Phylogeographic history of the ring distribution in *Malcus flavidipes* (Heteroptera: Malcidae) shaped by directional dispersal and island adaptation in the South China Sea region

Yan-Fei Li[1,#], Mei-Qi Guo[1,#], Ju-Hong Chen[1,2], De-Qiang Ai[1], Zi-Hao Li[1], Jing-Yao Zhang[1], Zhen Ye[1], Huai-Jun Xue[1], Shu-Jing Wang[1,*], Wen-Jun Bu[1,*]

[1] Institute of Entomology, College of Life Sciences, Nankai University, Tianjin 300071, China

[2] College of Life Sciences, Anhui Normal University, Wuhu 241002, Anhui Province, China

[#]Authors contributed equally to this work.

[*]Corresponding authors: E-mails: wangsj@nakai.edu.cn (S.-J. Wang); wenjunbu@nankai.edu.cn (W.-J. Bu).

Science Press

ZR Zoological Research

ABSTRACT:

The ring distribution pattern offers a unique perspective for analysing the mechanisms of species or population differentiation. The semi-closed South China Sea (SCS) region, with its islands, peninsulas, and continental margins, is ideal for studying the formation of ring distribution patterns and the evolutionary dynamics of island populations. This study focused on *Malcus flavidipes*, a species distributed in the SCS region. Based on double-digest restriction-associated DNA data and geographic distribution data, population genetic approaches and species distribution modeling were employed to investigate its distribution pattern formation and the evolutionary characteristics of island populations. The results revealed five geographic clades around the SCS which are located on the Yunnan–Indochina Peninsula, southern China, Hainan Island, Kalimantan and the Philippines. Genetic diversity evidence supported that the ring distribution pattern originated in the Yunnan-Indochina Peninsula and diffused through the land bridges exposed during the Pleistocene glaciation along two routes: the northern route extended from southern China to Hainan Island, whereas the southern route stretched from Kalimantan to the Philippines. Island populations show significantly lower genetic diversity and higher within-population inbreeding coefficients than mainland ones. The genetic structural differences among different island populations are closely related to their geographic distance from mainland, responses to historical climatic fluctuations, and selective pressures from island environments. This study provides important empirical evidence for the ring differentiation of species in the SCS region, and offers a new perspective for understanding the evolutionary mechanisms of island populations under the combined effects of geographical isolation, genetic drift, and environmental selection.

Keywords: South China Sea region; Ring distribution pattern; Island biogeography; Population genetics; ddRAD-seq; Pleistocene

INTRODUCTION

Geographically, ring distribution is a relatively special pattern of species distribution (Afzali et al., 2024; Irwin et al., 2001; Keightley et al., 2009; Kuchta et al., 2009). This distribution model not only allows exploration of the dispersal and evolutionary history of species distribution formation but also facilitates discussions on the historical processes underlying the formation of ring distributions (Bouzid et al., 2022; Wang & Bu, 2020).

The region surrounding the South China Sea (SCS) is a semiclosed geographic unit centred on the SCS and is composed of islands, peninsulas, and continental margins that encircle it (Hall, 2002; Yang et al., 2017). The vast expansion of the SCS forms a natural geographic barrier to some extent, providing a unique "ring distribution" terrain for terrestrial species around the central geographic barrier (Fuchs et al., 2015; Singham et al., 2017). The region is located at the junction of multiple tectonic plates, encompassing four important biodiversity hotspots: the Sunda region, the Philippines region, Wallacea, and the India-Burma region of the Indochinese Peninsula, which exhibits exceptionally high biodiversity (Myers et al., 2000). Moreover, the drastic climatic fluctuations during the Pleistocene caused significant sea-level changes around the SCS: land bridges emerged during glacial periods, providing corridors for terrestrial animal dispersal, whereas during interglacial periods, warmer climates and rising sea levels submerged these land bridges (Cros et al., 2020; Gorog et al., 2004; Lavery et al., 2023). The marine barriers cut off gene flow between populations, driving adaptive evolution of local populations and accelerating the process of population differentiation (Thörn et al., 2021). The SCS region has a complex geological and climatic history and environmental heterogeneity, complicating biological distribution and dispersal patterns.

However, currently, most of the research on the SCS region focuses on its related historical and geographical features and other aspects. There are relatively few studies focused on species differentiation, and most of these studies are carried out on plant groups (for example, research on the distribution pattern of mangrove species diversity

and its formation mechanism (Yang & Ren, 2023) and the formation mechanism of island plant diversity (Chiang et al., 2001)).

Insects are characterized by high sensitivity to environmental changes and short generation times, which gives them an advantage in regional biological evolution studies—they can quickly record the impacts of factors such as glacial-interglacial climate fluctuations and "land bridge emergence and submergence" on population dynamics. However, current phylogeographic research on insects within SCS region is mostly limited to local studies, covering some islands (Dusfour et al., 2007; Singham et al., 2017) or focusing on mainland China and adjacent islands (Damaska et al., 2021; Fu et al., 2024; Tian et al., 2015). This leads to a lack of understanding of its overall pattern, and there is an urgent need to supplement comprehensive evidence from insect taxa. Furthermore, multiple studies have shown that the dispersal and differentiation of insect island populations in Southeast Asia are temporally consistent with the dynamic changes of land bridges during glacial and interglacial periods (Hosoishi et al., 2023; Liu et al., 2015; Schutze et al., 2012; Singham et al., 2017). Based on this, we hypothesize that the dispersal of insects in the SCS region may have been significantly influenced by historical climatic fluctuations and the periodic emergence and submergence of land bridges.

Islands have long been regarded as model systems for exploring various aspects of ecology and evolutionary biology, from speciation to community assembly (Gyllenhaal et al., 2025; Valente et al., 2018). The evolutionary processes of species across different islands can be analogized to "natural experiments", where species with similar biological and ecological traits are placed in different environments (Gray & Cavers, 2014; Whittaker et al., 2017). This unique model provides an irreplaceable research framework for dissecting the impacts of environmentally selective pressures and historical contingency on species genetic diversity (Hirschfeld et al., 2023). In the SCS region, Hainan Island (located at the northern margin of the SCS, adjacent to the southern tip of the Chinese mainland), Borneo (at the southern margin of the SCS), and the Philippine Archipelago (at the eastern margin of the SCS) form a typical "semiring island network". The geographical location and environmental differences among these

islands provide a natural paradigm for explaining the mechanisms of population differentiation between islands.

Malcus flavidipes Stål, 1859 is widely distributed in the SCS region (Wang & Bu, 2020). As a phytophagous species with weak flight and migration capabilities, it lacks the ability to disperse across seas via wind. Therefore, we hypothesize that its historical dispersal and current distribution may have been influenced by local historical climatic fluctuations and changes in land bridges, making it an ideal material for studying the evolutionary history of species, their adaptation mechanisms, and population dynamic changes under the distribution pattern in the SCS region.

In this study, we integrated double-digest restriction-site associated DNA sequencing (ddRAD-seq) data and geographical distribution data for *M. flavidipes* and combined them with climate and environmental data. Through phylogeographic research methods, this study aims to: 1) reveal the phylogeographic pattern of this species within the SCS region; 2) elucidate the evolutionary history and population dynamic changes of this species, and verify the hypothesis of "the dispersal pattern driven by the periodic changes of land bridges"; 3) conduct comparative analyses of populations across different islands to explore the driving mechanisms of island geographical location, environmental factors, and historical climatic events on population differentiation.

MATERIALS AND METHODS

Sampling

This study included a total of 199 samples, consisting of 179 individuals (24 populations) of *M. flavidipes* from the SCS region, and 20 individuals (2 populations) of its sister species *M. inconspicuous*, which was selected as the outgroup (Figure 1a; Supplementary Table S1) (Wang & Bu, 2020). All individuals were immediately preserved in 100% ethanol and stored in a freezer at −20°C before DNA extraction.

Voucher samples were stored at the Institute of Entomology, College of Life Sciences, Nankai University, China (NKU).

Molecular laboratory procedures and single-nucleotide polymorphism (SNP) calling

Genomic DNA was extracted from muscle tissue via a Universal Genomic DNA Kit (CWBIO, Beijing, China).

