Specialization into Host Sea Anemones Impacted Clownfish Demographic Responses to Pleistocene Sea Level Changes

Alberto García Jiménez[1*], Théo Gaboriau[1], Lucy M. Fitzgerald[1], Sara Heim[1], Anna Marcionetti[1], Sarah Schmid[2], Joris Bertrand[3], Glenn Litsios[4], Abigail Shaughnessy[6], Carl Santiago[6], Ploypallin Rangseethampanya[7], Phurinat Ruttanachuchote[7], Wiphawan Aunkhongthong[7], Sittiporn Pengsakun[7], Makamas Sutthacheep[7], Bruno Frédérich[5], Fabio Cortesi[6], Thamasak Yemin[7] and Nicolas Salamin[1]

[1*]Department of Computational Biology, University of Lausanne, Lausanne, Switzerland.
[2]Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland.
[3]Laboratoire Génome et Développement des Plantes, UPVD, Perpignan, France.
[4]InfoFauna, Lausanne, Switzerland.
[5]Laboratory of Evolutionary Ecology, University of Liège, Liège, Belgium.
[6]Queensland Brain Institute, The University of Queensland, Brisbane, Australia.
[7]Marine Biodiversity Research Group, Ramkhamhaeng University, Bangkok, Thailand.

*Corresponding author(s). E-mail(s): agarcia26286@gmail.com;

Abstract

Fluctuating sea levels during the Pleistocene led to habitat loss and fragmentation, impacting the evolutionary trajectories of reef fishes. Species with specialized ecological requirements or habitat preferences, like clownfishes (Amphiprioninae), may have been particularly vulnerable due to their intricate dependence on sea anemones. The diverse host specializations within this group likely contributed distinct responses to sea-level fluctuations, differentially shaping their recent evolutionary histories. Leveraging a comprehensive genomic dataset, we reveal demographic patterns and connectivity dynamics across multiple populations of ten clownfish species under different host specializations. Host-generalist species demonstrated strong resilience to habitat perturbations, while those specialized on single hosts suffered dramatic bottlenecks linked to sea-level fluctuations. Spatial analyses revealed the significant role of oceanic currents in shaping clownfish genetic diversity landscapes. Dispersal barriers were driven by environmental variables, with the Coral Triangle emerging as a hub of genetic diversity. Our results reveal how clownfish associative behavior influences their population dynamics, holding major implications for their conservation such as the need to consider their mutualism with sea anemones, particularly on host-specialists, to ensure their survival in the face of climate threats. These findings extend broader principles of conservation, improving our understanding of species' responses to ecological constraints and environmental changes over evolutionary timescales.

Keywords: clownfish, Sea-level fluctuations, Host specialization, demographic history, population genomics

Introduction

Habitat loss and fragmentation pose significant threats to species biodiversity and population resilience across ecosystems, disrupting natural habitats and connectivity essential for species persistence [1]. While habitat loss negatively affects species biodiversity and population resilience [2, 3], habitat fragmentation effects are more complex and context-dependent [4, 5]. Habitat fragmentation can promote allopatric speciation and trigger diversification processes [6, 7], while limiting gene flow between previously connected populations, leading to loss of genetic diversity and compromised resilience of populations [8, 9]. The outcome depends on factors such as species' dispersal abilities, ecological traits, and habitat fragment configuration [10].

Pleistocene sea-level fluctuations caused a global fragmentation of coral reef habitats [11], resulting in habitat loss and increased reproductive isolation for reef species [12]. These changes are hypothesized to have caused significant population bottlenecks and genetic structuring among various reef fish species [13]. However, the impact of limited dispersal and strong ecological adaptations, such as habitat specialization, remains unexplored. This underscores the ongoing need to improve our understanding of how species with different ecological specializations respond to climate change, both historically and in future scenarios, crucial for developing more effective conservation strategies [14].

