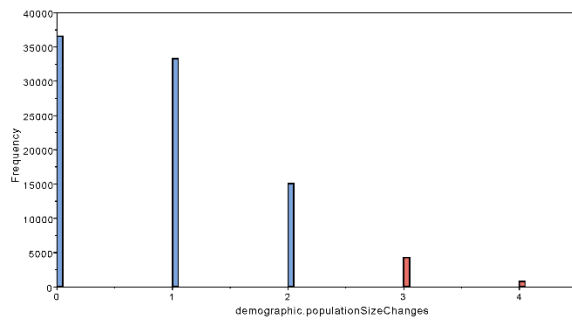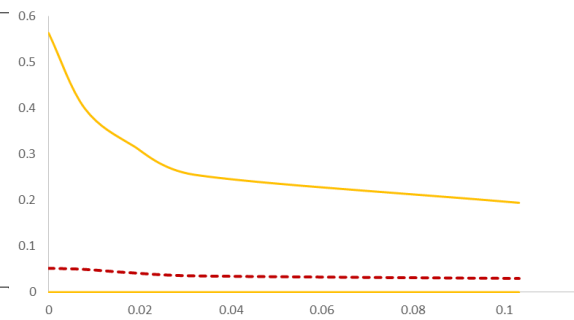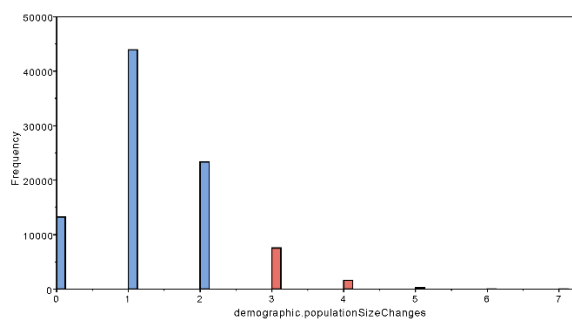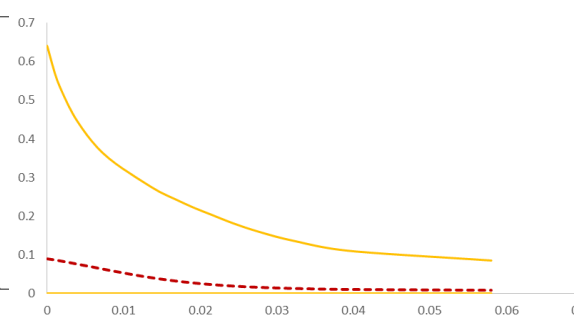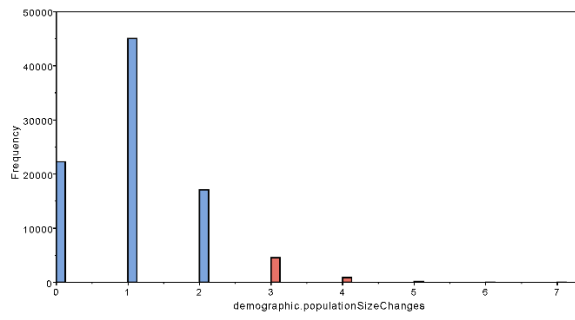
Camiguin

Camiguin
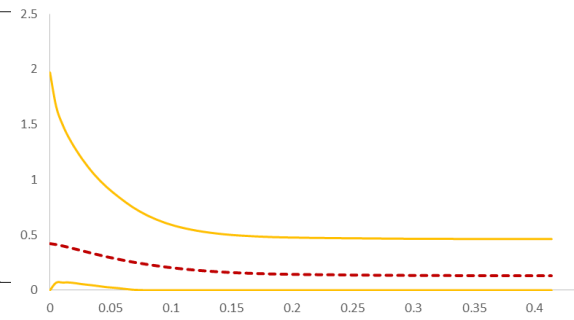
Fuga
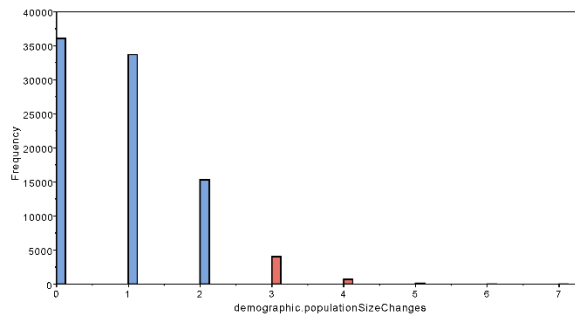
Fuga
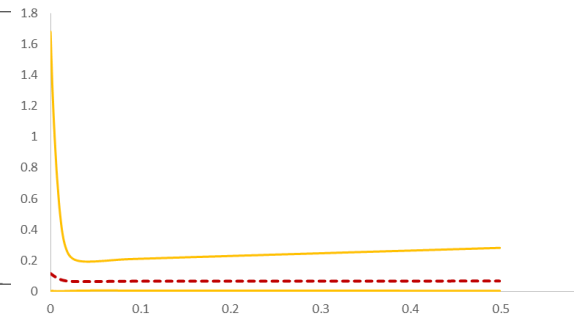
N. Luzon
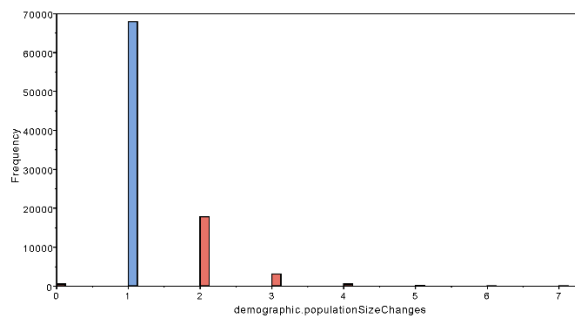
N. Luzon
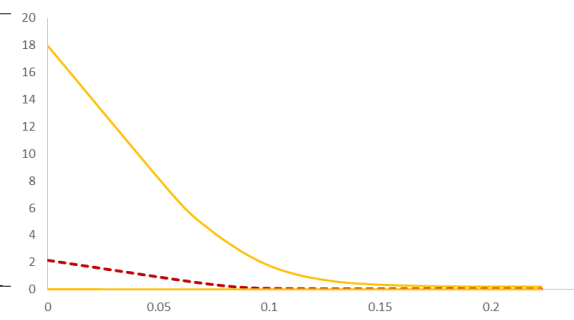
Appendix 9. Habitat and host plant records of the *Pachyrhynchus orbifer* complex.

Population	Habitat (sea shore/ inland)	Host plant	References
Yaeyama	seashore	Small-leaved barringtonia (*Barringtonia racemosa*, Lecythidaceae)	Personal observation
		Glochidion, Euphorbiaceae	Oshiro 1991
		Mango (*Mangifera indica*, Anacardiaceae)	Oshiro 1991
		Carambole (*Averrhoa carambola*, Oxalidaceae)	Oshiro 1991
Green	inland	Ceylon ardisia (*Ardisia elliptica*, Myrsinaceae)	Personal observation
		Elephant's ear (*Macaranga tanarius*, Euphorbiaceae)	Personal observation
		Iron-wood (*Casuarina equisetifolia*, Casuarinaceae)	Personal observation
		Melastoma (*Melastoma* prob. *Affine*, Melastomataceae)	Starr & Wang 1992
Orchid	sea shore, inland	Fish poison tree (*Barringtonia asiatica*, Lecythidaceae)	Personal observation
		Iron-wood (*Casuarina equisetifolia*, Casuarinaceae)	Personal observation
Itbayat	inland	Ceylon ardisia (*Ardisia elliptica*, Myrsinaceae)	Personal observation
Babuyan	?	?	
Camiguin	sea shore, inland	Fish poison tree (*Barringtonia asiatica*, Lecythidaceae)	Personal observation
Calayan	sea shore, inland	?	Personal observation
Dalupiri	sea shore	Elephant's ear (*Macaranga tanarius*, Euphorbiaceae)	Personal observation
Fuga	sea shore	?	Personal observation
Luzon	inland	Guava (*Psidium guajava*, Myrtaceae)	Personal observation

Chapter 4. The effect of data filtering in RAD sequences on the phylogenetic inference of *Pachyrhynchus orbifer* species complex

Abstract

Genome-wide single nucleotide polymorphisms (SNPs) obtained from Restriction Site-associated DNA sequencing (RAD-seq) method can provide a large numbers of genetic markers to resolve shallow phylogenetic relationships of closely related organisms. However, errors are prone to occur in large amount of SNPs and the sequence filtering process can influences the inference downstream analyses. RAD sequences of *Pachyrhynchus orbifer* species complex with different filtering parameter combinations in *Populations* of *Stacks* program were used to evaluate the stability of phylogenetic relationships, by comparing the number of retained loci, number of SNPs, heterozygosity, percentage of polymorphic site, bootstrap supports and tree topology. The parameters of the minimum percentage of individuals, minimum number of populations and a minimum depth of coverage for each individual significantly influenced the phylogenetic inference of the weevils. The problematic taxa from Ryukyu were detected because of the unusual pattern of heterozygosity and polymorphic sites. The most suitable parameters were identified after evaluating the number of retained taxa, heterozygosity, polymorphic sites, and ingroup monophyly among different settings. We recommend choosing the filtering parameters by testing different combination of parameters before downstream population genetic and phylogenetic analyses by identifying and removing the potentially erroneous samples.

Introduction

The remarkable advances in next-generation sequencing technology provide a powerful tool to study the functional expression of genomes (Morozova and Marra 2008), phylogenetics (Nadeau et al. 2013, Razkin et al. 2016), and population genomics (Veeramah and Hammer 2014, Kuhlwilm et al. 2016). These non-Sanger-based high-throughput DNA sequencing methods are extremely efficient for discovering thousands of markers cross the genome in non-model organisms (Davey et al. 2011). Employing a large number of genetic markers across the genome can build a more robust evolutionary history of populations or species (Miller et al. 2007, Baird et al. 2008). One of these approaches, Restriction Site-associated DNA sequencing (RAD-seq), is a genotyping method that allows subsampling of a genome, identifying homologous loci, and discovering genome-wide single nucleotide polymorphism (SNP) (Baird et al. 2008). This approach is especially powerful to reconstruct the relationships among recently divergent populations and estimate the phylogenies, which were unresolved using only a limited number of loci (Emerson et al. 2010, Wagner et al. 2013, Cruaud et al. 2014, Hou et al. 2015).

Although RAD-seq can provide large genetic information, analytical errors from various sources are inevitable. Errors can arise from the technical (e. g. the quality of reagents and equipments) and human manipulations (e. g. sample mislabelling and pipetting errors) (Bonin et al. 2004), wet laboratory procedures (e. g. PCR errors and bias) (Bonin et al. 2004, Peterson et al. 2012), next-generation-sequencing (e. g. sequencing error) (Loman et al. 2012, Davey et al. 2013) to bioinformatic processing (e. g. variation in coverage and mapping using a reference genome) (Catchen et al. 2013b, Davey et al. 2013). In bioinformatic analyses, several softwares dealing with sequence filtering and assembly have been widely applied, where *Stacks* (Catchen et al. 2011, Catchen et al.

2013b) is the most popular and frequently used to process RAD sequences. *Stacks* constructs a modular pipeline that assembles a large number of short-read sequences from different individuals, and aligns them onto a reference genome or *de novo* assembly (Catchen et al. 2011, Catchen et al. 2013b). In *Stacks* program package, the first stage is to clean raw Illumina reads by discarding reads with low quality scores, and then de-multiplex the data. The next stage includes building loci, calling SNP and creating catalogues of loci. The last stage is analyzing population genetic statistics of individuals and exporting suitable data formats for other analytical programs (Catchen et al. 2013b).

The combination of distance-based filtering parameters in the pipeline of *Stacks* results in different assemblies of RAD sequences (Mastretta-Yanes et al. 2015). After cleaning raw reads, three parameters in *ustacks* and *cstacks* (-*m*: minimum numbers of identical, raw reads required to create a stack; -*M*: number of mismatches allowed between stacks; -*n*: number of mismatches allowed between loci when building the catalogue) are used to identify loci. SNP, locus and allele error rates varied with different parameter settings by examine the individual sample replicates (Mastretta-Yanes et al. 2015). Not only the settings in *ustacks*, a number of filtering parameters in *Populations* (-*r*: minimum percentage of individuals within a population; -*p*: a minimum number of populations; -*m*: a minimum depth of coverage for each individuals) can be used to process different reads, therefore the sensitivity of filtering results needs to be evaluated (Catchen et al. 2013b). However, most recent studies using *Stacks* (e.g. Catchen et al. 2013a, Combosch and Vollmer 2015, Papadopoulou and Knowles 2015, Wang et al. 2016) did not consider the potential influences on the following analyses of phylogenetic inference or population genetics.