The ddRAD library was prepared according to Peterson′s protocol (Peterson et al., 2012). Briefly, the genomic DNA was digested with EcoRI and MspI restriction enzymes (New England Biolabs, Ipswich, MA, USA) at 37°C for 8 hours. Both ends of the digestion product were subsequently ligated with Illumina adaptors and unique 5 bp barcodes. Ligation products were pooled together, and fragments from 250 bp to 600 bp were selected through Pippin Prep (Sage Science, Beverly, MA, USA). Finally, the size-selected pools were amplified with 8–10 cycles via PCR with Illumina-indexed primers. All intermediate products, including enzyme-digested products, ligation products and amplified DNA fragments, were purified via AMPure XP magnetic beads (Beckman Coulter Inc., Brea, CA, USA) before proceeding to the next step. The ddRAD libraries were sent to the Novogene Sequencing Center & CAP and ISO Lab for sequencing, and paired-end 150 bp reads were generated on the Illumina HiSeqX10 platform.

We processed raw Illumina reads and called SNPs via the ‘API: ipyrad assembly workflow’ (Eaton & Overcast, 2020). For the ipyrad parameters, the de novo method was used for assembly, and the clustering threshold was set to 0.85 (to avoid the issue where homologous sequences might fail to cluster together at an excessively high clustering threshold-due to the presence of Ns, insertions/deletions (indels), sequencing errors, or polymorphisms). The average Phred score offset was set to 33 (the default

offset), the maximum number of low-quality, undetermined sites ('N') in filtered sequences was set to 5 (the default value and generally recommend against increasing this value greatly), the filter for the adapters was set to 2 (strict filter for adapters is applied to address the issue of adapter sequence contamination in ddRAD data caused by abnormal fragment lengths). The generated ddRAD dataset requires that each locus exists in at least 70% of individuals (i.e., 30% missing samples per locus). Two datasets were finally generated: 1) ALL70_SNP: composed of SNP data that includes all samples. 2) NOUT70_USNP: composed of unlinked single-nucleotide polymorphism (USNP) that includes all samples excluding the outgroup.

Phylogenetic analysis, population genetic structure, population genetic diversity and differentiation

The neighbour-net method based on NOUT70_USNP was implemented in SplitsTree v4.14.5 (Huson & Bryant, 2006). In addition, the maximum likelihood method based on ALL70_SNP was conducted with IQ-TREE v2 (Minh et al., 2020) with automatic model screening and additional ASC options, and the node support values were assessed via bootstrap resampling (BP) and calculated via 1 000 replicates.

Both Bayesian clustering and discriminant analysis of principal components (DAPC) were used for population genetic structure analysis. The Bayesian clustering method was performed in Structure v2.3.4 (Falush et al., 2003). We ran Structure ten times for each K value from 1 through 10 to determine the optimal number of groups (K). Each run was performed for 120 000 Markov chain Monte Carlo generations with a burn‐in period of 20 000 generations. The best K value was found via Structure Harvester (http://taylor0.biology.ucla.edu/structureHarvester/), and ten runs were combined into one output via CLUMPP v1.1.2 (Jakobsson & Rosenberg, 2007). The graphical representation was implemented in distruct v1.1 (Rosenberg, 2004). DAPC

analysis was performed via the R package adegenet (Jakobsson & Rosenberg, 2007). We used the five groups based on the results of Bayesian clustering analysis and phylogenetic analysis as prior. The optimal number of PCs was determined via the xvalDapc() command. The optimal number of PCs retained was 30, and we used the PCs number range of 25 to 35 for analysis.

The nucleotide diversity and within-population inbreeding coefficient (*FIS*) were calculated in VCFtools v0.1.17 (Danecek et al., 2011). The pairwise F_{ST} values among the sampling localities and clades were estimated via the R package hierfstat (Goudet, 2005). The analysis of "isolation by distance" (IBD) was performed via the R package ade4 (Stéphane & Anne-Béatrice, 2007) via the Mantel test method based on the paired geographic differences and paired F_{ST} values among the sampling localities. The minimum spanning networks (MSNs) were constructed in the R package popr via the Rogers distance calculation method.

Demographic model testing

The demographic history was simulated in Fastsimcoal v2.6.1 (Excoffier et al., 2013) from the site frequency spectra (SFS) via a coalescent simulation-based method. The observed SFS was generated in easySFS (https://github.com/isaac overast/easySFS). First, we tested the best-fit demographic model for the topological relationships among three clades with five divergence models. Second, we assume that there is gene flow between all populations. This model was run independently 50 times via 100 000 coalescent simulations and 20 optimization cycles to obtain a maximum likelihood that could be used to evaluate the results. The run with the smallest difference between MaxObsLhood and MaxEstLhood was chosen as the best run and as the representative of this model. The parameters of the model were estimated via parametric bootstrapping. First, we simulate 100 replicate SFS runs from the

*_maxL.par file for the best-fit run. We then performed 50 analyses as described above for each of the 100 newly simulated SFS files. Finally, we calculated the mean parameter estimates and 95% confidence intervals (CIs) from the 100 best-fit bootstrapping replicates. The nuclear mutation rate (3.5E-9 per site per generation) was specified following the estimates for *Drosophila melanogaster* (Keightley et al., 2009).

Historical demographic changes

The historical demographic reconstruction of the populations was performed in Stairway Plot v2 (Liu & Fu, 2020) for each clade. The population size changes over time were inferred based on a one-dimensional folded SFS. The SFS for three clades was generated via easySFS (https://github.com/isaac overast/easySFS). The nuclear mutation rate of $3.5E^{-9}$ per site per generation followed the estimates for *Drosophila melanogaster* (Keightley et al., 2009), and a generation time of 0.5 years was adopted for the analysis.

Estimated effective migration surfaces

Estimated effective migration surfaces (EEMS) analysis involves covering the study area with a dense equidistant triangular grid and each sample location is adjusted to the closest deme (vertex) on the grid. A Bayesian approach is then used to estimate the migration parameters via a stepping-stone model, which assumes that individuals migrate locally between demes and that each deme exchanges migrants only with its neighbors. We computed genetic dissimilarity matrices via EEMS and assigned geographical coordinates to each sample from each district to contrast geographic and genetic distances between demes. The migration surface contours were estimated via 500 demes. The MCMC analysis was run for 15 000 000 MCMC iterations including 14 000 000 burn-in iterations and repeated using 4 different seeds to ensure the convergence of the MCMC chains as well as the accepted distributions of parameters

as specified by the developers. Final spatial visualizations illustrating migratory surfaces were generated via R scripts provided by EEMS (Petkova et al., 2016).

Species distribution modeling

The current species distribution modeling (SDM) was constructed using the maximum entropy algorithm in Maxent v3.4.1 (Dormann et al., 2007) and then extrapolated to palaeoclimatic layers of the last interglacial (LIG) period, the Mid-Holocene period and the last glacial maximum (LGM) period.

The occurrence data of *M. flavidipes* were collected from fieldwork, dried specimens (NKUM), and published literature (Wang & Bu, 2020). The geographic coordinates were filtered out by enforcing a distance of 20 km between records via the R package Wallace (Team, 2021). Finally, 61 unique records were obtained for subsequent analyses (Supplementary Table S2). A total of 19 bioclimatic variables with a spatial resolution of 2.5 minutes were obtained from the WorldClim (https://worldclim.org/) website. To avoid the interference of multicollinearity among variables, we combined the cross-correlations of bioclimatic variables (with a threshold of $|r| < 0.8$) and the results of variable contribution analysis to select the final set of bioclimatic variables. Finally, eleven bioclimatic variables were selected for further analysis: Annual Mean Temperature (BIO1), Mean Diurnal Range (BIO2), Temperature Seasonality (BIO4), Max Temperature of Warmest Month (BIO5), Temperature Annual Range (BIO7), Mean Temperature of Wettest Quarter (BIO8), Annual Precipitation (BIO12), Precipitation Seasonality (BIO15), Precipitation of Wettest Quarter (BIO16), Precipitation of Warmest Quarter (BIO18), Precipitation of Coldest Quarter (BIO19).

Before modelling, the various regularization multipliers (the regularization multiplier was set from 0.5 to 5, and the intervals were set to 0.5) and five feature

combinations (F = L, LQ, LQP, LQPT, LQPTH; linear = L, quadratic = Q, product = P, threshold = T, hinge = H) were tested via the R package kuenm; as a result, the regularization multiplier = 1.5, feature = LQ was selected as the best combination (delta.AICc = 0, Supplementary Table S3). For modelling, the output format = logistic, the number of replicate runs = 20 and the maximum number of background points = 10 000. Seventy-five percent of the data were used for training, and 25% were used to test the model. The area under the curve (AUC) of the receiver operating characteristic (ROC) plot was used for model evaluation. The predicted binary layers were visualized in ArcGIS v10.8 (Esri, 2020).