The interplay between habitat dynamics and species' ecological adaptations is particularly evident in reef fishes, with clownfishes serving as a compelling example. Clownfishes are known for their unique and tight relationship with sea anemones, which has made them a focal point of marine evolutionary research [15]. This mutualistic relationship is crucial for clownfish survival and reproduction [16], with no observations of clownfishes living independently of sea anemones [17, 18]. Clownfish ecological

adaptations significantly impact their population dynamics and vulnerability to habitat changes [19]. Their strong attachment to their host anemones [16], reflects their heavy reliance on the coral reef ecosystem. This dependency is further emphasized by their relatively short pelagic larval dispersal stage of approximately 15 days [20, 21], with larvae rarely dispersing beyond 27 km [22]. Additionally, clownfishes have a unique social structure where only the two largest individuals within a social group form the breeding pair [23]. These factors make clownfish populations particularly sensitive to habitat disturbances, with their mutualistic relationship playing a crucial role in shaping their biogeography and population dynamics [24, 25].

Clownfish species exhibit a generalist-specialist continuum in host usage, reflecting a complex interplay between ecological adaptation and habitat preference. This gradient is primarily influenced by the habitat preferences of their associated sea anemones, indicating an ecological sorting along environmental gradients [26]. Specialized species form tight associations with *Radianthus magnifica*, *Entacmaea quadricolor*, or members of the *Stichodactyla* genus, while several clownfish species display generalist behavior, capable of inhabiting, socializing, and reproducing across different sea anemone species without clear preference [27]. Variations in sea anemone characteristics create a heterogeneous landscape of resources and protection for clownfish populations [28], likely influencing their ecological and evolutionary trajectories. Despite extensive research on clownfish ecology, the long-term evolutionary impacts of historical habitat changes on their population dynamics remain poorly understood. Quaternary climate fluctuations significantly affected reef communities, but the extent to which clownfishes' unique host associations may have buffered against these environmental changes is unclear. This knowledge gap highlights the need to investigate how clownfish-anemone mutualism has influenced resilience to environmental shifts over time, providing crucial insights for conservation strategies in the face of ongoing climate change and habitat degradation.

We hypothesize that clownfish populations have undergone distinct evolutionary trajectories shaped by their specialized relationships with sea anemone hosts and Pleistocene sea-level fluctuations. To test this, we analyze 382 geo-referenced genomes from ten clownfish species with varying host specializations. Our multifaceted approach investigates the impact of late Pleistocene sea-level fluctuations on clownfish demographic histories using Sequential Markovian Coalescent methods, aiming to reconstruct historical population size changes and identify potential bottlenecks or expansions. Through spatial genomic analyses, we explore the influence of environmental factors on genetic diversity and dispersal patterns, disentangling the contributions of geographical barriers, habitat preferences, and historical climate changes to current population structures. We also examine differences in demographic patterns and connectivity dynamics across host specializations, assessing the influence of host specificity on long-term population dynamics and evolutionary trajectories.

Results

We generated whole-genome sequencing data from 382 geo-referenced samples to analyze the population structure and reconstruct the demographic histories of multiple populations across ten clownfish species with different host-specialization regimes (Figure 1). The sequencing reads were mapped onto the *Amphiprion clarkii* reference genome, with mapping rates ranging from 81.05% to 97.74% (Supplementary Figure S1a). Average sequencing depths varied between 5.82X and 42.21X (Supplementary Figure S1b). Generated species-specific SNP datasets contained between 202,007 and 2,192,819 SNPs (detailed information in Supplementary Material).

Population structure

Phylogenetic relationships within each species were inferred with a GTR + Γ model and 100 bootstraps. These relationships were primarily structured according to geographic locations (Supplementary Figure S3). Some individuals showed greater genetic similarity to those from nearby populations than to con-specifics from their own population, suggesting localized gene flow or potential recent migration events between adjacent populations. Principal Component Analysis supported these findings, clustering individuals by geographic populations. Admixture analyses revealed limited genetic population structure, with the optimal number of genetic clusters (K) consistently lower than the number of geographically defined populations (Supplementary Figure S4). Specialist species showed higher genetic structuring relative to their sampled populations compared to generalists (Supplementary Table S2), highlighting the influence of ecological specialization on the genetic structure of their populations.