In chapter 3, a few basal nodes of *P. orbifer* phylogeny remained poorly resolved

based on the four gene sequences. This chapter evaluated the stability of phylogenetic relationships of *P. orbifer* complex using different filtering parameters combinations of a RAD sequencing data set. One data set was constructed in which each individual was assigned to different populations according to their island of origin. In the other data set, each individual was treated as a single population regardless of the sampling localities. In order to compare the filtered data and the associated phylogenies resulted from different settings, we used several characteristics of the genomic data and phylogenetic trees as criteria to evaluate the filtering parameter combination: 1) number of retained loci and SNPs that used to reconstruct the phylogeny: higher r, p and m will decrease the amount of data, which means lose characters for phylogenetic construction; however, higher r, p and m may represent more individuals and populations sharing these loci and SNPs, and may reflect the "true" variations among individuals and populations; 2) observed heterozygosity of variant sites and percentage of polymorphic site in each population: higher heterozygosity and polymorphic sites may provide more informative sites to reconstruct species relationships; 3) higher bootstrap supports of phylogenetic trees: the bootstrap value of the node represent the relative stability of the grouping within a phylogenetic tree, and we compared the overall bootstrap values among trees of different combinations. Because the genetic similarity among individuals from the same locality are usually higher than that between different localities, we also compare tree topologies among different settings, and exclude the trees that the outgroup taxa mix with the ingroup taxa. We aim to choose the most reliable filtering parameters by comparing a set of characteristics, and this filtered data set can then be used for further analyses.

Materials and Methods

Taxon sampling

    Pachyrhynchus orbifer species complex were collected from 11 localities in Yaeyama Islands (Ishigaki and Iriomote), Green Island, Orchid Island, Batanes Islands (Itbayat), Babuyan Islands (Babuyan, Calayan, Fuga, Dalupiri, Camiguin) and northern Luzon (Claveria) (Fig. 1 in Chapter 3). *Pachyrhynchus infernalis* distributes in Yaeyama Islands, *P. orbifer* distributes in northern Luzon, Batanes, and Fuga Island (Schultze 1923, Kano 1936), and *P. sonani* and *P. jitanasaius* distributes in Orchid and Green Islands, respectively (Chen et al. 2017). *Pachyrhynchus tobafolius* from Orchid and Green Islands was used as outgroups according to previous phylogenetic relationship (Tseng et al. unpublished data).

DNA extraction and SNP dataset collection

    Weevil specimens from 11 populations were stored in a −20 °C freezer, and their genomic DNAs were extracted from the legs using Qiagen Blood and Tissue DNA Extraction Kit (Qiagen, Madison USA). DNA was quantified using Qubit^{TM} (Invitrogen) and then diluted to 5 ng/ul in 10 ul volume. The quantified DNA was digested with restriction enzyme *Nde1* (recognition sequence: 5'-CA^TATG-3'), followed by unique barcodes and illumine adaptors being ligated to the digested fragments, and then these fragments were amplified by PCR. Pooled DNA samples from 350-450 bp were purified and then sequenced on Illumina HiSeq 2500 sequencing platform (Illumina, San Diego, CA, USA).

    Raw sequence reads were processed using *Stacks* pipeline (version 1.4) (Catchen et al. 2011), and the reads were de-multiplexed and filtered using the *process_radtags.pl* script which only retained reads with Phred score > 10. The loci were built *de novo*,

merged from multiple individuals, and matched against individuals using *denovo_map.pl* script. The minimum stack depth (*m*) was set to 3, and both the maximum distances between stacks (*M*) and the number of mismatches allow between individuals (*n*) were set to 2 (Mastretta‑Yanes et al. 2015). SNP data were output to phylip formats in *Populations* program of Stacks.

Filtering parameters

After assigning each individual to its population, three filtering parameters of *Populations* (*-r*, *-p*, *-m*) were used to examine the sensitivity of phylogenetic inference. Only one parameter was changed at a time, while the other two were fixed (*r* =0.5, *p*=5, and *m*=5). If each individual was treated as a single population, only two parameters need to be considered (*-p*, *-m*), and one parameter varied with the other one being fixed to *p*=50 and *m*=5.

Phylogenetic reconstruction and demographic inference

Maximum likelihood method (ML) was used to reconstruct the phylogenetic relationship of *P. orbifer* species complex based on SNP data. The SNP dataset, which fixed within all numbers of a population, but vary between at least two populations. The SNP dataset was transformed into the phylip format and used in RAxML Version 8.2.9 (Stamatakis 2014) for inferring phylogenies. The GTRCAT model with an ascertainment bias correction was used in ML analyses under the suggestion of Leaché et al. (2015), with the branch support assessed by 1,000 bootstrap sampling of the rapid hill-climbing algorithm.

Results

A total of 148,861,687 sequences were obtained from 96 individuals. After excluding those with ambiguous barcodes (12,846,715), ambiguous RAD-Tags (3,774,743), and low quality reads (708,681), 131,531,548 reads were retained for further analyses. The number of retained reads per individual ranged between 26,318 (an individual from Green Island) and 7870,247 (an outgroup from Orchid Island) (Appendix 1).

Number of loci and SNPs

The number of retained loci and SNPs fixed in the populations decreased when parameter *r*, *p* and *m* increased. The number of loci decreased more rapidly than that of SNPs (Fig. 1 and Table 1). In the population dataset, if *r, p* and *m* exceeded 0.5, 5 and 8, number of SNPs decreased to be lower than 1000, respectively (Table 1).

All the SNPs of Ryukyu were ambiguous sites after *r* was larger than 0.7 and Green Island populations contained all ambiguous sites when *r* was 1.0 (Table 1). All SNPs were ambiguous sites in Babuyan population when *p* and *m* was 10 and 60, separately (Table 1). When *m* was larger than 30, SNPs of Ryukyu population are all uncertain sites (Table 1). In individual dataset, if *p* was larger than 80, only very few loci and SNPs were retained (less than 100). When *m* was larger than 13, SNPs of several individuals were lost (Table 1). If the data set included all populations in the analyses, values of *r* ranged from 0.1-0.6, *p* from 1-9, and *m* from 2-27 retained higher number of loci and SNPs than the other settings in the population dataset. In individual dataset, when *p* was set from 10-70 and *m* from 2-10, the datasets retained all sampling individuals.

Heterozygosity and percentage of polymorphic sites

Heterozygosity of variant sites and percentage of polymorphic sites in different populations showed a similar pattern except for Ryukyu population and the outgroup, *P. tobafolius*. If *r* was larger than 0.6, heterozygosity of Babuyan, Itbayat, Green Island and Orchid Island decreased (Fig. 2a). Heterozygosity of outgroup decreased dramatically when *p* was larger than 2 (Fig. 2c), and the taxa of Babuyan and Ryukyu became unstable when *m* was larger than 11 (Fig. 2e). The percentage of polymorphic site was the highest when *r*= 0.6 and *p*= 2, 3, while *m* was smaller than 11 would yielded relatively stable results (Fig. 2b, d, f).

Tree topology and bootstrap support

In population dataset, all ingroup populations formed monophyletic lineages when *r* was set to 0.1 to 0.5 and 0.8 (Appendix 2), *p* was set from 2 to 5 (Appendix 3), and *m* was set from 3 and 5 (Appendix 4). In the phylogenetic trees resulted from other parameter settings, either Ryukyu or Babuyan populations were grouped with the outgroup or their relationships not resolved. In the individual datasets, ingroup individuals formed monophyletic group when *p* is from 10 to 50, 70 and *m*= 2 and 5 (Appendix 5-6). The phylogenetic trees under other parameter settings showed that some individuals from Babuyan or Ryukyu mixed with the outgroups. Regardless population or individual datasets, two consistent monophyletic groups [(Itbayat+Green+Orchid) and (Calayan+ Dalupiri)] were observed on 37 out of 41 trees, except for one individual from Ryukyu (A6) which was sometimes mixed with the individuals from Itbayat, Green Island and Orchid Island (Appendix 2-6).

Under different filtering settings, bootstrap values of trees in individual dataset showed a tendency toward lower bootstrap support when parameter *p* and *m* increased. However, various settings showed different trends in population dataset. Although the

106

highest mean and median appeared when $r= 0.1, 0.2$, $p= 5, 7$ and $m= 3, 5$, but the distributions of bootstrap values did not show obvious differences among different combinations (Fig. 3).

Discussion

Validation of filtered sequences posits a challenge for RAD-seq analyses because different parameters settings in program *Stacks* recover different sets of loci (Catchen et al. 2013b, Mastretta-Yanes et al. 2015, Rodríguez-Ezpeleta et al. 2016). However, how to choose the most suitable sets of parameters in the filtering process remains unclear. This study showed that even processing the loci in the population level caused different levels of variation in the phylogenetic inferences. After comparing the amount of retained loci and SNPs (retained all populations and individuals), heterozygosity, polymorphic sites, and tree characteristics (topology and bootstrap support), we found that setting higher values of *r*, *p* and *m* might not always be suitable for the best inference of weevils' phylogeny. If we considered recovery of the monophyletic relationships of ingroup populations as the most critical criterion, we proposed that a combination of *r*=0.5, *p*=2 and *m*=5 is most suitable for phylogenetic inference of the weevil's population dataset (Table 2). On the other hand, a combination of *p*=70, and *m*=5 is suitable for individual dataset because this combination of parameter values yielded the highest number of species monophylies in the trees. In this case study, because the outgroups are morphologically distinct species from *P. orbifer* complex, and the two outgroups are individuals of the same species. *P. tobafolius*, using the ingroup monophyly as a criterion for choosing among parameter combinations was reasonable. However, if the populations or individuals are collected from the potential hybrid zoon or area of sympatry, this monophyly criterion may not be suitable (Mastretta-Yanes et al. 2015).

This study also found that most phylogenetic analyses found consistent relationships (e.g., the monophyly of Green Island, Orchid Island and Itbayat; and the monophyly of Calayan and Dalupiri) under different parameter settings, while the relationships of the other populations were not stable (e.g., populations from Camiguin, Fuga, or Claveria).

By comparing different parameter settings, we can detect some potentially problematic sequences of populations or individuals which did not form a presumed monophyly within the same islands or displayed unusual patterns of heterozygosity and polymorphic sites, as in the specimens from Ryukyu. The heterozygosity and polymorphic sites of Ryukyu display a fluctuated pattern, and the most suitable parameter was difficult to be identified. Confirming the depth of coverage of individuals from Ryukyu (5-34), three (A6, G5 and H5) of eight individuals have coverage lower than seven (Appendix 7). The lower ranges of loci coverage yield higher error rates, especially for *de novo*-assembled datasets (Fountain et al. 2016).

Applying RAD-seq is promising to resolve difficult phylogenetic relationships among taxa (Cariou et al. 2013, Cruaud et al. 2014), however, unsuitable filtering SNPs may cause false inference in the downstream analyses. In this study, different filtering parameters in *Populations* of *Stacks* affect the results of phylogenetic inferences of *P. orbifer* species complex.   Therefore, we recommended using the number of retained taxa, SNPs, heterozygosity, polymorphic sites, and tree topology as a set of criteria to evaluate different combinations of parameters, and to detect the lower coverage samples. After removing the identified problematic sampling taxa, it formed a more appropriate dataset for further phylogeographical analyses.

References

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS one **3**:e3376.

Bonin, A., E. Bellemain, P. Bronken Eidesen, F. Pompanon, C. Brochmann, and P. Taberlet. 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology **13**:3261-3273.

Cariou, M., L. Duret, and S. Charlat. 2013. Is RAD‑seq suitable for phylogenetic inference? An in silico assessment and optimization. Ecology and Evolution **3**:846-852.

Catchen, J., S. Bassham, T. Wilson, M. Currey, C. O'Brien, Q. Yeates, and W. A. Cresko. 2013a. The population structure and recent colonization history of Oregon threespine stickleback determined using restriction‑site associated DNA‑sequencing. Molecular Ecology **22**:2864-2883.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013b. Stacks: an analysis tool set for population genomics. Molecular Ecology **22**:3124-3140.

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

Combosch, D. J., and S. V. Vollmer. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific Pocillopora corals. Molecular Phylogenetics and Evolution **88**:154-162.

Cruaud, A., M. Gautier, M. Galan, J. Foucaud, L. Sauné, G. Genson, E. Dubois, S. Nidelet, T. Deuve, and J.-Y. Rasplus. 2014. Empirical assessment of RAD sequencing for interspecific phylogeny. Molecular Biology and Evolution **31**:1272-1274.

Davey, J. W., T. Cezard, P. Fuentes-Utrilla, C. Eland, K. Gharbi, and M. L. Blaxter. 2013.

Special features of RAD Sequencing data: implications for genotyping. Molecular Ecology **22**:3151-3164.

Davey, J. W., P. A. Hohenlohe, P. D. Etter, J. Q. Boone, J. M. Catchen, and M. L. Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics **12**:499-510.