Niche comparison analysis

Niche similarity was tested between each phylogeographic group. For groups with fewer 5 geographic data points, we simulated geographic data based on real geographic records to reach at least 5 points (to reach the requirements of subsequent software analysis). First, we extracted the environmental variable information through real geographic records to obtain its range of ecological factors for suitable habitats. Next, we extracted the latitude and longitude coordinates of locations within a 1-degree range from real geographical records within a suitable range of ecological factors. The extraction of ecological factors and geographical coordinates was performed via the R packages rgeos, sp, rgdal and raster (Kaya et al., 2019).

Eleven bioclimatic variables of each geographical location were extracted based on latitude and longitude via the R packages rgdal and raster. The environmentally space-based niche equivalence statistic method was used to test niche similarity and was performed via the R package Humboldt (Brown & Carnaval, 2019). Schoener's D niche similarity value between environmental conditions ranges from 0 (no overlap) to

1 (identical predictions). We also calculated the correlation between genetic diversity and geographic and ecological information via Pearson correlation analysis.

RESULTS

Single-nucleotide polymorphism (SNP) calling

After filtering, the number of reads mapped to each individual ranged from 1,503,927 to 15,707,201 with 5,844,800 reads per individual on average. After the loci that failed to meet the filtering criteria were removed, two final datasets were constructed for analysis: 1) ALL70_SNP: dataset containing all samples that contained 11336 SNP loci. 2) NOUT70_USNP: dataset containing all samples excluding outgroups that contain 5193 USNP loci.

Phylogenic results and population genetic structure

Based on the NOUT70_USNP dataset, the results of the neighbour-net method revealed that *M. flavidipes* was divided into five independent genetic clades that corresponded to geography: 1) South China and Southwest clade (FLASW) was located mainly in South China and Southwest China; 2) Yunnan-Indochina clade (FLAYL) was located mainly in the Indo-China Peninsula and Malay Peninsula; 3) Hainan Island clade (FLAHN) was located only on Hainan Island, China; 4) East Malaysia clade (FLAEM) was located only in East Malaysia; and 5) Philippines clade (FLAPH) was located only in the Philippines. The ML tree strongly supported that the five clades mentioned above were monophyletic, and the phylogenetic relationship was (((FLASW+FLAYL)+FLAHN)+(FLAEM+FLAPH)) (Figures 1A & C).

The structure results revealed that the optimal K value was 5 (Supplementary Figures S1-2). FLASW and FLAEM mainly contained one genetic component (sky blue and orange, respectively). For FLAYL, the populations located on the Indo-China

Peninsula mainly contained one genetic component (green), and the populations located on the Malay Peninsula likely had mixed genetic components (yellow). FLAHN and FLAPH both had mixed genetic components with genetic components from neighboring island populations (Figure 1B). The DAPC results were consistent with the phylogenetic results, showing five obvious clusters (Figure 1D).

MSN showed that the FLAYL populations was connected to the FLASW populations and the FLAEM populations, respectively. Subsequently, the FLASW populations differentiated into the FLAHN populations, and the FLAEM populations then differentiates into the FLAPH populations (Figure 1E).

Genetic diversity, genetic differences and within-population inbreeding coefficient

The nucleotide diversity ranged from 0.0446 in flaYNJH population of FLAYL to 0.0121 in flaMSJJ of FLAEM (Supplementary Table S1)., with that of each branch was ranked in descending order: FLAYL > FLASW > FLAHN> FLAPH> FLAEM (Figure 2A).

The pairwise F_{ST} showed that the values between the populations ranged from 0.0228 to 0.1435 (Supplementary Figure S3; Supplementary Table S4). According to the F_{ST} values, the five clades could be divided into two groups, with FLAYL, FLASW, and FLAHN forming one group and FLAEM and FLAPH forming one group. Within the groups, there was relatively little genetic differentiation among the clades (F_{ST} values lower than 0.08), whereas between the groups, there was significant genetic differentiation between the clades (F_{ST} values greater than 0.11) (Supplementary Figure S4; Supplementary Table S5). In addition, IBD analysis revealed a significant positive correlation between genetic differences and geographic distance ($P < 0.001$) (Figure 2B).

The results of *FIS* revealed that flaMSJJ had the highest *FIS* at the population level, whereas FLAEM had the highest *FIS* at the clade level. At the clade level, the *FIS* decreased in the order of FLAEM (0.8173) > FLAPH (0.6816) > FLAHN (0.6648) > FLASW (0.5711) > FLAYL (0.4508) (Figure 2C; Supplementary Table S6).

Demographic model testing

M1 was determined to be the best-fitting model, and its topology was consistent with that of the ML tree (Supplementary Table S7). Based on an estimate of two generations per year, the common ancestor of FLAEM and FLAPH diverged from the common ancestor of FLAYL, FLASW, and FLAHN approximately 0.75 million years ago (Ma). FLAEM and FLAPH diverged approximately 0.31 Ma, FLAHN diverged from the common ancestor of FLAYL and FLASW approximately 0.56 Ma, and FLAYL and FLASW diverged approximately 0.41 Ma (Figure 2D; Supplementary Table S8). The gene flow between geographically adjacent clades (such as FLAYL&FLASW, FLASW&FLAHN, FLAEM&FLAPH) was relatively high, whereas the gene flow with other clades was relatively low (Figure 2E; Supplementary Table S8).

Historical demographic changes

The stairway plots revealed that the FLAYL populations experienced rapid expansion 1 Ma, and remained stable thereafter. The FLASW populations experienced rapid expansion during the LIG period. The FLAHN population began to shrink approximately 1 Ma, reaching its lowest point 0.4 Ma, followed by rapid expansion, reaching its peak during the LIG period, and then remaining stable. The FLAEM population experienced rapid shrinkage 0.3 Ma, reaching its lowest point 0.2 Ma, followed by rapid expansion and reaching stability during the LIG period. Later, a small

expansion event occurred approximately 0.07 Ma and has remained stable. The FLAPH population experienced a rapid expansion 0.5 Ma, reaching its peak 0.2 Ma, followed by a slow contraction. After the LGM period, a sudden and intense contraction occurred, until 0.007 Ma, when it reached its lowest point and then remained stable (Figure 2F).

Estimated effective migration surfaces

EEMS contours illustrate relatively high or low effective migration within the main distribution area of *M. flavidipes* and its neighbouring regions. The results of the effective migration area revealed the following obvious barriers to gene exchange within its distribution area: barrier 1: high-altitude areas in the Ailao Mountains–Honghe River Fault Shear Zone; barrier 2: the Qiongzhou Strait and the Beibu Gulf; barrier 3: the Isthmus of Kra; and barrier 4: the SCS and the Sulu Sea. These potential barriers also showed high posterior probabilities > 0.90 in the Bayesian estimation of the migration parameters. Notably, among the results with high-posterior probabilities > 0.90, there was a significant barrier fracture between geographical barriers 1 and 2, and there is an obvious communication channel nearby (Figures 3A &3B).

Species distribution modeling

During the LIG period, suitable habitats were basically consistent with the current distribution. With the start of the LGM period, suitable habitats expanded to the surrounding areas, while large connected land bridges emerged. Moreover, the exposed continental shelf connecting islands and continents provided a large area of suitable habitat. In the middle Holocene, the suitable habitat significantly shifted northwards, and there was almost no suitable habitat in the southern region of the Malay Peninsula. Moreover, there was no suitable habitat on Kalimantan Island. Under current conditions, the range of the suitable habitat was similar to that of suitable habitat in the LIG period (Figures 3C-3F).

Niche comparisons analysis

In terms of the similarity in Niche comparisons, from the highest to the lowest, FLAYL and FLASW have the highest Schoener's D niche similarity value (0.175) (Figure 4D). The Schoener's D niche similarity value between FLAYL and FLAPH is 0.07(Figure 4C). The Schoener's D niche similarity values between FLAYL and both FLAEM and FLAHN are 0.022 (Figures 4A &4B). The Schoener's D niche similarity values between the remaining combinations were 0 (Figure 4E). Pearson correlation analysis showed that BIO2 and BIO7 were significantly positively correlated with genetic diversity (BIO2: $r = 0.6491$, $P<0.001$; BIO7: $r = 0.4343$, $P<0.05$) (Figure 4F), while nucleotide diversity exhibited a significant negative correlation with longitude ($r = -0.869$; $P<0.001$) (Figures 2A & S5; Supplementary Table S9).