Patterns of genetic diversity

We used the Estimated Effective Migration Surfaces *eems* [29] to analyze spatial patterns of genetic diversity and gene flow across species. The method uses a Bayesian approach to estimate posterior probabilities of genetic diversity and migration rates from pairwise genetic dissimilarities in geo-referenced genomic data, allowing for the identification of areas with significantly higher or lower genetic diversity than expected under an isolation-by-distance model.

The spatial distribution of genetic diversity varied among species and regions. Widely distributed generalists, *A. clarkii* and *A. chrysopterus*, exhibited extensive areas of high genetic differentiation within their populations (Figure 2). These species, along with *A. melanopus* and *A. perideraion* specialists, displayed higher-than-expected genetic diversity within the Coral Triangle, contrasting with predictions based on Isolation-by-distance ($P_{\log(q)>0} > 0.9$; Supplementary Figure S5a,c,d and f). In contrast, specialist species, such as *A. akallopisos* in the West Indian Ocean, *A. perideraion* in the Great Barrier Reef, and *Stichodactyla* specialists *A. polymnus* and *A. sandaracinos* in the Coral Triangle, predominantly exhibited lower levels
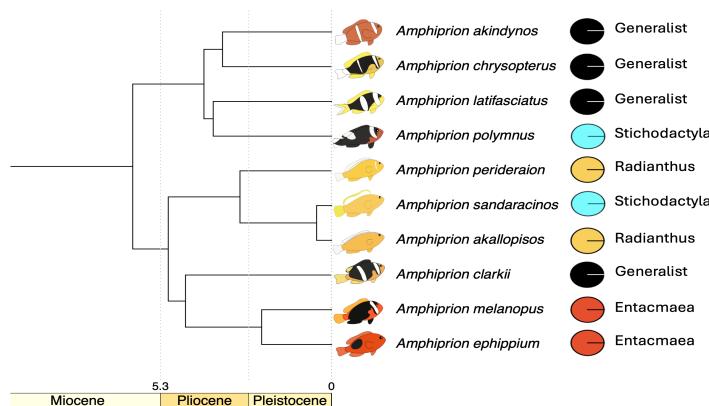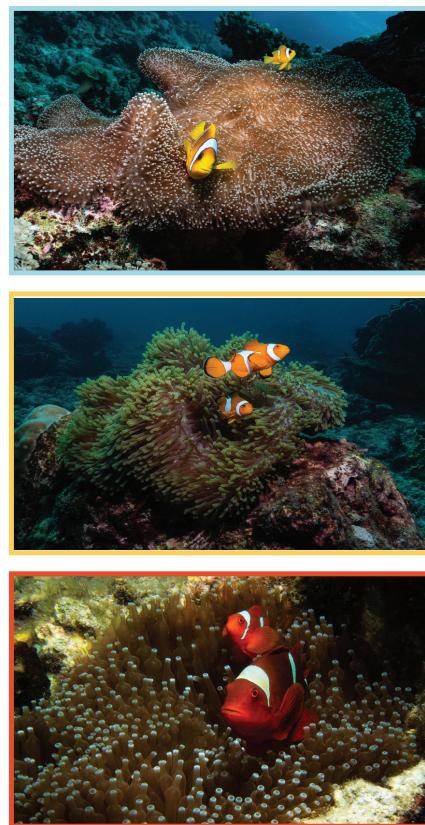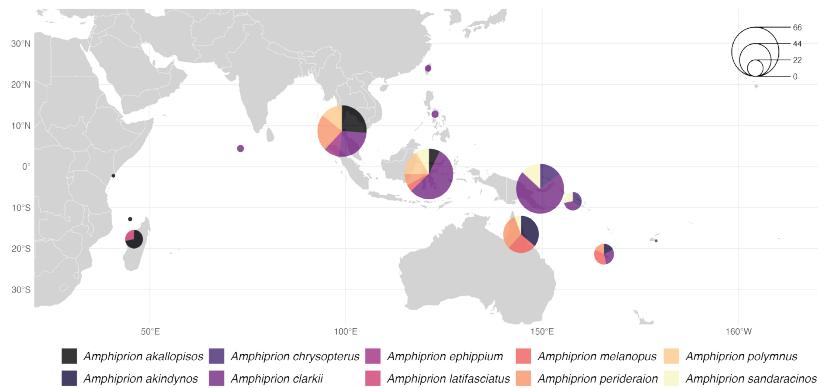
Fig. 1: Species in this study. a) Subset of phylogenetic relationships from [27], showing studied species and host specialization types (colored circles). b) Sea anemone exemplars for each host specialization: *Stichodactyla mertensii* (light blue), *Radianthus magnifica* (orange), and *Entacmaea quadricolor* (red). Credit: Lucy Fitzgerald. c) Sample distribution by species. Pie charts show species proportions at each location; size indicates total sample number (see legend). Colors represent species (indicated below).