Emerson, K. J., C. R. Merz, J. M. Catchen, P. A. Hohenlohe, W. A. Cresko, W. E. Bradshaw, and C. M. Holzapfel. 2010. Resolving postglacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences **107**:16196-16200.

Fountain, E. D., J. N. Pauli, B. N. Reid, P. J. Palsbøll, and M. Z. Peery. 2016. Finding the right coverage: the impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. Molecular Ecology Resources **16**:966-978.

Hou, Y., M. D. Nowak, V. Mirré, C. S. Bjorå, C. Brochmann, and M. Popp. 2015. Thousands of RAD-seq loci fully resolve the phylogeny of the highly disjunct arctic-alpine genus *Diapensia* (Diapensiaceae). PloS one **10**:e0140175.

Kano, T. 1936. Some problems concerning the biogeography of Kôthshô (5). Geographical review of Japan **10**:59-83.

Kuhlwilm, M., I. Gronau, M. J. Hubisz, C. de Filippo, J. Prado-Martinez, M. Kircher, Q. Fu, H. A. Burbano, C. Lalueza-Fox, and M. de La Rasilla. 2016. Ancient gene flow from early modern humans into Eastern Neanderthals. Nature **530**:429-433.

Leaché, A. D., B. L. Banbury, J. Felsenstein, A. N.-M. de Oca, and A. Stamatakis. 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Systematic Biology **64**:1032-1047.

Loman, N. J., R. V. Misra, T. J. Dallman, C. Constantinidou, S. E. Gharbia, J. Wain, and M. J. Pallen. 2012. Performance comparison of benchtop high-throughput sequencing platforms. Nature biotechnology **30**:434-439.

Mastretta-Yanes, A., N. Arrigo, N. Alvarez, T. Jorgensen, D. Piñero, and B. Emerson. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. Molecular Ecology Resources **15**:28-41.

Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, and E. A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome research **17**:240-248.

Morozova, O., and M. A. Marra. 2008. Applications of next-generation sequencing technologies in functional genomics. Genomics **92**:255-264.

Nadeau, N. J., S. H. Martin, K. M. Kozak, C. Salazar, K. K. Dasmahapatra, J. W. Davey, S. W. Baxter, M. L. Blaxter, J. Mallet, and C. D. Jiggins. 2013. Genome-wide patterns of divergence and gene flow across a butterfly radiation. Molecular Ecology **22**:814-826.

Papadopoulou, A., and L. L. Knowles. 2015. Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. Evolution **69**:1501-1517.

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

Razkin, O., G. Sonet, K. Breugelmans, M. J. Madeira, B. J. Gómez-Moliner, and T. Backeljau. 2016. Species limits, interspecific hybridization and phylogeny in the cryptic land snail complex *Pyramidula*: The power of RADseq data. Molecular Phylogenetics and Evolution **101**:267-278.

Rodríguez-Ezpeleta, N., I. R. Bradbury, I. Mendibil, P. Álvarez, U. Cotano, and X. Irigoien. 2016. Population structure of Atlantic mackerel inferred from RAD-seq-

derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection. Molecular Ecology Resources **16**:991-1001.

Schultze, W. 1923. A monograph of the pachyrrhynchid group of the Brachyderinae, Curculionidae: part 1. The Philippine Journal of Science **23**:609-673.

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

Veeramah, K. R., and M. F. Hammer. 2014. The impact of whole-genome sequencing on the reconstruction of human population history. Nature Reviews Genetics **15**:149-162.

Wagner, C. E., I. Keller, S. Wittwer, O. M. Selz, S. Mwaiko, L. Greuter, A. Sivasundar, and O. Seehausen. 2013. Genome - wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular Ecology **22**:787-798.

Wang, J., D.-X. Xue, B.-D. Zhang, Y.-L. Li, B.-J. Liu, and J.-X. Liu. 2016. Genome-wide SNP discovery, genotyping and their preliminary applications for population genetic inference in spotted sea bass (*Lateolabrax maculatus*). PloS one **11**:e0157809.

Table 1. Number of retained loci, SNPs, and all SNPs of population or individual contained ambiguous sites with varied parameters.

	r	p	m	Retained loci	SNP	Ambiguous sites
Population dataset	0.1	5	5	19874	5343	
	0.2	5	5	15063	3160	
	0.3	5	5	11582	2117	
	0.4	5	5	8784	1407	
	0.5	5	5	7599	1185	
	0.6	5	5	5796	919	
	0.7	5	5	4053	629	Ryukyu
	0.8	5	5	2322	371	Ryukyu
	0.9	5	5	587	82	Ryukyu
	1	5	5	231	35	GI and Ryukyu
	0.5	2	5	58080	4699	
	0.5	3	5	19448	3016	
	0.5	4	5	11267	1882	
	0.5	5	5	7599	1185	
	0.5	6	5	3990	631	
	0.5	7	5	2171	310	
	0.5	8	5	927	113	
	0.5	9	5	211	30	
	0.5	10	5	53	12	Babuyan
	0.5	5	2	9153	1479	
	0.5	5	3	9153	1479	
	0.5	5	4	8359	1358	
	0.5	5	5	7599	1185	
	0.5	5	6	6979	1114	
	0.5	5	7	6465	1055	
	0.5	5	8	6068	1007	
	0.5	5	9	5749	969	
	0.5	5	10	5485	935	
	0.5	5	11	5235	903	
	0.5	5	12	5016	866	

Table 1. (Continued)

	r	p	m	Retained loci	SNP	Ambiguous sites
Population dataset	0.5	5	15	4453	784	
	0.5	5	18	3975	710	
	0.5	5	21	3506	633	
	0.5	5	24	3083	574	
	0.5	5	27	2693	486	
	0.5	5	30	2346	421	Ryukyu
	0.5	5	40	1380	275	Ryukyu
	0.5	5	50	800	174	Ryukyu
	0.5	5	60	443	86	Babuyan, Ryukyu
Individual dataset	1	10	5	68835	59180	
	1	20	5	23114	33233	
	1	30	5	13835	20261	
	1	40	5	7489	12223	
	1	50	5	3774	6439	
	1	60	5	1770	2750	
	1	70	5	561	622	
	1	80	5	68	57	P113
	1	90	5	6	4	A1,A2,A6,G5,H5,P109,P110,P111,P112,P113,P114
	1	100	5	0	0	
	1	50	2	5003	8994	
	1	50	5	3774	6439	
	1	50	8	2746	4793	
	1	50	10	2327	4161	
	1	50	13	1813	3119	A6
	1	50	17	1355	2296	A2,A6
	1	50	20	1075	1726	A2,A6,G5,H5,p113
	1	50	23	843	1285	A1,A2,A6,G5,H5,P113
	1	50	27	642	922	A1,A2,A3,A6,G5,H5,P113
	1	50	30	530	719	A1,A2,A3,A6,F11,G5,H5,P113

Table 2. Compared the most suitable parameter in different criterions.

	Population dataset			Individual dataset	
	r	*p*	*m*	*p*	*m*
Number of retained loci/SNPs	-	-	-	-	-
Number of retained taxa	0.1-0.6	2-9	2-27	10-70	2-10
Heterozygosity	0.1-0.6	2	2-12		
% of polymorphic site	0.1-0.7	2-3	2-12		
ingroup monophyly	0.1-0.5, 0.8	2-5	3-5	10-50,70	2-5
bootstrap value (mean and median)	-	-	-	10	2
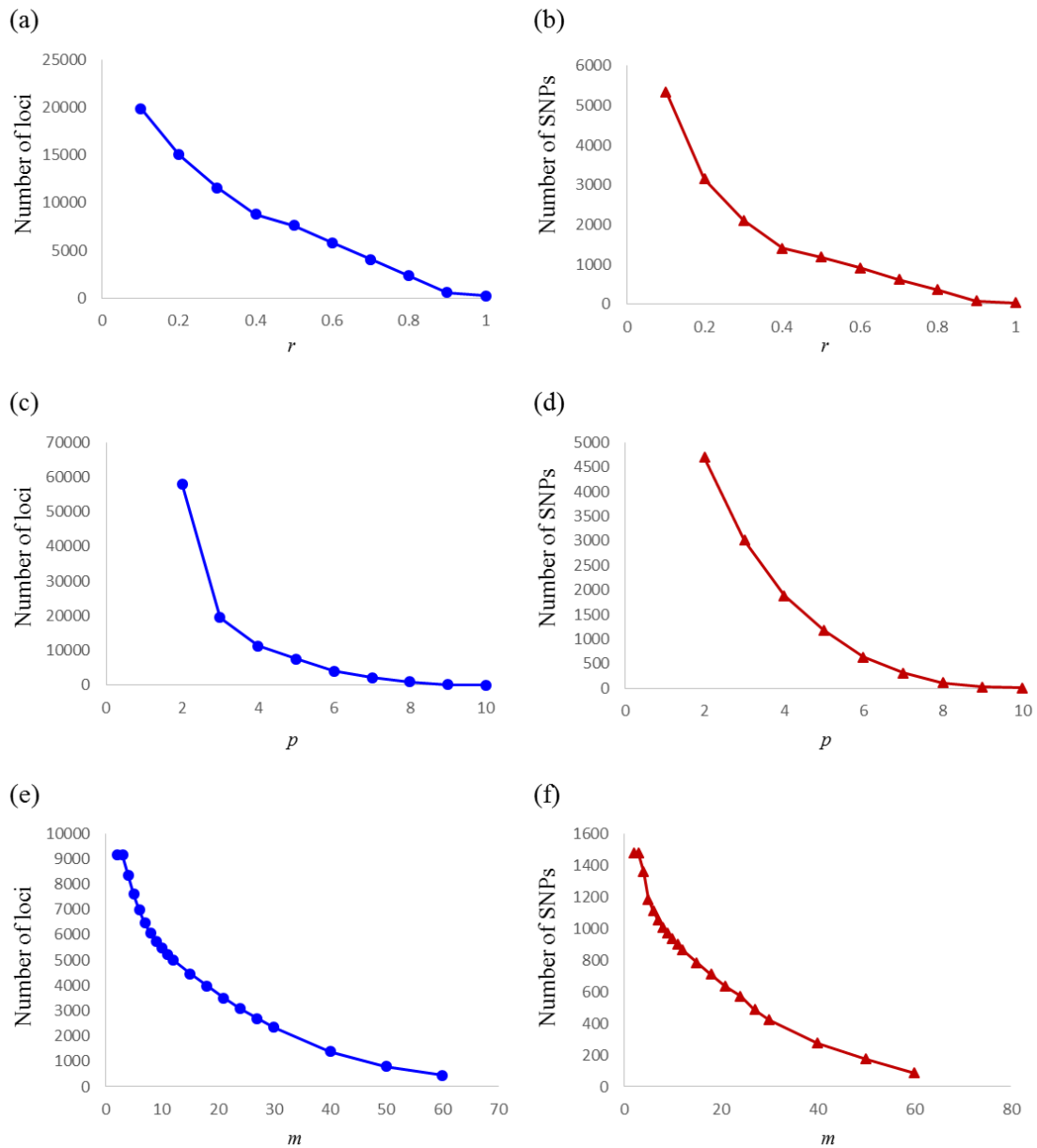
Figure 1. Effects of different filtering settings on (a, c, e, g, i) number of retained loci and (b, d, f, h, j) number of SNPs. (a)-(f) are population datasets, and (g)-(j) are individual datasets. (a, b) r varied from 0.1 to 1.0; (c, d) p varied from 2 to 10; (e, f) m varied from 2 to 60; (g, h) p varied from 10 to 100; (i, j) m varied from 2 to 30.