DISCUSSION

Ring divergent phylogeographic patterns and historical diffusion routes in opposite directions around the South China Sea

A clear ring divergent phylogeographic pattern has been revealed in *M. flavidipes* around the SCS (Figure 1A). Each island population converged into an independent group (FLAPH, FLAHN, and FLAEM), reflecting that the ocean and straits are important geographical barriers for population differentiation and play important roles in geographic isolation for terrestrial organisms (Figures 1 & 3B) (Lim et al., 2020; Liu et al., 2020). Among the mainland populations, FLAYL and FLASW formed obvious population differences, with the Ailao Shan–Red River shear zone as the geographical barrier (Figures 3A &3B). The Ailao Shan–Red River shear zone is composed of a crustal discontinuity over 1800 kilometres long and extends from the southeastern edge of the Qinghai-Tibet Plateau to the SCS forming an important geographical barrier to the local biota (Che et al., 2010; Liu & Fu, 2020).

Ancestral populations tend to accumulate relatively high levels of genetic variation, and the degree of genetic variation gradually decreases along the diffusion path due to

399 the founder effect (Qiao et al., 2018; Soderquist et al., 2025). The FLAYL populations,
400 especially the populations from the Indo-China Peninsula, had relatively high genetic
401 diversity values, and the diversity centre of the Malcidae species was located precisely
402 on the southeastern edge of the Qinghai-Tibet Plateau (Li et al., 2022). Therefore, we
403 speculated that the ancestral populations of *M. flavidipes* should be located in the Indo-
404 China Peninsula.

405 According to the genetic diversity values (a significant decrease in genetic
406 diversity from west to east), IBD and MSN result (Figures 1D & 2B), the ancestral
407 population diffused on the land around the SCS probably through two routes in opposite
408 directions: southwards and northwards.

409 On the northwards diffusion route, the ancestral population first spread northwards
410 to the southern part of mainland China and then spread from west to east within this
411 region. The population in southern China subsequently crossed the Qiongzhou Strait
412 and spread to Hainan Island. The population dispersal process along this pathway
413 perfectly aligns with the founder effect and newly colonized populations experience a
414 certain degree of genetic bottleneck during their differentiation from the original
415 population. After the population dispersed to Hainan Island, gene flow with the
416 mainland population was blocked as land bridges disappeared, and the environmental
417 differences on the island further accelerated the differentiation rate of the FLAHN
418 (Figures 4B & 4D). In contrast, the terrestrial isolation barrier between the FLAYL and
419 FLASW is weaker than the marine barrier, and the degree of environmental difference
420 between the two regions is relatively low. Therefore, the differentiation time of these
421 two populations is later than that of the FLAHN.

422 On the southwards diffusion line, the ancestral populations first spread southwards
423 to the Malay Peninsula region. These populations then spread from west to east along
424 the island chain around the SCS, successively spreading to the Kalimantan Island region
425 and the Philippine Island region. Notably, although the Philippine Islands geologically
426 should be the final stop (along the dispersal route), their populations exhibit
427 significantly greater genetic diversity than those on Kalimantan Island do. This paradox

suggests that the two island populations may have undergone more complex population dynamics following their initial divergence and dispersal.

In addition, the divergence times among island populations of *M. flavidipes* are concentrated between 0.3 and 0.8 million years ago, a period that coincides with the Middle Pleistocene (0.13 – 0.8 million years ago) when glacial activity was intense. Previous studies have confirmed that during the Middle Pleistocene, sea levels fluctuated repeatedly with glacial cycles (De Boer et al., 2014; Klicka & Zink, 1997). The periodic land bridges formed by these fluctuations provided critical conditions for the transmarine dispersal of various biological groups in East and Southeast Asia, including snakes, plants and insects, to islands (Guo et al., 2019; Liang et al., 2024; Liu et al., 2020). Unlike some plants whose fruits and seeds can float on the sea surface, enabling strong long-distance dispersal (Yang & Ren, 2023) and insect groups that accomplish transmarine migration via monsoons (Yang et al., 2022), *M. flavidipes* has weak flight ability and cannot cross marine barriers by flying. Therefore, the transmarine dispersal of its populations is more likely to rely on the emergence of land bridges. This result supports the previously proposed hypothesis of the "dispersal mode driven by the periodic changes of land bridges in the SCS region".

Genetic-ecological correlations and evolutionary trajectories in island populations

Island populations, due to their smaller population sizes, are more prone to the risks of inbreeding and loss of genetic diversity (Dierickx et al., 2020). Compared with the mainland populations (FLASW and FLAYL), the three island populations (FLAEM, FLAPH, FLAHN) all exhibit a typical pattern of "high *FIS* and low genetic diversity" in terms of genetic characteristics. Meanwhile, the population historical dynamics shows that the mainland populations exhibited a pattern of "first increasing to a peak and then remaining stable", while the effective population sizes of the island populations experienced multiple fluctuations during the Middle Pleistocene. It is directly associated with the dispersal and differentiation events of island populations

during this period. Meanwhile, island populations belong to late-colonizing groups with restricted genetic pools and are constrained by the narrow living ranges of island ecosystems; this makes them more sensitive to climatic fluctuations and exhibits stronger genetic fragility. Therefore, climatic fluctuations in the Middle Pleistocene not only provided opportunities for transoceanic dispersal of island populations but also posed challenges to their survival in coping with climatic shocks due to the inherent fragility of their genetic structure.

Notably, although the three island populations share the same genetic characteristic pattern, their genetic traits still exhibit uniqueness, which is closely related to the different geographical locations of the three islands and their evolutionary histories.

The Hainan population (FLAHN) diverged from its ancestor 550 000 years ago (the second stage of the third last glacial period) (Yi et al., 2005). During this period, the effective population size experienced a brief contraction, followed by a rapid increase to a stable state, which is consistent with the typical founder effect model. However, it is worth noting that during the glacial period, the exposure of the continental shelf connected Hainan Island both with southern China and the Beibu Gulf. However, why did the Hainan population show a dispersal pattern of crossing the Qiongzhou Strait via southern China? (The MSN results show that the genetic distance between the FLAHN and GDTS geographical populations is the closest. Moreover, the structure results also indicate that there is no genetic admixture between FLAHN and FLAYL, while there is a strong gene exchange signal between FLAHN and FLASW.) Instead of a population dispersal event where FLAYL crossed the Beibu Gulf from the Indochina Peninsula to Hainan. In terms of topography, the shortest straight - line distance between Hainan Island and the Leizhou Peninsula in southern China is merely approximately 18 kilometres, while the shortest distance between Hainan Island and the Indochina Peninsula is nearly 200 kilometres. Meanwhile, the ecological niche results from the Last Glacial Maximum indicate that the suitable habitats on the land bridge formed in the Beibu Gulf area are fragmented, while the land bridge connecting the Leizhou Peninsula has a connected suitable habitat corridor. Therefore, we believe that the advantage of geographical distance and the continuous suitable habitats on the land

bridge probably together facilitated the dispersal of FLASW to Hainan Island and the subsequent genetic exchange. Thus, it can be seen that land bridges do provide an opportunity for populations to cross marine barriers. However, not all geographical populations adjacent to land bridges can successfully disperse. The ecological corridors within the suitable habitats on the land bridges determine the "lucky ones" that can successfully disperse.

Meanwhile, we found that the Hainan population exhibited the lowest *FIS* compared to other island populations. Geographically, Hainan Island is the closest to the mainland, enabling frequent gene flow with mainland populations (Stanback et al., 2022). This finding aligns with the "Island Equilibrium Theory", which posits a negative correlation between island isolation and immigration rate from the mainland. The land-island adjacency pattern protects island populations from genetic vulnerability to some extent by facilitating the influx of mainland individuals (Beaugrand et al., 2024; Losos & Schluter, 2000; Schoville et al., 2013).

However, the other two island populations, FLAEM and FLAPH, showed contradictions with theoretical expectations in multiple genetic characteristic comparisons: ① They exhibited genetic traits conflicting with the founder effect (see 4.1 for details); ② FLAEM showed an abnormal combination of the lowest genetic diversity but a high effective population size. According to the neutral theory ($\pi = 4Ne\mu$), genetic diversity should be positively correlated with effective population size when the mutation rate is constant. Therefore, it can be inferred that the two populations have experienced a more complex population evolutionary history after initial differentiation.