of genetic diversity than expected ($P(\log(q) < 0) > 0.9$; Supplementary Figure S5e, f, g and h).

Indian Ocean populations, including *A. clarkii* in the Maldives and Thailand, and *A. akallopisos* in Madagascar, Mayotte, and Kenya, displayed low genetic differentiation compared to expectations under Isolation-by-distance ($P_{\log(q)<0} > 0.9$; Supplementary Figure S5a and e). Similar patterns were observed in the Pacific region, such as *A. chrysopterus* in East Melanesia and Polynesia, and *A. melanopus* and *A. perideraion* in the Great Barrier Reef ($P_{\log(q)<0} > 0.9$; Supplementary Figure S5d, and f). The Gulf of Thailand also exhibited lower genetic diversity than expected in *A. perideraion* and *A. polymnus* populations ($P_{\log(q)<0} > 0.9$; Supplementary Figure S5f, and g).

This analysis revealed diverse spatial patterns of genetic diversity and gene flow across species, highlighting distinct genetic clusters in specialist and generalist species within the Coral Triangle and Indian Ocean regions. We conducted spatial auto-regressive models to test the explanatory power of oceanographic, temperature, nutrient, and water quality variables, along with clownfish and sea anemone species richness, on genetic diversity patterns. Combined explanatory variables better explained genetic diversity distribution across species than single

environmental variables, underscoring the complex interplay shaping population genetic structures in clownfish. Best models included temperature, salinity, and current velocity, while ecological variables, such as species diversity of clownfish and sea anemone, had limited influence (Supplementary Table S4).

Population connectivity and dispersal barriers

eems analyses inferred higher migration rates among demes within the Coral Triangle than expected under a stepping-stone model for multiple species (Figure 3), with notable rates observed in *A. clarkii*, *A. perideraion*, and *A. polymnus* ($P_{\log(m)>0} > 0.9$; Supplementary Figure S6a, f and g). Lower migration rates were identified in regions like Melanesia, the Gulf of Thailand, the Sunda Shelf, the Sahul Shelf, and the area connecting Thailand and East African populations of *A. akallopisos* ($P_{\log(m)<0} > 0.9$; Supplementary Figure S6).