(g)

(h)

(i)

(j)
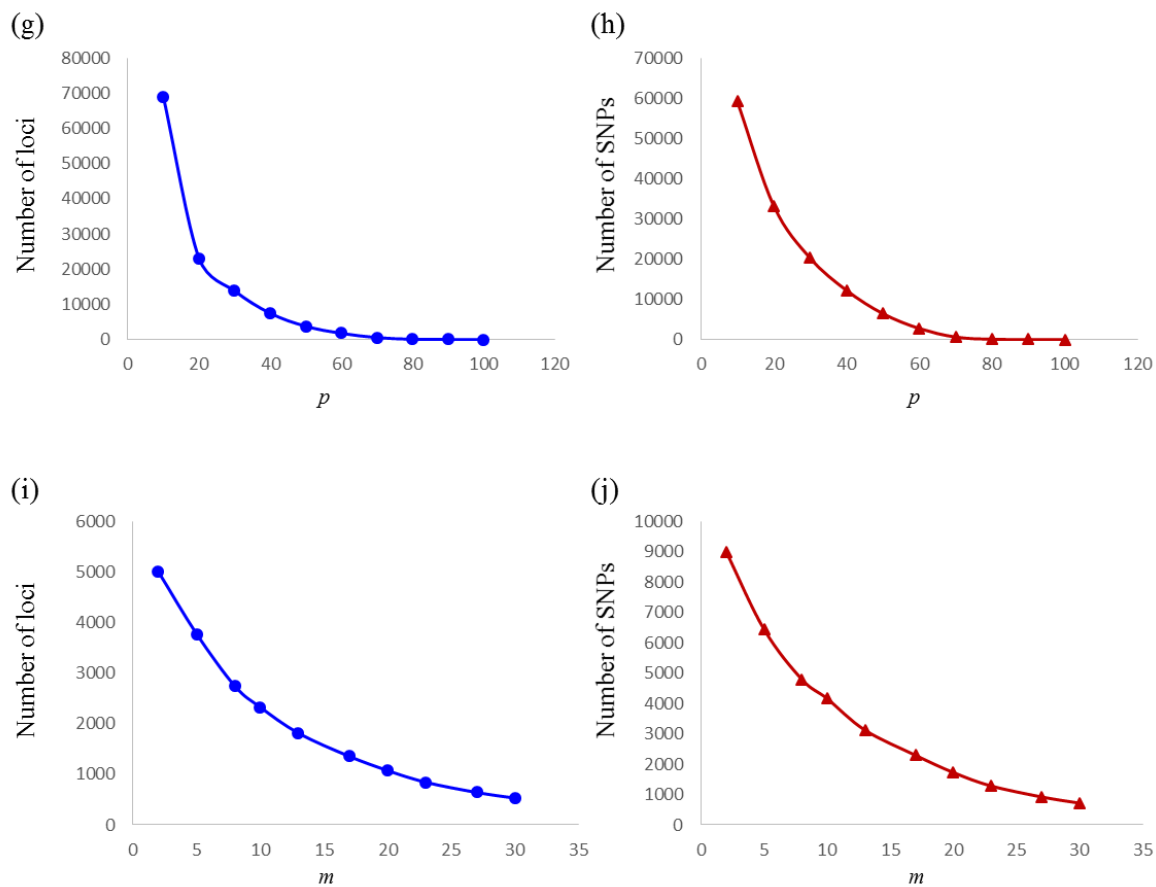
Figure 1. (Continued)

118
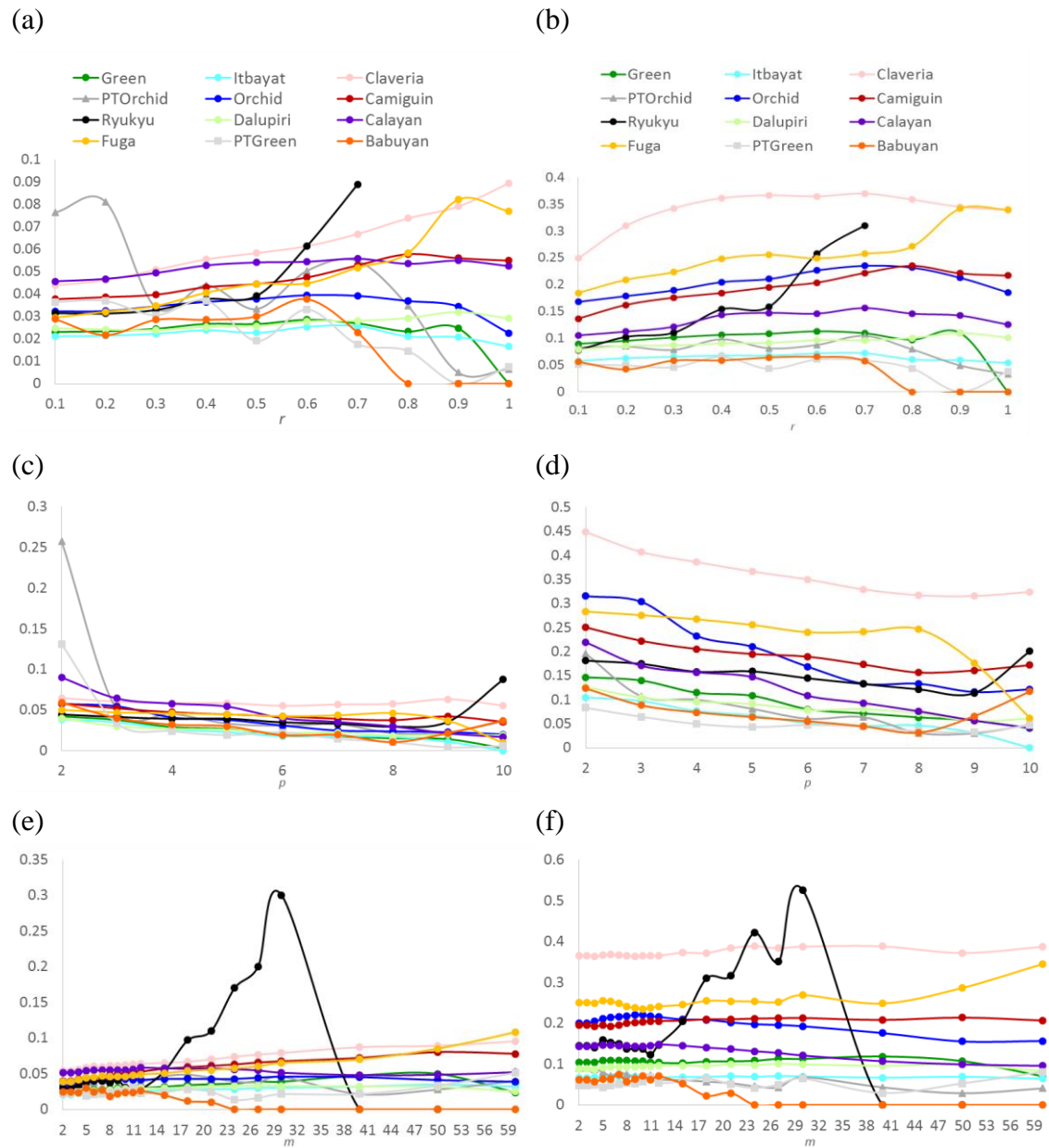
Figure 2. The effect of different parameter settings on average heterozygosity and percentage of polymorphic sites of each population. (a, c, e) Heterozygosity of different population with varied *r*, *p* and *m* on x-axis, respectively; (b, d, e) Percentage of polymorphic site of different population with varied *r*, *p* and *m* on x-axis, respectively. Colors represent different populations.
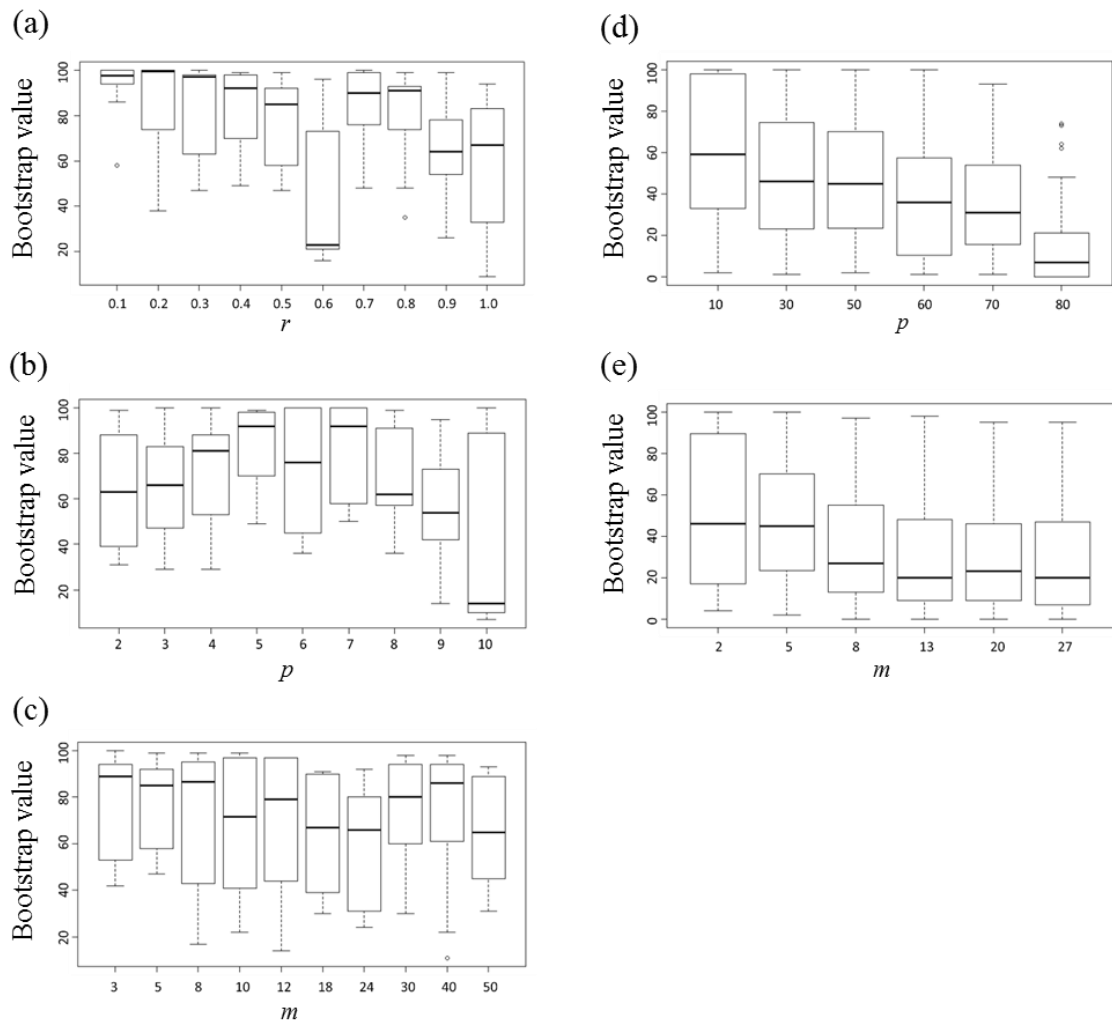
Figure 3. Bootstrap value distribution among phylogenetic trees with varied filtering parameters. (a), (b) and (c) represent different parameters *r*, *p* and *m* in population dataset, respectively and (d) and (e) show the distribution pattern under parameter *p* and *m* in individual dataset.

Supporting Information

Appendix 1. Summary of barcodes, dropped and retained sequences of RAD sequences of the specimens.

Appendix 2. Phylogenetic trees of the *P. orbifer* complex reconstructed from SNPs of population dataset filtered by different *r* setting.

Appendix 3. Phylogenetic trees of the *P. orbifer* complex reconstructed from SNPs of population dataset filtered by different *p* setting.

Appendix 4. Phylogenetic trees of the *P. orbifer* complex reconstructed from SNPs of population dataset filtered by different *m* setting.

Appendix 5. Phylogenetic trees of the *P. orbifer* complex reconstructed from SNPs of individual dataset filtered by different *p* setting.

Appendix 6. Phylogenetic trees of the *P. orbifer* complex reconstructed from SNPs of individual dataset filtered by different *m* setting.

Appendix 7. Loci coverage of individuals in different populations.

Appendix 1. Summary of barcodes, dropped and retained sequences of RAD sequences of the specimens.