A correlation analysis between genetic diversity and environmental factors revealed a strong positive correlation between the genetic diversity of FLAEM and FLAPH and the ecological factors Bio2 (diurnal temperature range) and Bio7 (annual temperature range) (Figure 2A & Supplementary Figure S5, Supplementary Table S9).

Given the critical impact of temperature on the growth, reproduction, and other developmental processes of insects, we hypothesize that environmental selective pressures may further drive directional adaptive evolution in the two populations. Meanwhile, FLAEM experienced a severe genetic bottleneck event 0.2-0.3 million years ago (Ma), which likely led to homogenization of its gene pool. Additionally, ENM revealed that suitable habitats in eastern Malaysia vanished from the Last Glacial Maximum to the Mid-Holocene, yet population size remained stable during this period. In contrast, the effective population size of FLAPH declined rapidly despite its extensive suitable habitat area, with a certain degree of gene flow existing between the two populations. It is hypothesized that FLAEM was forced to migrate northwards for gene exchange with FLAPH during this period, then returned as habitats expanded southward. Thus, FLAPH probably enhanced genetic diversity via the "refuge effect", while FLAEM experienced genetic drift again amid habitat shifts, leading to a further reduction in its genetic diversity (Leite et al., 2016; Mikkelsen et al., 2025). Therefore, for FLAEM, although the effective population size has recovered, repeated genetic drift, and inbreeding constraints (FLAEM has the highest *FIS*), and environmental selection pressure may make it difficult for FLAEM's genetic diversity to increase effectively (Costanzi & Steifetten, 2019).

Moreover, it is noteworthy that during the differentiation of island populations, FLAHN and FLAEM experienced genetic bottlenecks 500 000 and 300 000 years ago, respectively, while the effective population size of FLAPH increased rather than decreased; this indicates that its geographical location might have provided a more

stable and suitable habitat, buffering the impact of climatic fluctuations on the population.

In summary, island populations exhibit complex responses to climatic fluctuations and environmental pressures since the Middle Pleistocene, attributed to their small population size, fragile genetic structure, and geographic isolation. Meanwhile, their genetic characteristics are closely linked to geographic location, dispersal history, and environmental pressures: gene flow from adjacent continents can partially mitigate genetic vulnerability, whereas far-island populations develop more complex population dynamic evolutionary histories due to high geographic isolation, pronounced environmental selection pressures, and sensitivity to climatic oscillations.

The conduct of our study also provides important insights for the conservation of SCS islands. As fragile regions significantly impacted by climate change and habitat fragmentation, prioritizing the protection of offshore oceanic islands is essential. For instance, establish long-term genetic monitoring mechanisms to real-time track changes in population genetic structure, preventing further loss of genetic diversity caused by inbreeding depression or environmental selection pressure. Our study found that the Philippine Islands, with their stable habitats, serve as important refuges for populations responding to climate change. Therefore, in the conservation of SCS islands, SDMs can be used to predict potential refuges under future climate warming scenarios, and core protected areas within islands can be demarcated to further enhance the resilience of island populations in responding to long-term climate fluctuations. As key components of island ecosystems, insects' conservation can drive the stability of the entire ecosystem. They can be employed as important evaluation indicators to regularly monitor the effectiveness of conservation measures through parameters such as population density and changes in genetic diversity, thereby dynamically optimizing

conservation plans and providing a scientific basis for the sustainable management of SCS island ecosystems.

Science Press
ZR Zoological Research

FOUNDATION ITEMS: This project is supported by the National Natural Science Foundation of China (No. 32130014 and 32300369).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

AUTHORS' CONTRIBUTIONS

W.J.B., and S.J.W. conceived and designed the study. Y.F.L. conducted the experimental design, data analysis and prepared the manuscript. M.Q.G., J.H. C., and D.Q.A. performed the gene expression and population genomic analyses with input from Z.H.L. and J.Y.Z. collected the samples. Z.Y. and H.J.X. interpreted the results. All the authors contributed to the manuscript and approved the submitted version.

COMPETING INTERESTS

The authors declare no conflicts of interest.

DATA AVAILABILITY

Raw sequence reads are deposited in the SRA: BioProject PRJNA1028310; GenBank accession nos. SRR26408936-SRR26409111.

REFERENCES

Afzali YA, Naderloo R, Keikhosravi A, et al. 2024. Phylogeography of the freshwater crab *Potamon persicum* (Decapoda: Potamidae): an ancestral ring species? *Journal of Heredity*, **115**(3): 277–291.

Beaugrand G, Kleparski L, Luczak C, et al. 2024. A niche-based theory of island biogeography. *Ecology and Evolution*, **14**(6): e11540.

Bouzid NM, Archie JW, Anderson RA, et al. 2022. Evidence for ephemeral ring species formation during the diversification history of western fence lizards (*Sceloporus occidentalis*). *Molecular Ecology*, **31**(2): 620–631.

Brown JL & Carnaval AC. 2019. A tale of two niches: methods, concepts, and evolution. *Frontiers of Biogeography*, **11**(4): e44158.

Che J, Zhou W, Hu J, et al. 2010. Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(31): 13765–13770.

Chiang TY, Chiang YC, Chen YJ, et al. 2001. Phylogeography of *Kandelia candel* in East Asiatic mangroves based on nucleotide variation of chloroplast and mitochondrial DNAs. *Molecular Ecology*, **10**(11): 2697–2710.

Costanzi J-M & Steifetten O. 2019. Island biogeography theory explains the genetic diversity of a fragmented rock ptarmigan (*Lagopus muta*) population. *Ecology and Evolution*, **9**(7): 3837–3849.

Cros E, Chattopadhyay B, Garg KM, et al. 2020. Quaternary land bridges have not been universal conduits of gene flow. *Molecular Ecology*, **29**(14): 2692–2706.

Damaska AF, Konstantinov A, Lee C-F, et al. 2021. Molecular phylogeny of moss-inhabiting flea beetles from the *Chabria* group (Coleoptera: Chrysomelidae: Alticini) reveals multiple colonizations and radiations in Taiwan. *Systematic Entomology*, **46**(4): 915–925.

Danecek P, Auton A, Abecasis G, et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156–2158.

De Boer B, Lourens LJ, Van De Wal RSW. 2014. Persistent 400,000-year variability of Antarctic ice volume and the carbon cycle is revealed throughout the Plio-Pleistocene. *Nature Communications*, **5**: 2999.

Dierickx EG, Sin SYW, Van Veelen HPJ, et al. 2020. Genetic diversity, demographic history and neo-sex chromosomes in the Critically Endangered Raso lark. *Proceedings of the Royal Society B-Biological Sciences*, **287**(1922): 20192613.

Dormann CF, Mcpherson JM, Araujo MB, et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**(5): 609–628.

Dusfour I, Michaux JR, Harbach RE, et al. 2007. Speciation and phylogeography of the Southeast Asian *Anopheles sundaicus* complex. *Infection Genetics and Evolution*, **7**(4): 484–493.

Eaton DaR & Overcast I. 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, **36**(8): 2592–2594.

Esri R. 2020. ArcGIS desktop: release 10.8. Environmental Systems Research Institute.