Migration patterns differed between generalist and specialist species. The most widespread and generalist species, *A. clarkii*, displayed the highest migration rates in the Coral Triangle, with values two orders of magnitude higher than those expected under a stepping-stone model (Supplementary Figure 3a). In contrast, specialists
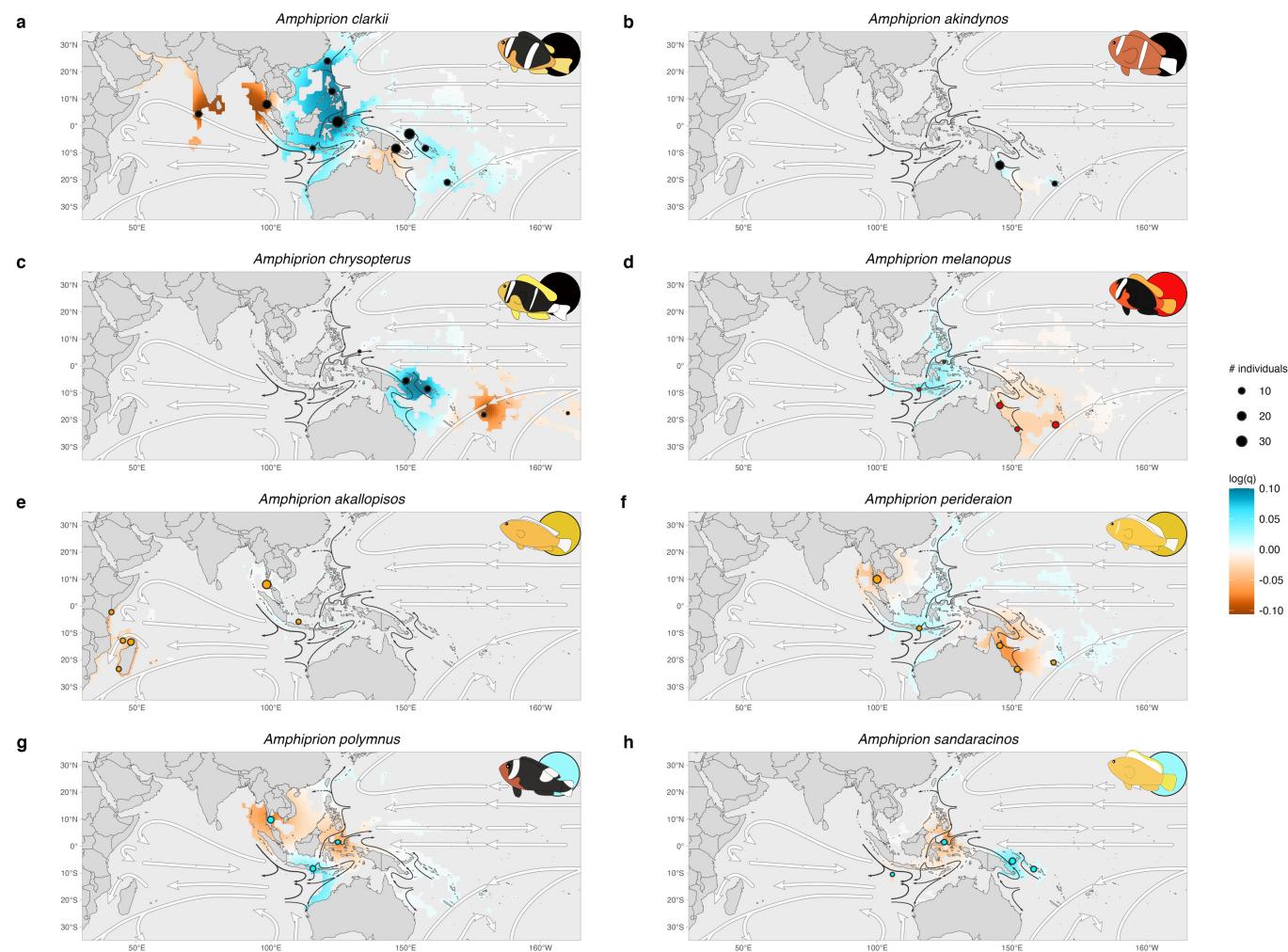
Fig. 2: Spatial patterns of genetic diversity. Maps show the clownfish clade distribution range. Sampling locations are indicated by circles, with size representing the number of individuals, and color the host category (black-generalists, orange-*Radianthus*, red-*Entacmaea*, blue-*Stichodactyla*). Effective genetic diversity surfaces are masked by species distributions. White arrows indicate global oceanic currents, while dark arrows indicate Coral Triangle currents. Values reflect log-transformed genetic diversity rates (blue-higher, red-lower than expected under Isolation-by-Distance). Species names are shown at the top, and vector illustrations and host category on the top-right corner.

exhibited reduced migration rates, with *A. melanopus*, *A. akallopisos* and *A. perideraion* displaying migration rates from 30 to 80 times lower than expected under a stepping-stone process (Figure 3d, e, and f).