Population	Barcode	Filename	Total sequences	Ambiguous RAD-Tag drops	Low Quality drops	Retained reads
Calayan	AAACGG	A8	2006096	27148	12086	1966862
Calayan	CTTGCG	G7	511533	7124	2534	501875
Calayan	TTCGAG	H7	4747329	31840	25157	4690332
Camiguin	TAACAG	A5	163874	4107	931	158836
Camiguin	GATTCG	B5	258813	6584	1486	250743
Camiguin	AACAGG	C5	2659985	22657	15514	2621814
Camiguin	GAGATG	D5	733202	6460	3766	722976
Camiguin	CAGGTG	E5	706302	9676	3279	693347
Camiguin	TTACGG	F5	1215197	18289	7199	1189709
Camiguin	CTTCGG	G4	2379982	16344	13429	2350209
Camiguin	ATCCAG	H4	2241692	37113	12685	2191894
Claveria	AAAGCG	A11	4791473	43562	27998	4719913
Claveria	GATGAG	B11	2225416	25164	11283	2188969
Claveria	AACCTG	C11	1179978	12821	7034	1160123
Claveria	CAGTGG	D11	1299260	8593	6837	1283830
Claveria	CTATTG	E11	1605730	11537	9155	1585038
Claveria	CTTATG	F10	1326150	23512	7171	1295467
Claveria	GTTACG	F11	151661	9591	739	141331
Claveria	ATCACG	G10	1457984	14760	8589	1434635
Claveria	GTGAAG	H10	5558997	33839	28666	5496492
Dalupiri	GAACCG	A7	1437029	108656	7479	1320894
Dalupiri	CATCAG	B7	880752	28720	4754	847278
Dalupiri	GACTTG	C7	1268310	24802	6885	1236623
Dalupiri	TAGAGG	D7	2614503	222636	12896	2378971
Dalupiri	GAGGCG	E6	488523	13286	2348	472889
Dalupiri	TTAATG	E7	1119042	86496	5700	1026846
Dalupiri	GTAGAG	F6	1185943	56204	5929	1123810
Dalupiri	ATAGTG	F7	1614209	118359	7754	1488096
Dalupiri	TTTGTG	G6	1991404	266396	8756	1716252
Dalupiri	GTCCGG	H6	2041265	75945	11046	1954274
Fuga	CAAGAG	A9	3646243	36365	19360	3590518
Fuga	GATCTG	B8	634587	10572	3222	620793
Fuga	AATCCG	B9	258022	7674	1405	248943

Population	Barcode	Filename	Total sequences	Ambiguous RAD-Tag drops	Low Quality drops	Retained reads
Fuga	TACTGG	C9	1560308	34804	8186	1517318
Fuga	GAGTAG	D8	2400559	21938	11277	2367344
Fuga	AAGTTG	D9	1423264	13391	6942	1402931
Fuga	CTAACG	E8	658460	7118	3953	647389
Fuga	GTAAGG	E9	2667124	225855	13171	2428098
Fuga	TTAGCG	F8	539343	42614	2543	494186
Fuga	TTTAAG	F9	3174834	331150	15428	2828256
Fuga	CTCAAG	G8	2780557	27770	15355	2737432
Fuga	CTCGTG	H8	1085904	11293	5577	1069034
Green	AAATAG	A1	89463	5736	686	83041
Green	TAATTG	A2	33531	7052	161	26318
Green	TATATG	B1	695303	262420	2381	430502
Green	CATACG	B2	211881	5593	1231	205057
Green	TATGCG	C1	466867	136652	1798	328417
Green	TACAAG	C2	1223559	22668	6712	1194179
Green	AACGAG	D1	1958022	14410	10508	1933104
Green	TACGTG	D2	240173	3911	1257	235005
Green	TAGCTG	E1	570851	5944	2968	561939
Green	CAGCCG	E2	120814	3365	600	116849
Green	ATATGG	F1	2230630	47531	12393	2170706
Green	CTACAG	F2	889239	13432	4799	871008
Green	GTTCAG	G1	798090	13073	4115	780902
Green	ATCTTG	H1	915470	10985	5006	899479
Itbayat	GAAGTG	A10	1457396	17783	6952	1432661
Itbayat	TATCGG	B10	1040858	135007	5016	900835
Itbayat	GACCAG	C10	1035623	15550	5582	1014491
Itbayat	TAGTCG	D10	114907	3628	623	110656
Itbayat	TTATAG	E10	1595783	40841	8711	1546231
Itbayat	GTCATG	G9	1876298	18748	9754	1847796
Itbayat	GTCGCG	H9	266550	4281	1561	260708
Orchid	CAATCG	A3	73899	4498	572	68829
Orchid	GAATGG	A4	270033	5844	1558	262631
Orchid	GATAGG	B3	412497	14724	2202	395571
Orchid	TATTAG	B4	825001	193735	3510	627756

Appendix 1. (Continued)

Population	Barcode	Filename	Total sequences	Ambiguous RAD-Tag drops	Low Quality drops	Retained reads
Orchid	CACATG	C3	488810	9977	2709	476124
Orchid	GACACG	C4	195425	3195	1072	191158
Orchid	CACGCG	D3	120698	2354	627	117717
Orchid	CAGAAG	D4	2330391	24773	12391	2293227
Orchid	GAGCGG	E3	1162378	13104	5797	1143477
Orchid	TAGGAG	E4	2252276	20619	11181	2220476
Orchid	GTACTG	F3	1729460	13607	9361	1706492
Orchid	ATACCG	F4	615837	10853	3736	601248
Orchid	ATTCTG	G2	1127055	13020	6003	1108032
Orchid	TTTCCG	G3	675111	95835	3039	576237
Orchid	TTCTCG	H2	967671	10572	5698	951401
Orchid	CTCTGG	H3	3032809	34605	17292	2980912
PTGreen	AAGCAG	D12	999881	7298	4541	988042
PTGreen	GTATCG	E12	787506	3287	4118	780101
PTGreen	ATTAGG	F12	2075730	36420	9923	2029387
PTGreen	GTCTAG	G12	1425987	20584	6503	1398900
PTGreen	TTGACG	H12	742191	6456	3424	732311
PTOrchid	AATAAG	A12	8076648	166691	39710	7870247
PTOrchid	AATGTG	B12	2353258	19826	10533	2322899
PTOrchid	CACCGG	C12	2005385	13955	9843	1981587
PTOrchid	TTCAGG	G11	4468721	19248	21488	4427985
PTOrchid	ATGATG	H11	3167344	16237	14774	3136333
Ryukyu	CAACTG	A6	34925	1661	296	32968
Ryukyu	AATTGG	B6	684558	17226	4126	663206
Ryukyu	CACTAG	C6	1179487	14027	7190	1158270
Ryukyu	AAGACG	D6	551291	4761	3267	543263
Ryukyu	ATTGAG	G5	55076	11417	291	43368
Ryukyu	TTCCTG	H5	70540	7708	340	62492
Total			136014972	3774743	708681	131531548

Appendix 2. Phylogenetic trees of the *Pachyrhynchus orbifer* complex reconstructed from SNPs filtered by different *r* setting. (a) to (j) are individuals assigned to collecting localities and parameter *r* from 0.1 to 1.0.

(a) *r*=0.1
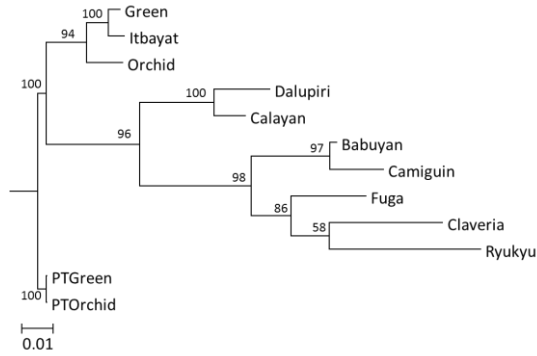
(b) *r*=0.2
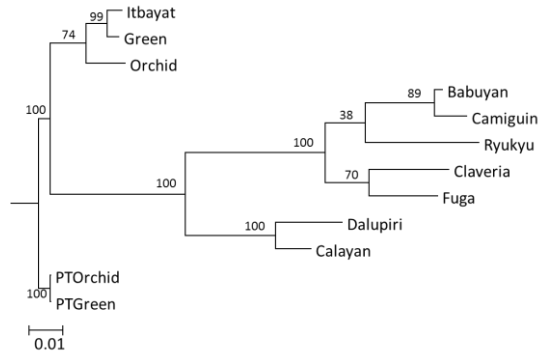
(c) *r*=0.3
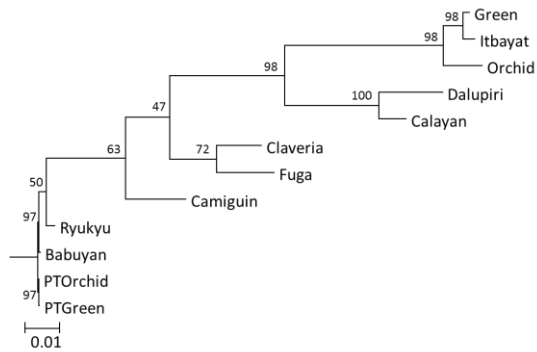
(d) *r*=0.4
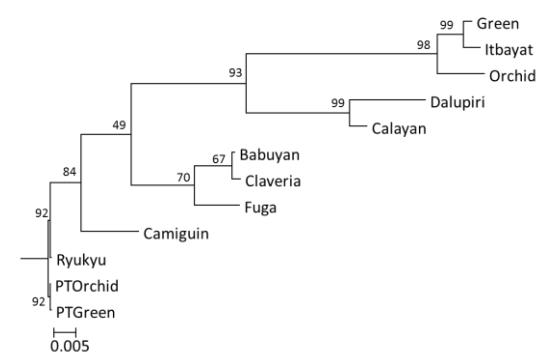
(e) *r*=0.5
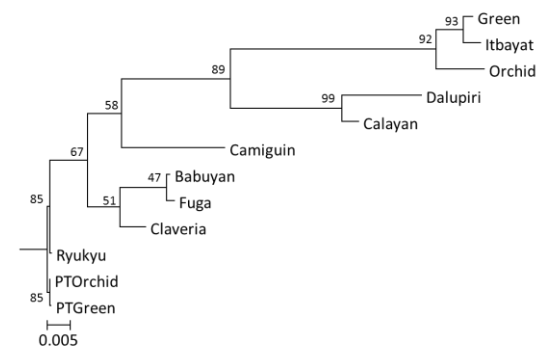
(f) *r*=0.6
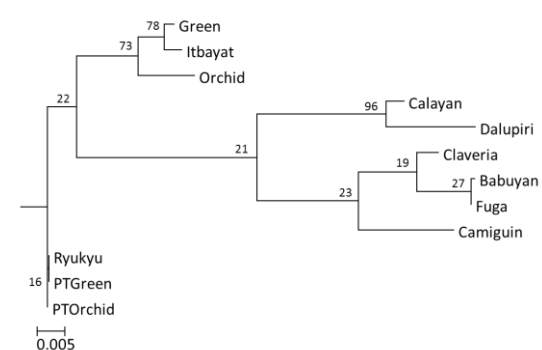
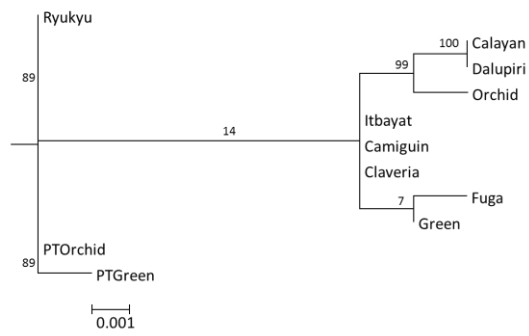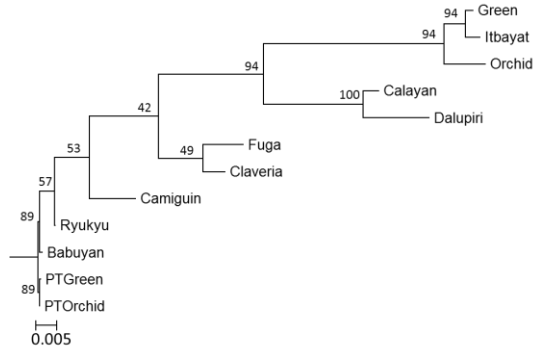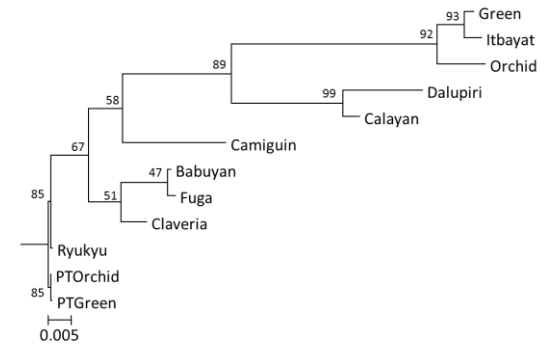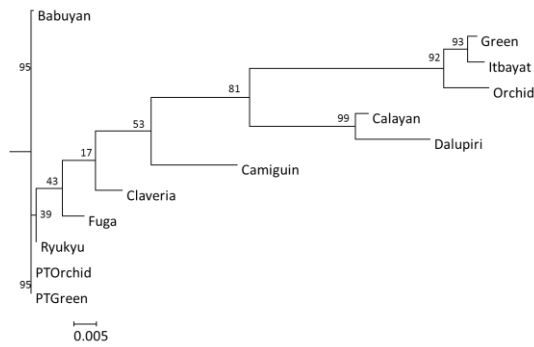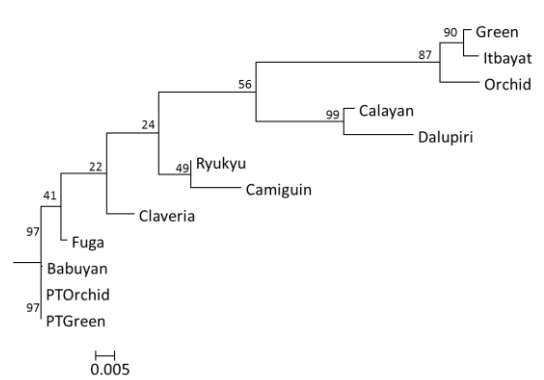Appendix 2. (Continued)