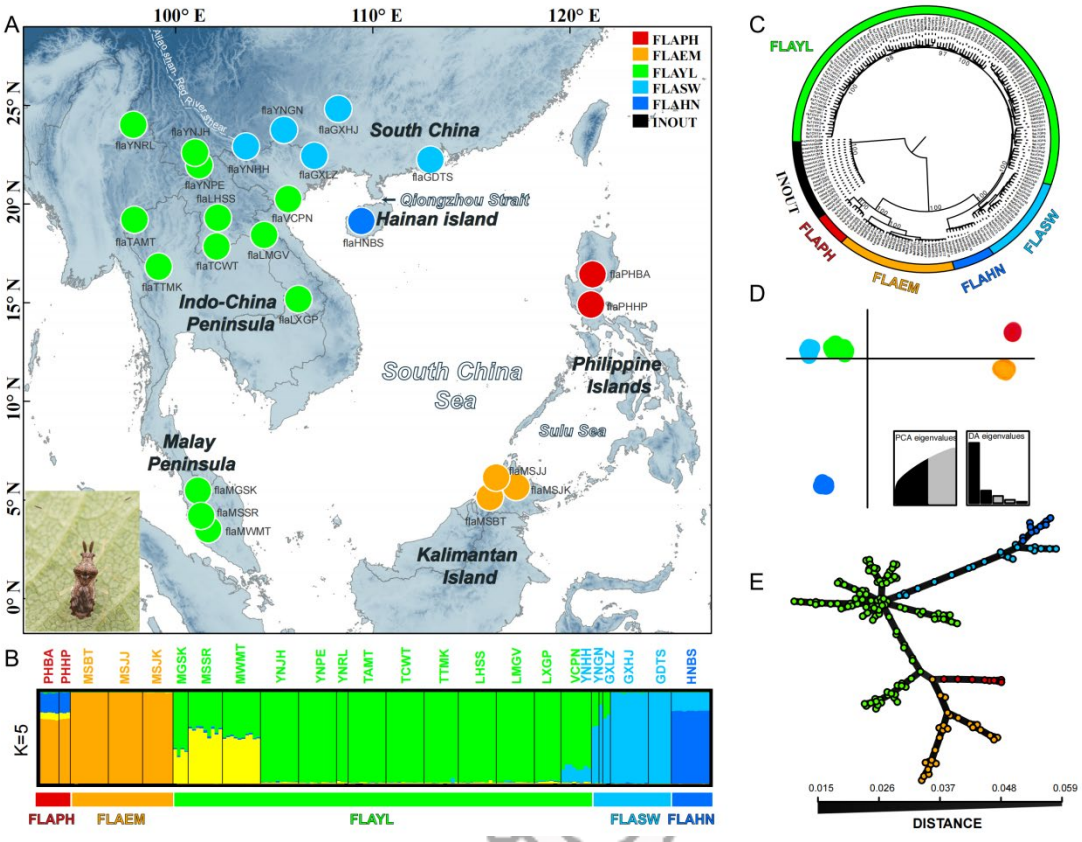
Excoffier L, Dupanloup I, Huerta-Sanchez E, et al. 2013. Robust Demographic Inference from Genomic and SNP Data. *Plos Genetics*, **9**(10): e1003905.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus

genotype data: Linked loci and correlated allele frequencies. *Genetics*, **164**(4): 1567–1587.
Fu S, Chen X, Wang K, et al. 2024. Shared phylogeographic patterns and environmental responses of co-distributed soybean pests: Insights from comparative phylogeographic studies of *Riptortus pedestris* and *Riptortus linearis* in the subtropics of East Asia. *Molecular Phylogenetics and Evolution*, **195**: 108055.
Fuchs J, Ericson PGP, Bonillo C, et al. 2015. The complex phylogeography of the Indo-Malayan *Alophoixus* bulbuls with the description of a putative new ring species complex. *Molecular Ecology*, **24**(21): 5460–5474.
Gorog AJ, Sinaga MH, Engstrom MD. 2004. Vicariance or dispersal?: Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). *Biological Journal of the Linnean Society*, **81**(1): 91–109.
Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, **5**(1): 184–186.
Gray A & Cavers S. 2014. Island Biogeography, the Effects of Taxonomic Effort and the Importance of Island Niche Diversity to Single-Island Endemic Species. *Systematic Biology*, **63**(1): 55–65.
Guo P, Liu Q, Zhu F, et al. 2019. Multilocus phylogeography of the brown-spotted pitviper *Protobothrops mucrosquamatus* (Reptilia: Serpentes: Viperidae) sheds a new light on the diversification pattern in Asia. *Molecular Phylogenetics and Evolution*, **133**: 82–91.
Gyllenhaal EF, Andersen MJ, Moyle RG, et al. 2025. Island size shapes genomic diversity in a great speciator (Aves: *Zosterops*). *Biology Letters*, **21**(3).
Hall R. 2002. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, **20**: 353–431.
Hirschfeld M, Barnett A, Sheaves M, et al. 2023. What Darwin could not see: island formation and historical sea levels shape genetic divergence and island biogeography in a coastal marine species. *Heredity*, **131**(3): 189–200.
Hosoishi S, Maruyama M, Yamane S, et al. 2023. Multilocus phylogeny and historical biogeography of the *Crematogaster inflata*-group (Hymenoptera: Formicidae) in South-East Asia. *Zoological Journal of the Linnean Society*, **198**(3): 901–922.
Huson DH & Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**(2): 254–267.
Irwin DE, Irwin JH, Price TD. 2001. Ring species as bridges between microevolution and speciation. *Genetica*, **112**: 223–243.
Jakobsson M & Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**(14): 1801–1806.
Kaya E, Agca M, Adiguzel F, et al. 2019. Spatial data analysis with R programming for environment. *Human and Ecological Risk Assessment*, **25**(6): 1521–1530.
Keightley PD, Trivedi U, Thomson M, et al. 2009. Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Research*, **19**(7): 1195–1201.
Klicka J & Zink RM. 1997. The importance of recent ice ages in speciation: A failed paradigm. *Science (Washington D C)*, **277**(5332): 1666–1669.
Kuchta SR, Parks DS, Mueller RL, et al. 2009. Closing the ring: historical biogeography of the

666 salamander ring species *Ensatina eschscholtzii*. *Journal of Biogeography*, **36**(5): 982–995.
667 Lavery TH, Decicco LH, Olson KV, et al. 2023. Phylogeography of Solomon Islands blossom bats
668 reflects oceanic divides and Pleistocene connections. *Journal of Biogeography*, **50**(5): 920–
669 931.
670 Leite YLR, Costa LP, Loss AC, et al. 2016. Neotropical forest expansion during the last glacial
671 period challenges refuge hypothesis. *Proceedings of the National Academy of Sciences of the*
672 *United States of America*, **113**(4): 1008–1013.
673 Li Y, Wang S, Zhou J, et al. 2022. Integrative taxonomy of the stalk-eyed bug genus *Chauliops*
674 (Heteroptera: Malcidae: Chauliopinae) reveals orogeny-driven speciation. *Journal of*
675 *Systematics and Evolution*, **61**(5): 932–947.
676 Liang Y, Du S-J, Zhong Y-J, et al. 2024. Molecular phylogeny and identification of agromyzid
677 leafminers in China, with a focus on the worldwide genus *Liriomyza* (Diptera: Agromyzidae).
678 *Journal of Integrative Agriculture*, **23**(2): 739–739.
679 Lim HC, Shakya SB, Harvey MG, et al. 2020. Opening the door to greater phylogeographic
680 inference in Southeast Asia: Comparative genomic study of five codistributed rainforest bird
681 species using target capture and historical DNA. *Ecology and Evolution*, **10**(7): 3222–3247.
682 Liu F-GR, Tsaur S-C, Huang H-T. 2015. Biogeography of *Drosophila* (Diptera: Drosophilidae) in
683 East and Southeast Asia. *Journal of Insect Science*, **15**: 69.
684 Liu X-M & Fu Y-X. 2020. Stairway Plot 2: demographic history inference with folded SNP
685 frequency spectra. *Genome Biology*, **21**(1): 280.
686 Liu Y, Hong Thai P, He Z, et al. 2020. Phylogeography of the cicada Platypleura hilpa in subtropical
687 and tropical East Asia based on mitochondrial and nuclear genes and microsatellite markers.
688 *International Journal of Biological Macromolecules*, **151**: 529–544.
689 Losos JB & Schluter D. 2000. Analysis of an evolutionary species-area relationship. *Nature*,
690 **408**(6814): 847–850.
691 Mikkelsen E, Lavareda D, Vallinoto M, et al. 2025. Phylogeography of the Red-Headed Manakin
692 Supports the River-Refuge Hypothesis. *Journal of Biogeography*, **52**(1): 92–107.
693 Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: New Models and Efficient Methods
694 for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, **37**(5):
695 1530–1534.
696 Myers N, Mittermeier RA, Mittermeier CG, et al. 2000. Biodiversity hotspots for conservation
697 priorities. *Nature*, **403**(6772): 853–858.
698 Peterson BK, Weber JN, Kay EH, et al. 2012. Double Digest RADseq: an inexpensive method for
699 de novo SNP discovery and genotyping in model and non-model species. *Plos One*, **7**(5):
700 e37135.
701 Petkova D, Novembre J, Stephens M. 2016. Visualizing spatial population structure with estimated
702 effective migration surfaces. *Nature Genetics*, **48**(1): 94–100.
703 Qiao L, Wen G, Qi Y, et al. 2018. Evolutionary melting pots and reproductive isolation: A ring-
704 shaped diversification of an odorous frog (*Odorrana margaratea*) around the Sichuan Basin.
705 *Molecular Ecology*, **27**(23): 4888–4900.
706 Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure.
707 *Molecular Ecology Notes*, **4**(1): 137–138.
708 Schoville SD, Uchifune T, Machida R. 2013. Colliding fragment islands transport independent
709 lineages of endemic rock-crawlers (Grylloblattodea: Grylloblattidae) in the Japanese