Migration rates among clownfish species in the Coral Triangle exceeded expectations based on Isolation-by-distance, while shallow continental shelves and vast bodies of water posed significant geographical barriers to gene flow. The inferred migration patterns are closely linked to the Indonesian Through-Flow and prevailing oceanic currents. Analogous to spatial analysis of genetic diversity patterns, spatial auto-regressive models revealed ocean currents, temperature and water quality as the main variables explaining effective migration rates across most clownfish species. Nevertheless, the most comprehensive model, integrating physical, chemical, and ecological variables, frequently outperformed others, as indicated by the lowest corrected Akaike Information Criterion (Supplementary Table S5).

Demographic history reconstruction

We employed *msmc2* (Multiple Sequentially Markovian Coalescent; [30]) to infer historical changes in effective population size (N_e) from whole-genome sequences. To ensure the robustness of our reconstructions, we generated ten independent *msmc2* runs per species per population, each using three randomly selected individuals.

N_e reconstruction through time revealed distinct demographic patterns between specialist and generalist clownfish species during the Quaternary period. Specialists species exhibited a decrease in N_e, regardless of their host association (Figure 4b). This declines coincided with the trend of decreasing sea level during the late Pleistocene, reaching a minimum approximately 20,000 years ago (Figure 4a). Notably, specialist species (*Entacmaea*, *Radianthus*, and *Stichodactyla* specialists) initially exhibited higher historical N_e than generalists (Figure 4c), with peak values observed between 1 million and 200,000 years ago.
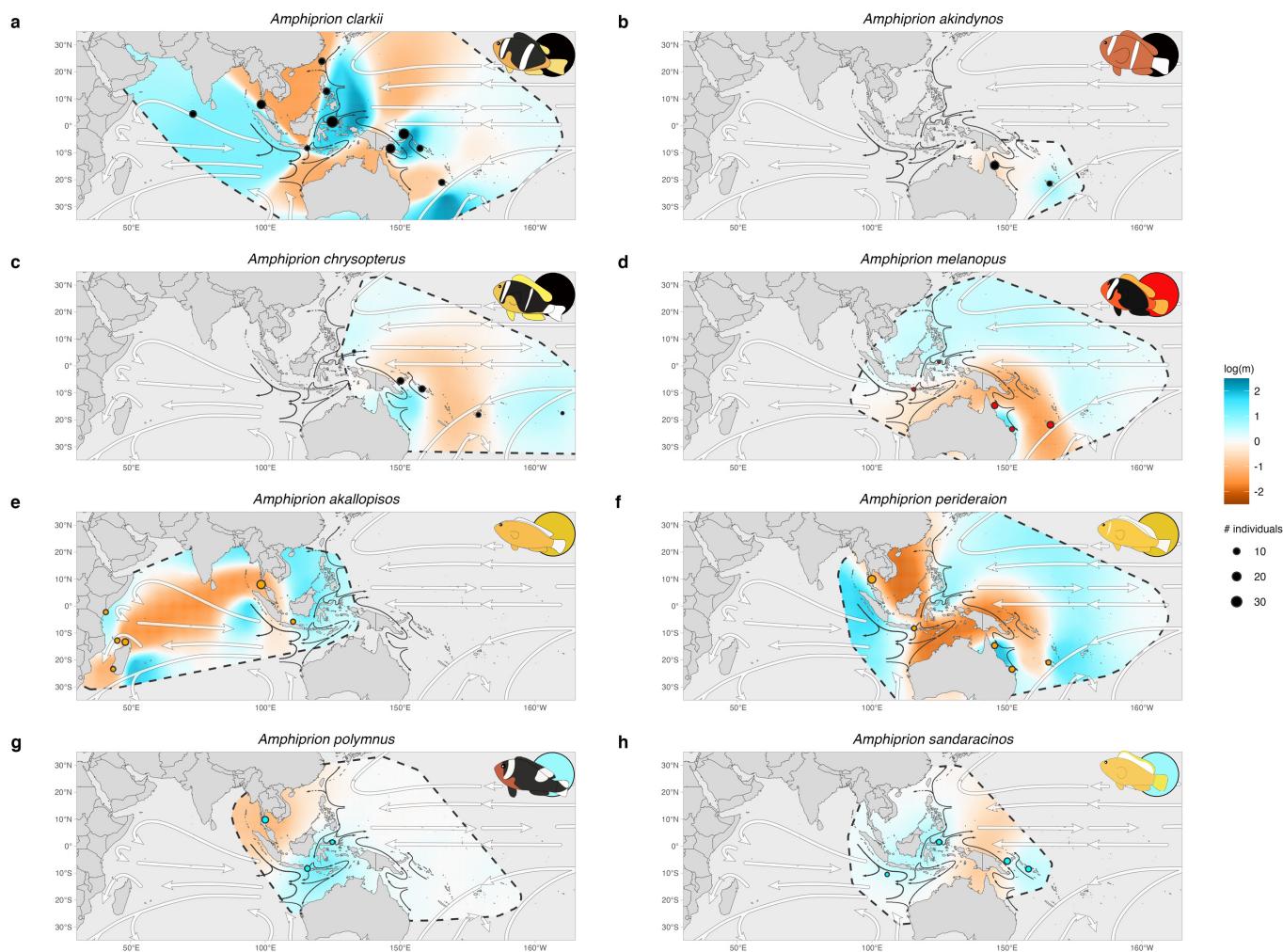
Fig. 3: Spatial patterns of population connectivity. Maps show the clownfish clade distribution range. Sampling locations are indicated by circles, with size representing the number of individuals, and color the host category (black-generalists, orange-*Radianthus*, red-*Entacmaea*, blue-*Stichodactyla*). White arrows indicate global oceanic currents, while dark arrows indicate Coral Triangle currents. Values reflect log-transformed effective migration rates (blue-higher, red-lower than expected under a stepping-stone process). Species names are shown at the top, and vector illustrations and host category on the top-right corner.