(g) *r*=0.7
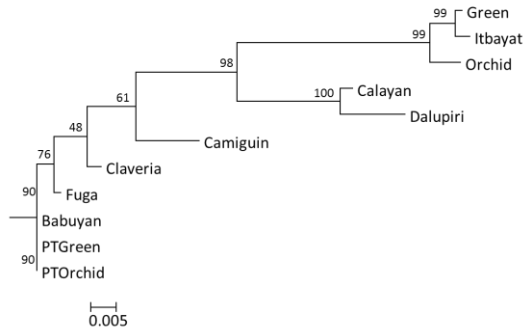
(h) *r*=0.8
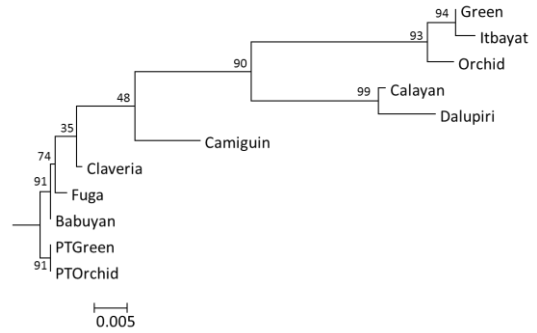
(i) *r*=0.9
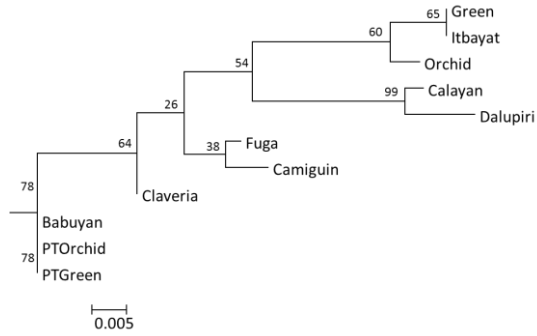
(j) *r*=1.0
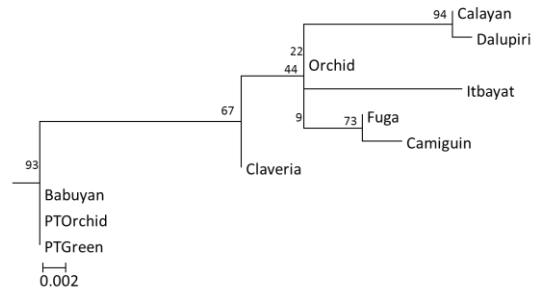
Appendix 3. Phylogenetic trees of the *Pachyrhynchus orbifer* complex reconstructed from SNPs filtered by different *p* setting. (a) to (i) are individuals assigned to collecting localities and parameter *p* from 2 to 10.

(a) *p*=2
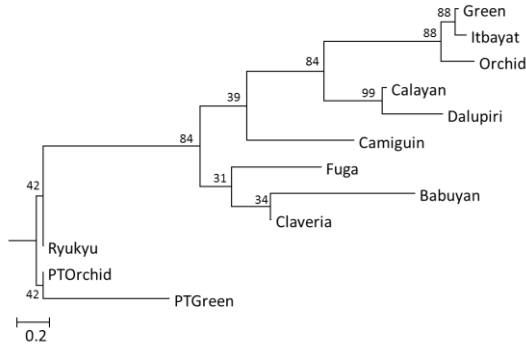
(b) *p*=3
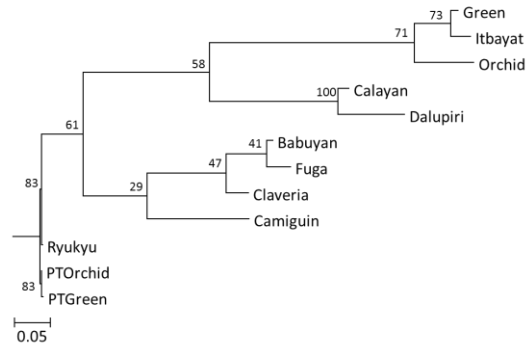
(c) *p*=4
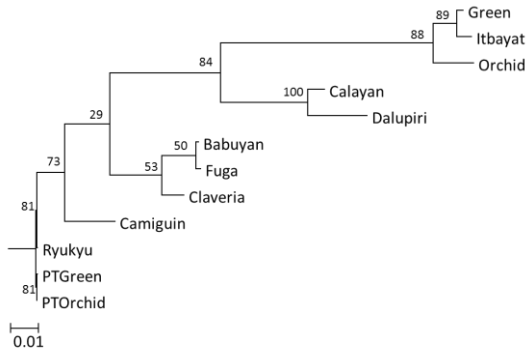
(d) *p*=5
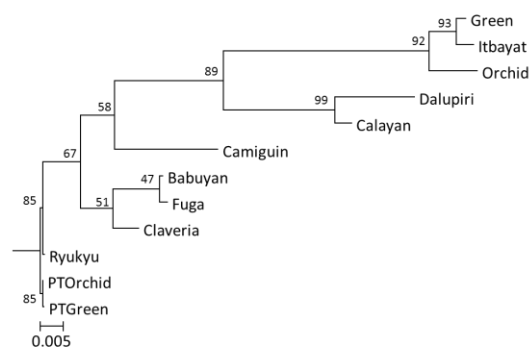
(e) *p*=6
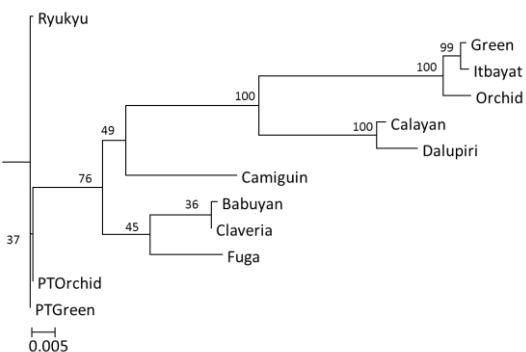
(f) *p*=7
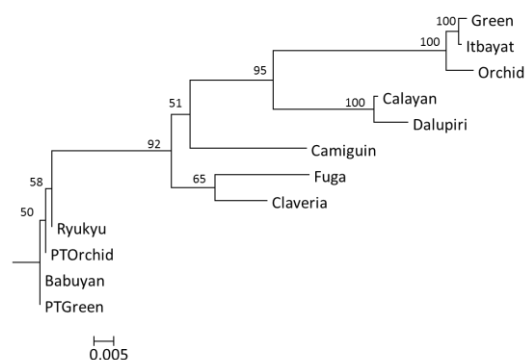
Appendix 3. (Continued)

(g) *p*=8
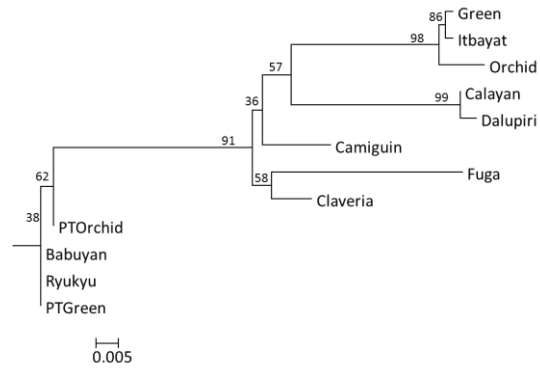
(h) *p*=9
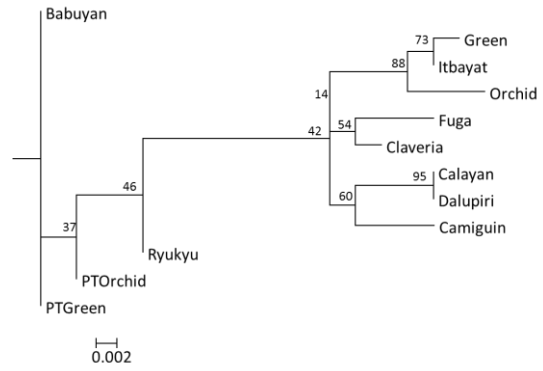
(i) *p*=10

Appendix 4. Phylogenetic trees of the *Pachyrhynchus orbifer* complex reconstructed from SNPs filtered by different *m* setting, and individuals were assigned to collecting localities.

(a) *m*=3

(b) *m*=5

(c) *m*=8

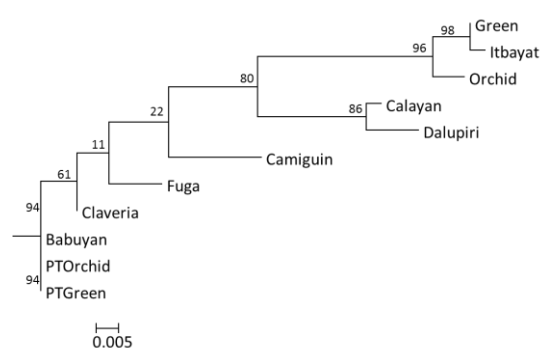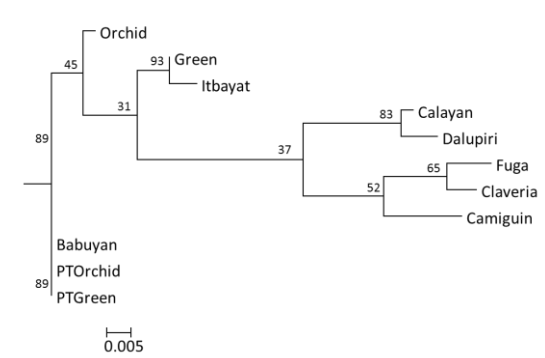(d) *m*=10

(e) *m*=12
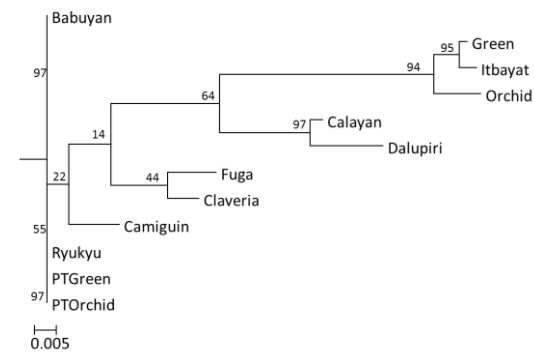
(f) *m*=18
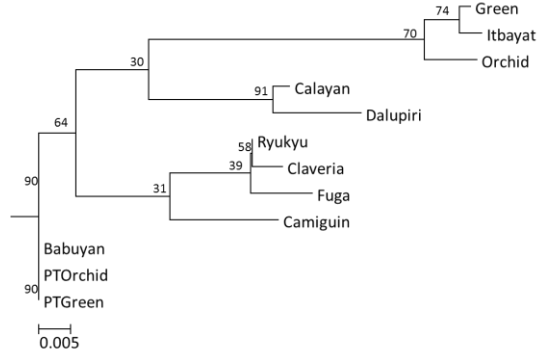
Appendix 4. (Continued)

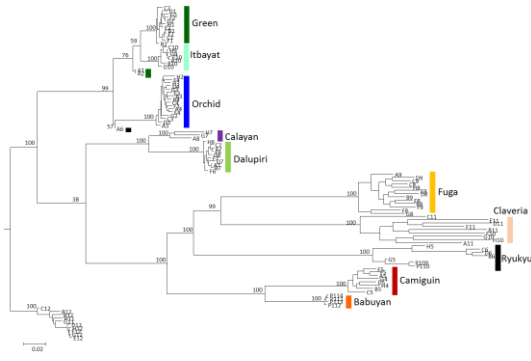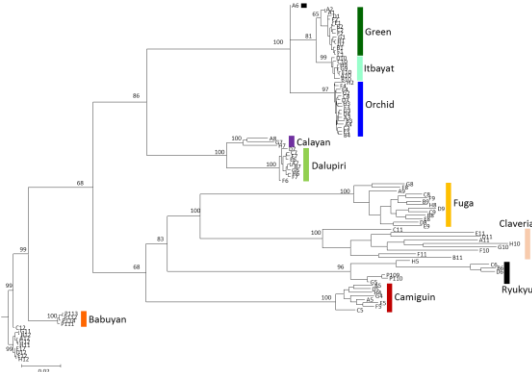(g) *m*=24
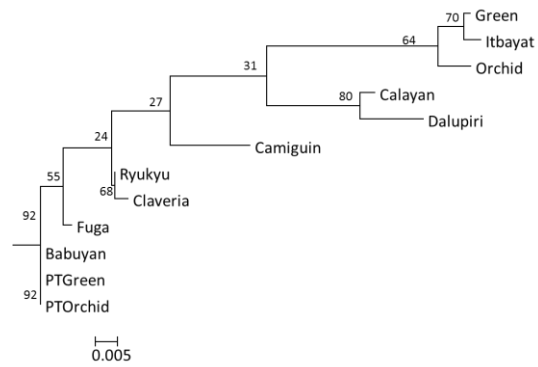
(h) *m*=30
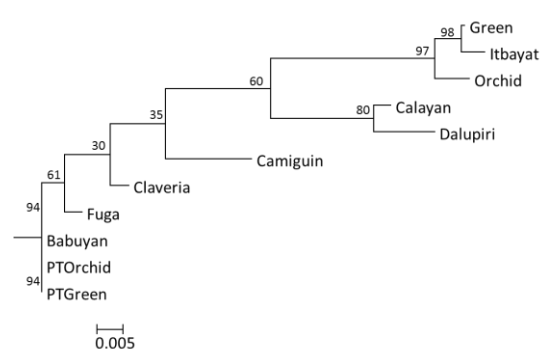
(i) *m*=40

(j) *m*=50

Appendix 5. Phylogenetic trees of the *Pachyrhynchus orbifer* complex reconstructed from SNPs filtered by different *p* setting, and each individual was treated as a single population.