archipelago. *Molecular Phylogenetics and Evolution*, **66**(3): 915–927.
Schutze MK, Krosch MN, Armstrong KF, et al. 2012. Population structure of *Bactrocera dorsalis s.s.*, *B. papayae* and *B. philippinensis* (Diptera: Tephritidae) in southeast Asia: evidence for a single species hypothesis using mitochondrial DNA and wing-shape data. *Bmc Evolutionary Biology*, **12**: 130.
Singham GV, Othman AS, Lee C-Y. 2017. Phylogeography of the termite *Macrotermes gilvus* and insight into ancient dispersal corridors in Pleistocene Southeast Asia. *Plos One*, **12**(11): e0186690.
Soderquist L, Joffard N, Scofield DG, et al. 2025. Genetic structure and diversity of the declining orchid *Gymnadenia conopsea* in Scandinavia: implications for conservation and management. *Ecography*, **7**: e07628.
Stanback M, Vaughan G, Libke Z. 2022. Nesting of Carolina Chickadees (*Poecile carolinensis*) on islands in a large freshwater lake not significantly influenced by island size or distance to mainland. *Wilson Journal of Ornithology*, **134**(3): 541–545.
Stéphane D & Anne-Béatrice D. 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, **22**(4): 1–20.
Team RC. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Thörn F, Rödin-Mörch P, Cortazar-Chinarro M, et al. 2021. The effects of drift and selection on latitudinal genetic variation in Scandinavian common toads (Bufo bufo) following postglacial recolonisation. *Heredity*, **126**(4): 656–667.
Tian E, Nason JD, Machado CA, et al. 2015. Lack of genetic isolation by distance, similar genetic structuring but different demographic histories in a fig-pollinating wasp mutualism. *Molecular Ecology*, **24**(23): 5976–5991.
Valente L, Phillimore AB, Etienne RS. 2018. Using molecular phylogenies in island biogeography: it's about time. *Ecography*, **41**(10): 1684–1686.
Wang S & Bu W. 2020. A key to species of genus *Malcus* Stål, with descriptions of four new species from China (Hemiptera: Haeroptera: Malcidae). *Zootaxa*, **4759**(1): 31–48.
Whittaker RJ, Maria Fernandez-Palacios J, Matthews TJ, et al. 2017. Island biogeography: Taking the long view of nature's laboratories. *Science*, **357**(6354): 885–892.
Yang SJ, Bao YX, Zheng XF, et al. 2022. Effect of the Asian monsoon on the northward migration of the brown planthopper to northern South China. *Ecosphere*, **13**(10).
Yang X & Ren M. 2023. Species distribution pattern and formation mechanism of mangrove plants around the South China Sea. *Chinese Journal of Plant Ecology*, **47**(8): 1105–1115.
Yang Y, Li J, Yang S, et al. 2017. Effects of Pleistocene sea-level fluctuations on mangrove population dynamics: a lesson from *Sonneratia alba*. *Bmc Evolutionary Biology*, **17**: 22.
Yi C, Cui Z, Xiong H. 2005. Numerical periods of quaternary glaciations in China. *Quaternary Sciences*, **5**: 609–619.



749

750 **FIGURE 1** Phylogeographic analysis results based on ddRAD data of *Malcus*

751 *flavidipes* ring distribution around the South China Sea. A. Geographical location of

752 sampling, different colored dots represent five clades marked in the panel (Red-FLAPH,

753 Orange-FLAWM, Yellow-FLAMM, Cyan-FLASW, Blue-FLAHN). Bottom-left panel:

754 Ecological photograph of *Malcus flavidipes* (provided by Fan Gao, taken in Wuzhishan

755 City, Hainan Province on January 26, 2022). B. Structure result with best K value (K=5).

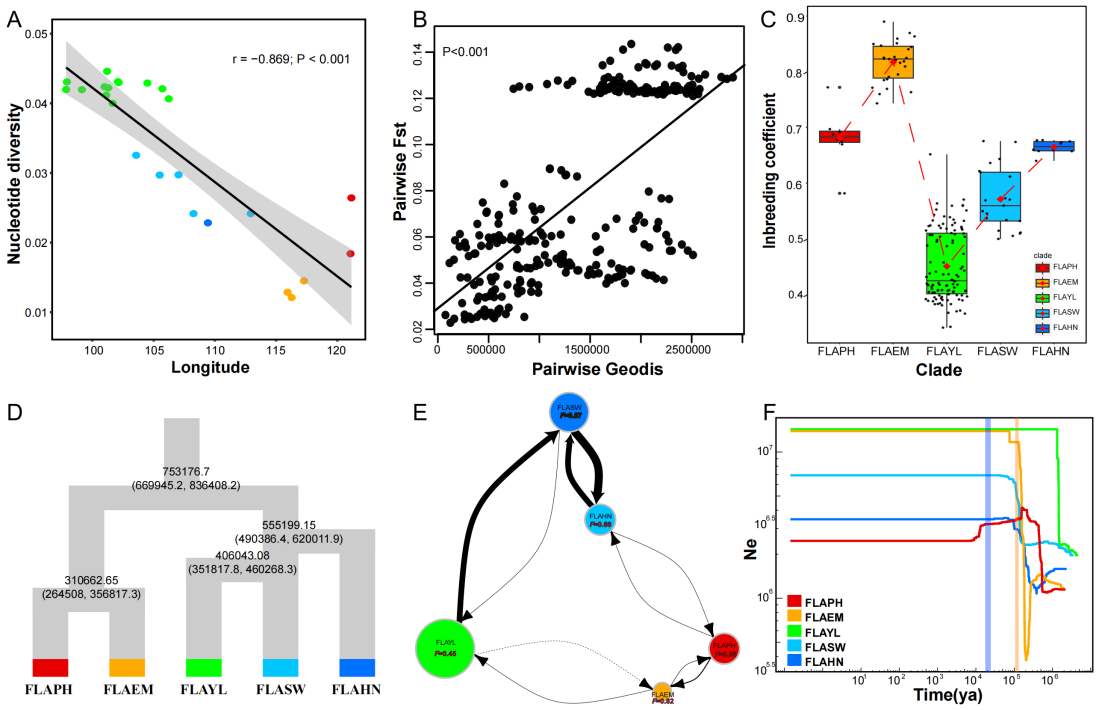
756 C. Phylogenetic topology based on neighbor-net method. D. The discriminant analysis

757 of principal components (DAPC) results. E. Minimum spanning networks of five clades

758 (Red-FLAPH, Orange-FLAWM, Yellow-FLAMM, Cyan-FLASW, Blue-FLAHN).

759

760



761

762 **FIGURE 2** The population genetic diversity and differentiation based on ddRAD data

763 among the five clades of *Malcus flavidipes*. A. The correlation between Nucleotide

764 diversity and Longitude, different colored dots represent five clades. B. The correlation

765 between genetic differentiation (Pairwise F_{ST}) and geographic distance (Pairwise

766 Geodis). C. The within-population inbreeding coefficient of each clade (The black dots

767 represent the within-population inbreeding coefficien values of different populations

768 within the clade; the red diamonds indicate the mean values, and the red dashed lines

769 reflect the changes in the mean values within each lineage). D. The historical divergent

770 time based on the best-fit demographic model. E. Gene flow and within-population

771 inbreeding coefficient among the five genetic clades. F. Historical demographic

772 changes in the five clades inferred from stairway plots.

773

774

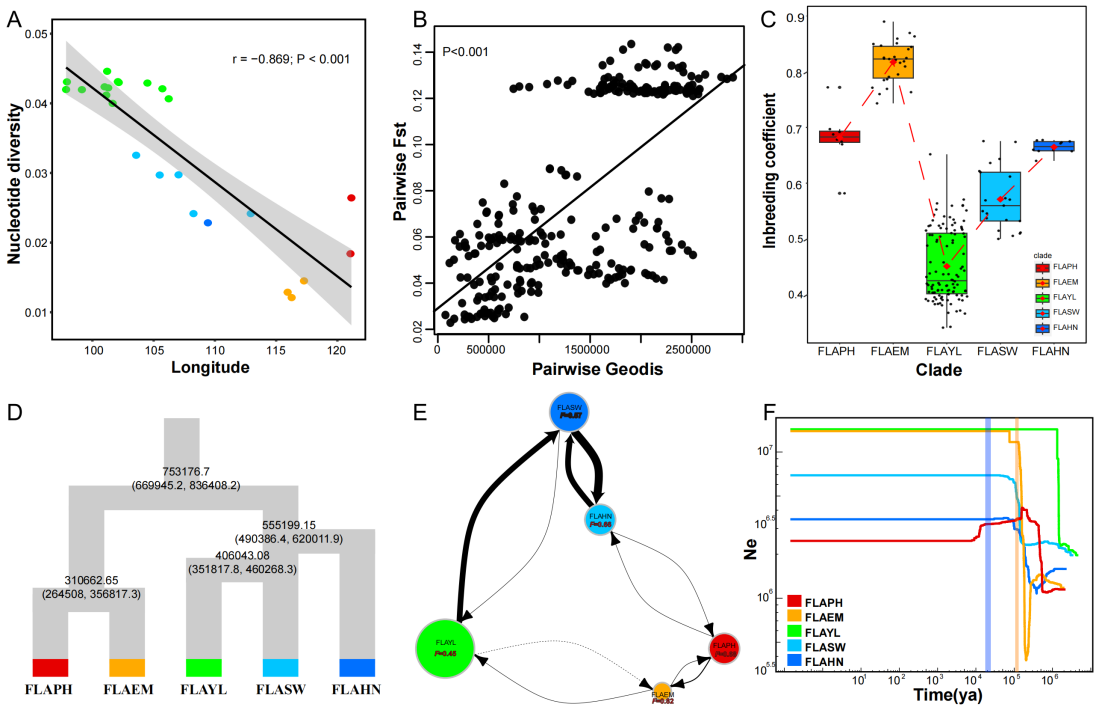


FIGURE 3 The estimated effective migration surfaces and predicted distribution based on the species distribution modeling of *Malcus flavidipes*. A. The contours illustrating relative migration (the threshold value is 0.1). Each circle is geographical projection of the STRUCTURE results (K = 5). B. The contours illustrating relative migration (the threshold value is 0.9). Each circle represents a geographical population, and the size of the circle is proportional to the number of samples in the geographical population. (C-F) Predicted distribution based on the species distribution modeling, different color indicates areas of habitat suitability during different periods. Green: current; purple: the Mid-Holocene period; blue: the last glacial maximum (LGM) period; red-the last interglacial (LIG) period.

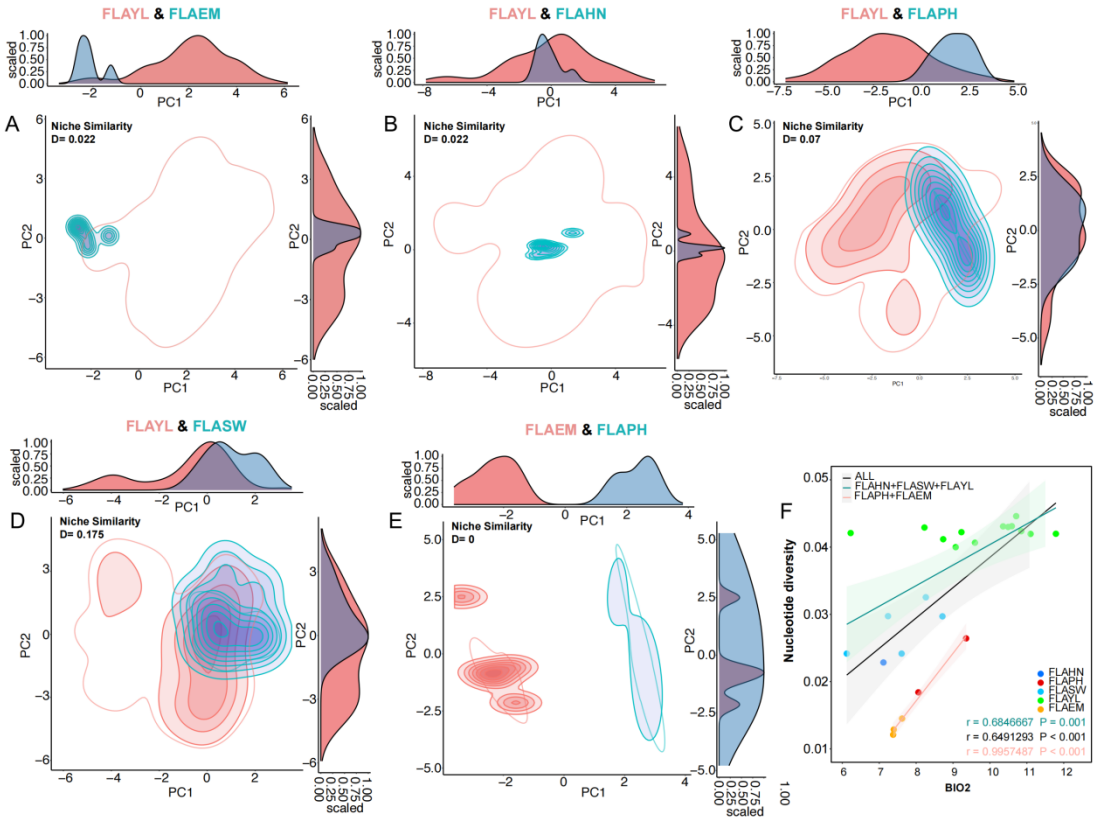


FIGURE 4 (A-E) The niche similarity between genetic clades. The Histogram Density Plots on the top and right showed the density of E-space for different PCs of each clade. Filled Kernel Density Isopleths: Lines representing the kernel density isopleths from 1-100% kernel densities. Empty Kernel Density Isopleths: Lines representing the 1% kernel density isopleths of the environments (not species) depicting outside boundaries of E space. (F) The correlation between Nucleotide diversity and Mean Diurnal Range (BIO2), different colored dots represent five clades (Red-FLAPH, Orange-FLAWM, Yellow-FLAMM, Cyan-FLASW, Blue-FLAHN).

中国南海地区环状分布黄足束长蝽（异翅目：束长蝽科）的系统地理学历史——受定向扩散与岛屿适应性影响

李艳飞[1,#]，郭美琪[1,#]，陈菊红[1.2]，艾德强[1]，李子豪[1]，张靖瑶[1]，叶瑱[1]，薛怀君[1]，王树景[1,*]，卜文俊[1,*]

[1] 南开大学生命科学学院昆虫学研究所，天津，300071

[2] 安徽师范大学生命科学学院，芜湖，241002

[#]作者对文章具有相同贡献

[*]通讯作者邮箱: wangsj@nakai.edu.cn (王树景); wenjunbu@nankai.edu.cn (卜文俊).

摘要：

环状分布模式为解析物种或种群分化机制提供了独特视角。中国南海及周边地区的半封闭海域因其岛屿、半岛及大陆边缘的地理特征，成为研究环状分布模式形成及岛屿种群演化动态的理想区域。本研究以分布于南海的黄足束长蝽（*Malcus flavidipes*）为研究对象，基于简化基因组数据与地理分布数据，运用群体遗传学方法及物种分布模型，系统探究其分布格局形成机制及岛屿种群的演化历史。研究结果显示，黄足束长蝽环绕南海形成五个遗传支系，分别位于云南-印度支那半岛、中国南方、海南岛、加里曼丹岛及菲律宾，呈现出显著的谱系地理结构。遗传多样性证据表明，环状分布模式起源于云南-印度支那半岛，在更新世冰期期间，借助海平面下降暴露的陆桥沿两条路径扩散：北向路径从中国南方延伸至海南岛，南向路径则从加里曼丹岛延伸至菲律宾。对比分析发现，岛屿种群遗传多样性显著低于大陆种群，且种群内近交系数更高；不同岛屿种群间的遗传结构差异与岛屿种群距大陆的地理距离、对历史气候波动的响应差异以及岛屿环境的选择压力密切相关。本研究为南海地区物种的环形分化提供了重要实证依据，并为理解地理隔离、遗传漂变与环境选择协同作用下的岛屿种群的进化机制提供了新视角。

关键词：南海地区；环状分布模式；岛屿生物地理学；群体遗传学；简化基因组；更新世

Science Press ZR Zoological Research