(a) *p*=10

(b) *p*=30

(c) *p*=50

(d) *p*=60

(e) *p*=70

(f) *p*=80

Appendix 6. Phylogenetic trees of the *Pachyrhynchus orbifer* complex reconstructed from SNPs filtered by different *m* setting, and each individual was treated as a single population.
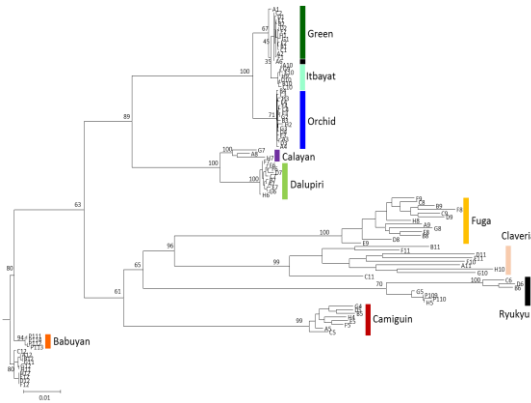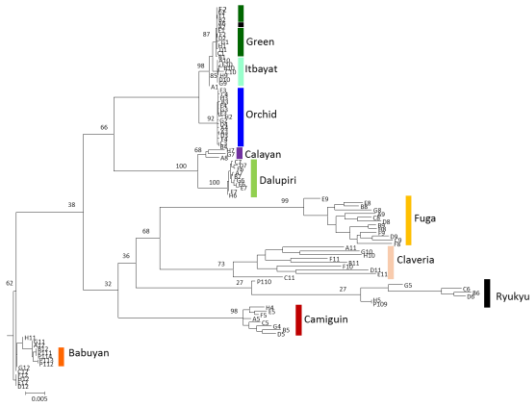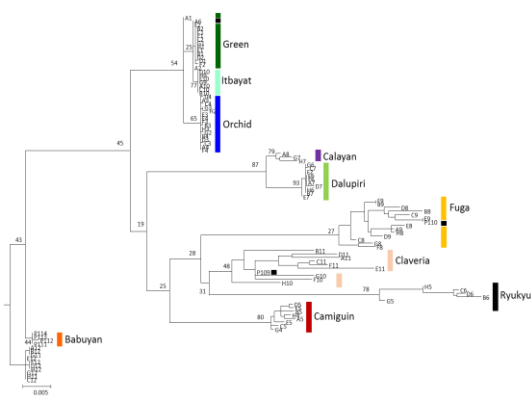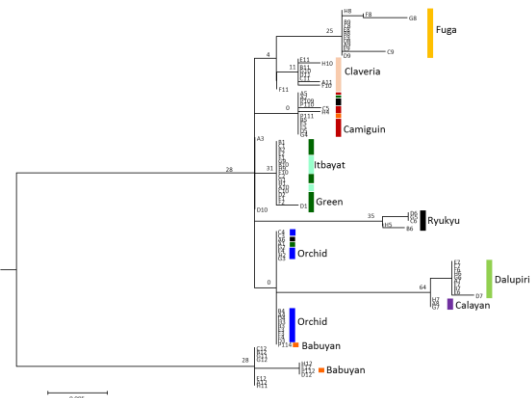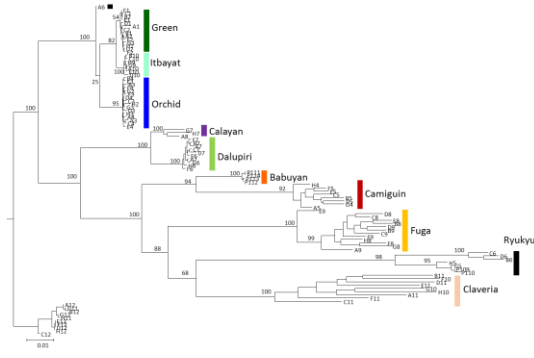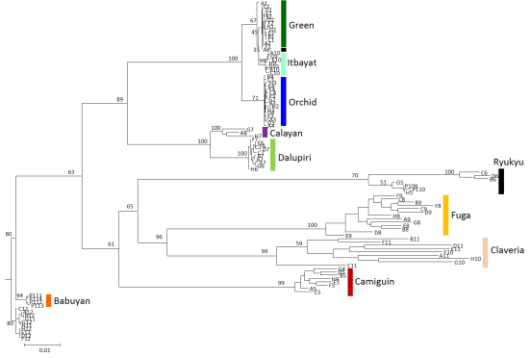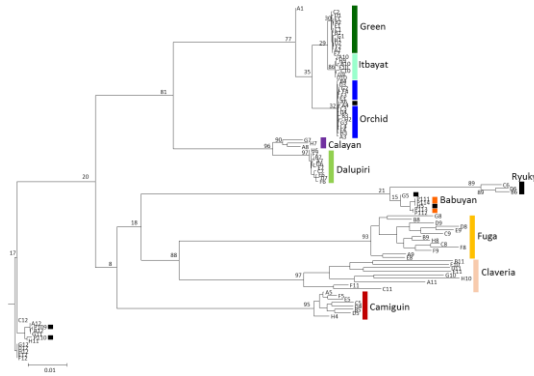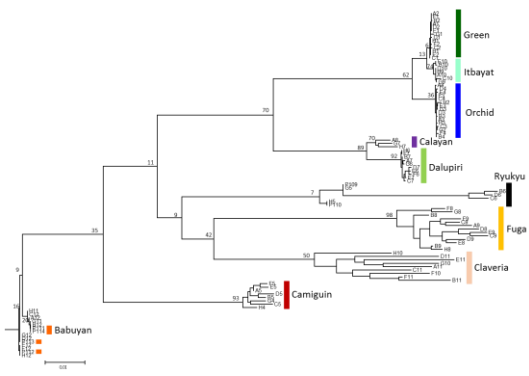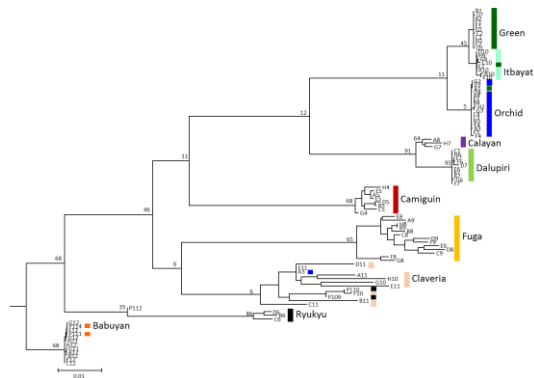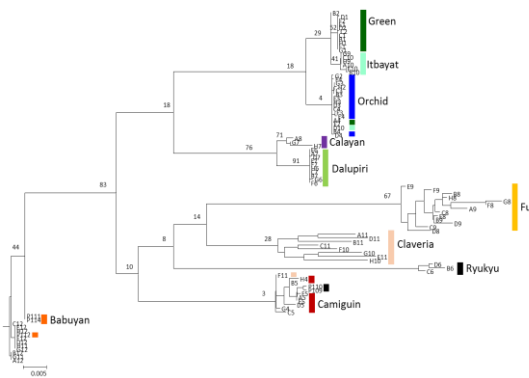
(a) *m*=2

(b) *m*=5

(c) *m*=8

(d) *m*=13

(e) *m*=20

(f) *m*=27

Appendix 7. Loci coverage of individuals in different populations.

ID	coverage	population	ID	coverage	population	ID	coverage	population
P113	10.7149	Babuyan	B9	9.05102	Fuga	A3	5.69845	Orchid
P114	41.0757	Babuyan	F8	12.7392	Fuga	D3	6.97991	Orchid
P111	46.3616	Babuyan	C8	14.7712	Fuga	C4	8.98968	Orchid
P112	49.126	Babuyan	E8	15.5066	Fuga	A4	10.6115	Orchid
G7	15.1019	Calayan	B8	16.9544	Fuga	C3	14.1108	Orchid
A8	38.4737	Calayan	H8	24.2635	Fuga	B3	14.333	Orchid
H7	76.8018	Calayan	E9	30.746	Fuga	G3	18.3473	Orchid
A5	7.89459	Camiguin	D9	30.8356	Fuga	F4	18.6996	Orchid
B5	10.0315	Camiguin	C9	31.083	Fuga	B4	20.2343	Orchid
E5	18.3164	Camiguin	D8	42.9615	Fuga	H2	27.7781	Orchid
D5	19.8192	Camiguin	G8	43.4655	Fuga	E3	29.8412	Orchid
F5	28.4903	Camiguin	F9	46.3359	Fuga	G2	30.6253	Orchid
C5	40.5782	Camiguin	A9	54.5344	Fuga	F3	41.5681	Orchid
H4	41.0121	Camiguin	A2	5.10634	Green	D4	46.183	Orchid
G4	46.1441	Camiguin	A1	5.73662	Green	E4	46.2852	Orchid
F11	6.5275	Claveria	E2	6.89197	Green	H3	59.8667	Orchid
C11	23.3206	Claveria	B2	10.365	Green	H12	18.9319	PTGreen
F10	25.7621	Claveria	D2	11.3883	Green	E12	19.3576	PTGreen
D11	26.4857	Claveria	C1	12.8973	Green	D12	24.3419	PTGreen
G10	28.9196	Claveria	B1	15.378	Green	G12	30.192	PTGreen
E11	30.3162	Claveria	E1	18.8408	Green	F12	40.2159	PTGreen
B11	37.2596	Claveria	G1	22.45	Green	C12	34.8727	PTOrchid
A11	62.626	Claveria	F2	25.2993	Green	B12	44.9906	PTOrchid
H10	71.7279	Claveria	H1	25.9706	Green	H11	56.1976	PTOrchid
E6	14.4594	Dalupiri	C2	31.0287	Green	G11	72.5483	PTOrchid
B7	23.0771	Dalupiri	D1	37.0329	Green	A12	97.0256	PTOrchid
E7	23.9045	Dalupiri	F1	46.2076	Green	G5	5.45059	Ryukyu
F6	24.4477	Dalupiri	D10	6.89763	Itbayat	H5	5.74993	Ryukyu
F7	28.5031	Dalupiri	H9	11.3143	Itbayat	A6	6.89415	Ryukyu
A7	29.8182	Dalupiri	C10	27.6118	Itbayat	D6	14.6816	Ryukyu
C7	30.331	Dalupiri	B10	28.1975	Itbayat	B6	18.1882	Ryukyu
G6	30.6509	Dalupiri	E10	37.8017	Itbayat	C6	25.0917	Ryukyu
D7	35.2528	Dalupiri	A10	38.3924	Itbayat	P109	27.545	Ryukyu
H6	40.9162	Dalupiri	G9	44.105	Itbayat	P110	34.9008	Ryukyu

Publications

SCI 期刊論文

1. Pike, D.A., R.W. Clark, A. Manica, **<u>Hui-Yun Tseng,</u>** J.-Y. Hsu, & W.-S. Huang. 2016. Surf and turf: predation by egg-eating snakes has led to the evolution of parental care in a terrestrial lizard. *Scientific Reports*, 6, 22207.

2. Lin, W.-L., S.-M. Lin, J.-W. Lin, Y. Wang, **<u>Hui-Yun Tseng*</u>**. 2015. Breeding performance of Crested Goshawk *Accipiter trivirgatus* in urban and rural environments of Taiwan. *Bird Study* 62(2): 177-184.（通訊作者）

3. Lin, W.-L., S.-M. Lin, **<u>Hui-Yun Tseng*</u>**. 2014. Colour morphs in the Collared Pygmy Owl Glaucidium brodiei are age-related, not a polymorphism. *Ardea* 102: 95–99.（通訊作者）

4. **<u>Tseng, Hui-Yun</u>**, C.-P. Lin, J.-Y. Hsu, D. A. Pike, W.-S. Huang. 2014. The functional significance of aposematic signals: geographic variation in the responses of widespread lizard predators to colourful invertebrate prey. *PLoS ONE* 9(3): 91777（第一作者）

5. Lin, W.-L., L. L. Severinghaus, **<u>Hui-Yun Tseng*</u>**, S.-M. Lin. 2013. Genetic differentiation between migratory and sedentary populations of the Northern Boobook (*Ninox japonica*), with the discovery of a novel cryptic sedentary lineage. *Journal of Ornithology* 154: 987-994.（通訊作者）

6. Lin, W.-L., S.-M. Lin, **<u>Hui-Yun Tseng*</u>**. 2012. Breeding ecology of the Northern Boobook *Ninox japonica totogo* in central Taiwan. *Forktail* 28: 164-166.（通訊作者）

Peer review 期刊論文

1. Ananjeva, N. B., E. E. Golynsky, S.-M. Lin, N. L. Orlov, **Hui-Yun Tseng**. 2015. Modeling habitat suitability to predict the potential distribution of the Kelung Cat Snake *Boiga kraepelini* Steineger, 1902. Russian Journal of herpetology 22(3):197-206.

2. **Tseng, Hui-Yun**, W.-S. Huang. 2012. *Sphenomorphus incognitus* (Brown Forest Skink). Parental care. *Herpetological Review* 43(1): 141. （第一作者）

3. **曾惠芸**、林文隆、王盈涵、趙子敬。2010。新興都會生態系--都市蛇類現況初探。台灣林業 36(3): 63-69。（第一作者）

4. 林文隆、吳居穎、**曾惠芸***。2009。森林"鴞"雄-談大雪山 200 林道的夜行性猛禽資源。台灣林業 35(2): 19-25。（通訊作者）

5. 林文隆、**曾惠芸***、王穎、陳明德。2008。喜馬拉雅林鴞繁殖生態學：繁殖、食性與捕食行為描述。特有生物研究 10(2): 13-24。（通訊作者）

6. Lin, W.-L., Y. Wang, **Hui-Yun Tseng***. 2007. Initial investigation on the diet of Eastern Grass Owl (*Tyto longimembris*) in southern Taiwan. *Taiwania* 52(1): 100-105. （通訊作者）

7. 林文隆、**曾惠芸***。2007。為貓頭鷹造一個家-談貓頭鷹巢箱設計與試驗過程。台灣林業 33(6): 70-76。（通訊作者）

8. 林文隆、**曾惠芸***。2007。九九峰自然保留區猛禽類資源。台灣林業 33(2): 49-54。（通訊作者）

9. 林文隆、**曾惠芸***、王穎。2007。台灣中部中高海拔山區道路邊緣灰林鴞食性研究。特有生物研究 9(1): 13-21。（通訊作者）

10. 林文隆、**曾惠芸***、曾啟政。2006。蛇類體長的測量：快速、精確與安全的電腦影像處理技術。臺大實驗林研究報告 20(4): 279-283。（通訊作者）

11. 林文隆、**曾惠芸***。2006。九九峰自然保留區陸域脊椎動物資源調查。台灣林

業 32(4): 47-55。(通訊作者)

研討會發表

1. **Hui-Yun Tseng**, J.-Y. Hsu, L.-Y. Wang, W.-S. Huang. 2016. Larger bite force driven by parental care in an insular skink population. The 8th World Congress of Herpetology. P489. August 15-21. Hanzhou, China. (國際研討會)

2. Wang, L.-Y., **Hui-Yun Tseng**, J.-Y. Hsu, Wen-San Huang. 2016. Denser the intraspecific competition, higher the bite force? Comparison the bite force among three populations of *Japalura swinhonis* in Taiwan. The 8th World Congress of Herpetology. P476. August 15-21. Hanzhou, China. (國際研討會)

3. Hsu, J.-Y. Hsu, D. A. Pike, **Hui-Yun Tseng**, W.-S. Huang. 2016. Clutch size variation of an insular lizard under high predation pressure from egg-eating snakes. The 8th World Congress of Herpetology. P472. August 15-21. Hanzhou, China. (國際研討會)

4. Poyarkov, N. A., H.-Y. Hou, E. Solovyeva, J.-S. Lai, **Hui-Yun Tseng**, S.-M. Lin. 2016. A comprehensive DNA COI-barcode reference library for the identification of amphibians in Taiwan and adjacent islands. The 8th World Congress of Herpetology. P554. August 15-21. Hanzhou, China. (國際研討會)

5. **曾惠芸**、鄭明倫、Reagan Joseph T. Villanueva、Olga M. Nuñeza、林仲平、黄文山。2016。Multi-directional colonization of Pachyrhynchid weevils in Taiwan-Luzon Arc。動物行為暨生態學研討會。中興大學，台中。

6. 李昶毅、**曾惠芸**、郭育瑋、許嫆雅、黄文山。2016。斯文豪氏攀蜥如何辨認小圓斑球背象鼻蟲為不好吃之獵物？動物行為暨生態學研討會。中興大學，台中。

7. 莊伯孟、廖伯盛、**曾惠芸**、林仲平、黄文山。2015。Body size and coloration

do not affect mating performance of *Eupyrgops waltonianus*。動物行為暨生態學研討會摘要集 88 頁。嘉義大學，嘉義。

8. **<u>Tseng, Hui-Yun</u>**, W.-S. Huang, M.-L. Jeng, R. J. T. Villanueva, O. M. Nuñeza, C.-P. Lin. 2014. Evolution of coloration and colonization history of *Pachyrhynchus* weevils in the Philippine archipelago. 第三十五屆台灣昆蟲學會年會。October 17-18, 2014。國立臺灣師範大學，台北市。

9. **<u>Tseng, Hui-Yun</u>**, W.-S. Huang, C.-P. Lin. 2014. Evolution on the Philippine islands: a case of flightless pachyrrhynchid weevils. 2014 動物行為暨生態學研討會。January 20-21, 2014。東海大學，台中市。

10. **<u>Tseng, Hui-Yun</u>**, C.-P. Lin, W.-S. Huang. 2013. Conspicuous colors of pachyrrhynchid weevils function as aposematic visual cues for agamid lizards. 第一屆臺灣兩棲爬行動物研討會。August 31, 2013。國立自然科學博物館，台中市。

11. Lin, Y.-P., **<u>Hui-Yun Tseng</u>**, S.-M. Lin. 2013. Inter- and intra-island differentiation: a case from a Japanese treefrog widely distributed in Taiwan and Ryukyus. 2013 年動物行為暨生態學研討會。January 25-26, 2013。國立東華大學，花蓮縣。

12. **<u>Tseng, Hui-Yun</u>**, C.-P. Lin, W.-S. Huang. 2012. To eat or not to eat? The relationship between predatory behavior of Swinhoe's tree lizard (*Japalura swinhonis*) and the coloration in Pachyrrhynchid weevils in Taiwan. 5th Asian Herpetological Conference (p. A015). June 1-4, 2012. Cheengdu, China. （國際研討會）

13. **<u>Tseng, Hui-Yun</u>**, C.-P. Lin, W.-S. Huang. 2012. To eat or not to eat?–Different predation response between allopatric and sympatric Swinhoe's tree lizards (*Japalura swinhonis*) to a conspicuous colored pachyrrhynchid weevil (*Kashotonus multipunctatus*). 2012 年動物行為、生態與全球變遷研討會-中國生物學會、台灣哺乳動物學會聯合年會。January 17-18, 2012。中國文化大學，台北市。

14. Lin, W.-L., S.-M. Lin, L. L. Severinghaus, **Hui-Yun Tseng**. 2012. Genetic differentiation between migratory and residential populations of the Northern Boobook (Ninox japonica, Strigidae), with a special note on a cryptic lineage. 2012 年動物行為、生態與全球變遷研討會-中國生物學會、台灣哺乳動物學會聯合年會。January 17-18, 2012。中國文化大學，台北市。

15. **Tseng, Hui-Yun**, N.-H. Jang-Liaw, W.-H. Chou. 2011. Phylogeography and genetic differentiation of Sylvirana latouchii in Taiwan and China. 1st International Symposium on east Asian vertebrate species diversity (p. 35). August 8-12, 2011. Guangzhou, China. （國際研討會）

16. 林德恩、杜銘章、陳昱凱、陳志耘、**曾惠芸**。2011。低海拔山區蜥蜴族群動態及生活史。2011 動物行為、生態暨生物及環境教育研討會。國立臺南大學，台南市。

17. **曾惠芸**、林文隆。2010。台灣褐鷹鴞留棲屬性探討。第四屆台灣猛禽生態研討會。Novemver 22, 2010。國立台灣師範大學，台北市。

18. 林文隆、王穎、**曾惠芸**。2009。都會鵂鶹繁殖生物學研究。2009 年台灣鳥類論壇。國立台灣大學，台北市。

19. 林文隆、王穎、**曾惠芸**。2009。越來越多的夜鷹-台灣夜鷹在都市繁殖的優勢。2009 年台灣鳥類論壇。國立台灣大學，台北市。

20. 林德恩、**曾惠芸**。2008。淡水龜的爬坡能力。2008 動物行為暨生態研討會。2008 動物行為暨生態研討會。January 28-29, 2008。國立台灣海洋大學，基隆市。

21. 林文隆、**曾惠芸**、王穎。2008。台灣中部都會區與原始環境鳳頭蒼鷹繁殖與食性研究。第七屆海峽兩岸鳥類學術研討會。January 26-27, 2007。國立台灣師範大學，台北市。

22. 曾啟政、**曾惠芸**、林文隆、廖翔凱。2007。陳氏鰍鮀生殖生物學與棲地環境初探。2007 動物行為暨生態研討會。2007 動物行為暨生態研討會-中國生物學

會聯合年會。January 27-28, 2007。國立花蓮師範大學，花蓮市。

23. 林文隆、**曾惠芸**。2007。黃嘴角鴞繁殖生物學初探。2007 動物行為暨生態研討會。2007 動物行為暨生態研討會-中國生物學會聯合年會。January 27-28, 2007。國立花蓮師範大學，花蓮市。

24. 尤少彬、**曾惠芸**、張心怡。2003。空中覓食性同功群鳥類形態特徵之初探。2003 動物行為暨生態研討會。January 23-24, 2003。高雄醫學大學，高雄市。

科普文章

1. **曾惠芸**。2016。越東邊境的岩虎。國立自然科學博物館館訊 343 期。

2. **曾惠芸**。2015。市區王者—鳳頭蒼鷹。國立自然科學博物館館訊 334 期。

3. **曾惠芸**、林仲平、許嫆雅、黃文山。2014。球背象鼻蟲-島嶼上的瑰麗珠寶。科學人雜誌 150 期。

4. **曾惠芸**、林思民。2014。戰鬥民族的博物館：俄羅斯國家科學院動物館藏。國立自然科學博物館館訊 323 期。

5. **曾惠芸**。2013。DNA 在爬蟲類研究的應用。國立自然科學博物館館訊 312 期。

6. **曾惠芸**。2013。吃素的巨蜥。國立自然科學博物館館訊 304 期。

7. 張育誠、高梅婷、陳姿言、**曾惠芸**、許鴻龍。2012。地下通勤族—淺談臺灣鼯鼠生態習性及飼養方法。自然保育季刊 80: 43-51。

8. **曾惠芸**。2012。蛇中熊貓-莽山烙鐵頭。國立自然科學博物館館訊 293 期。

9. **曾惠芸**、林文隆。2012。台灣褐鷹鴞族群分類之謎。黃山雀會訊 186: 8-9.

10. 黃文山、**曾惠芸**。2011。小島嶼開啟爬蟲研究的新視野。NOW 野生動物保育彙報及通訊 15(4): 35-42。

11. **曾惠芸**、林文隆。2011。都市新住民-南亞夜鷹。國立自然科學博物館館訊 281 期。

12. **曾惠芸**、林德恩。2008。外來入侵種多線南蜥分布北界的調查記錄。自然保育季刊 61: 37-42。

13. 林德恩、陳元龍、**曾惠芸**、蔡忠穎。2007。「不蜥──入侵種多線南蜥研究現況介紹及防治建議」解說摺頁。

14. **曾惠芸**、高孝偉。2004。與月光共舞的羽翼-台灣夜鷹。自然保育季刊 46: 56-59。

國際研討會獲獎記錄

Most meritorious oral presentation award. **Tseng, Hui-Yun**, C.-P. Lin, W.-S. Huang. 2012. To eat or not to eat? The relationship between predatory behavior of Swinhoe's tree lizard (*Japalura swinhonis*) and the coloration in Pachyrrhynchid weevils in Taiwan. 5th Asian Herpetological Conference (p. A015). June 1-4, 2012. Cheengdu, China.

Award for the best presentation for young scientist. **Tseng, Hui-Yun**, Jang-Liaw, Nian-Hong, Chou, Wen-Hao. 2011. Phylogeography and genetic differentiation of Sylvirana latouchii in Taiwan and China. 1st International Symposium on east Asian vertebrate species diversity (p. 35).August 8-12, 2011. Guangzhou, China.

學位論文

曾惠芸。2006。鯊魚、線鰻與雞磷酸葡萄糖異構酶的選殖，並推論磷酸葡萄糖異構酶在脊椎動物的演化。國立中興大學生命科學系碩士論文。

曾惠芸。2017。球背象鼻蟲斑紋功能與島嶼間演化歷史。東海大學生命科學系博士論文。