While specialist species showed significant declines in N_e estimates over time (*Radianthus* specialists: W = 2278; *p-value* < 0.001; *Entacmaea* specialists: W = 55; *p-value* < 0.001; and *Stichodactyla* specialists: W = 496; *p-value* < 0.001), generalists maintained relatively stable N_e with no significant variations across the periods of sea level change (W = 3665; *p-value* = 0.679; Figure 5a and Supplementary Figure S8a). Generalist species showed an upward trend of N_e starting around 100,000 years ago, with the exception of *A. chrysopterus*. Moreover, N_e dynamics in generalist species appeared more variable, displaying species-specific fluctuations over time compared to the consistent declining trends observed in all specialist species.

The most recent N_e estimates, within the last 1,000 years, revealed critically low values for all specialist species. The most extreme case was observed in the Papua New Guinea population of *A. sandaracinos*, where as few as 177 ± 146 individuals were estimated to contribute to the next generation approximately 180 years ago. An exception is the Indonesian population of *A. melanopus*, with relatively high N_e of $29,182 \pm 2.49$ individuals around 3,000 years ago, comparable to the N_e values of the generalist *A. clarkii* in Indonesia during the same period (Supplementary Figure S11).

Average N_e volatility across 10,000-year periods, a measure of demographic fluctuation intensity, was initially similar across all groups (Figure S9), ranging between 1,250 and 2,500 individuals contributing to the genetic diversity pool. However, following the period of sea level declines, all specialist species exhibited a significant decrease in N_e volatility (*Radianthus* specialists: W = 7, *p-value* < 0.001; *Entacmaea* specialists: W = 158, *p-value* < 0.001; and *Stichodactyla* specialists: W = 10, *p-value* < 0.001). In contrast, populations of generalist species showed significant increases in N_e over time after this period (W = 4678; *p-value* < 0.01; Figure 5b).
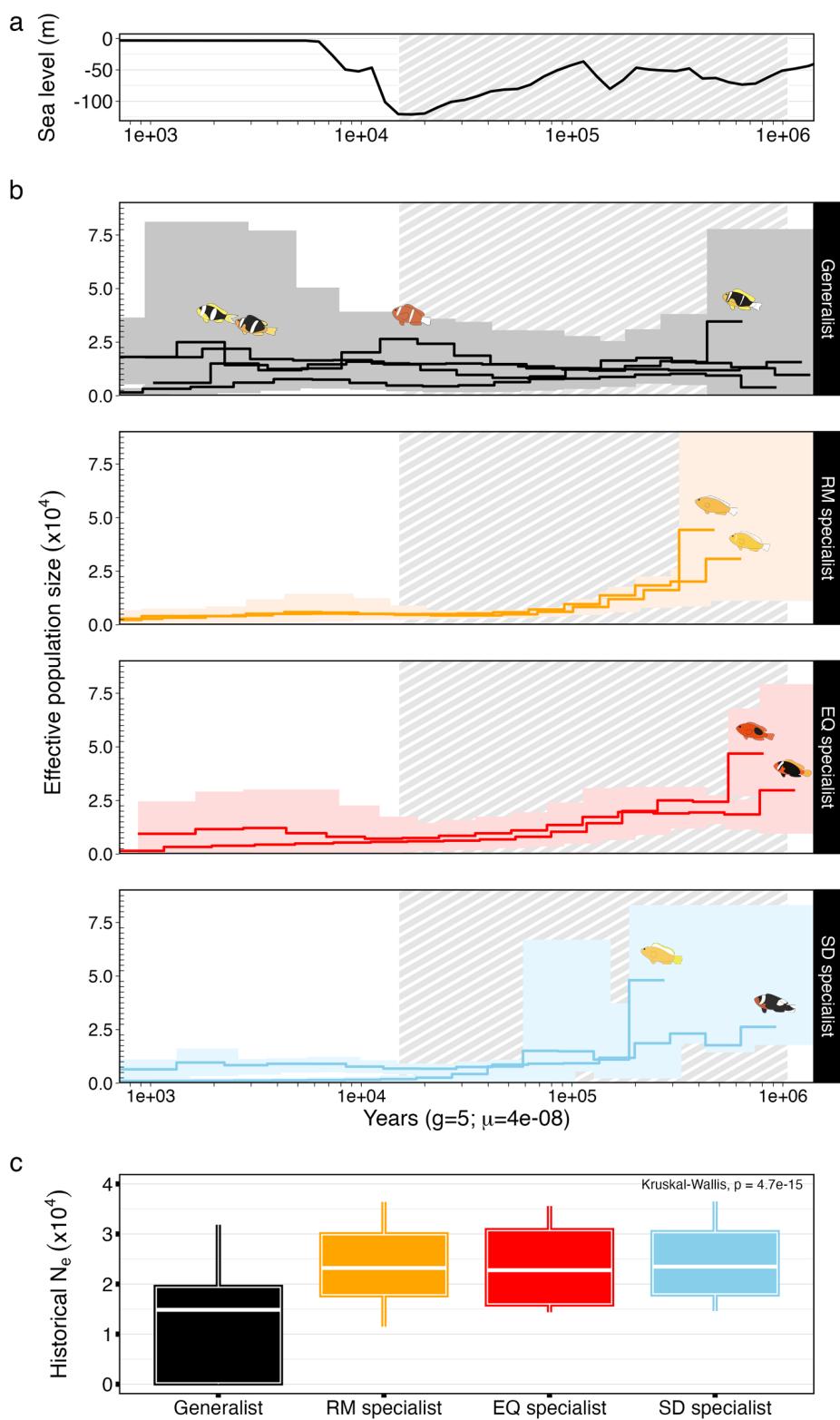
Fig. 4: Comparison of demographic reconstructions among host specialization groups in the context of late Pleistocene sea level fluctuations. (a) Sea level dynamics from 1 million years ago to present, with declines indicated by striped grey lines. (b) Demographic history reconstruction using *msmc2* [30] for ten species. Panels show different host associations: generalists (black), *Radianthus* (orange), *Entacmaea* (red), and *Stichodactyla* (blue). Step-lines represent average demographic reconstructions per species, with species illustrations near the highest N_e values. Light-colored areas show the range of N_e values across all populations. Period of sea-level declines is highlighted by striped grey lines. (c) Historical N_e for each group, calculated as the weighted harmonic mean of N_e estimates. The white line is the median, colored boxes show the IQR, and whiskers denote bounds. Kruskal-Wallis significance is indicated at the top right.
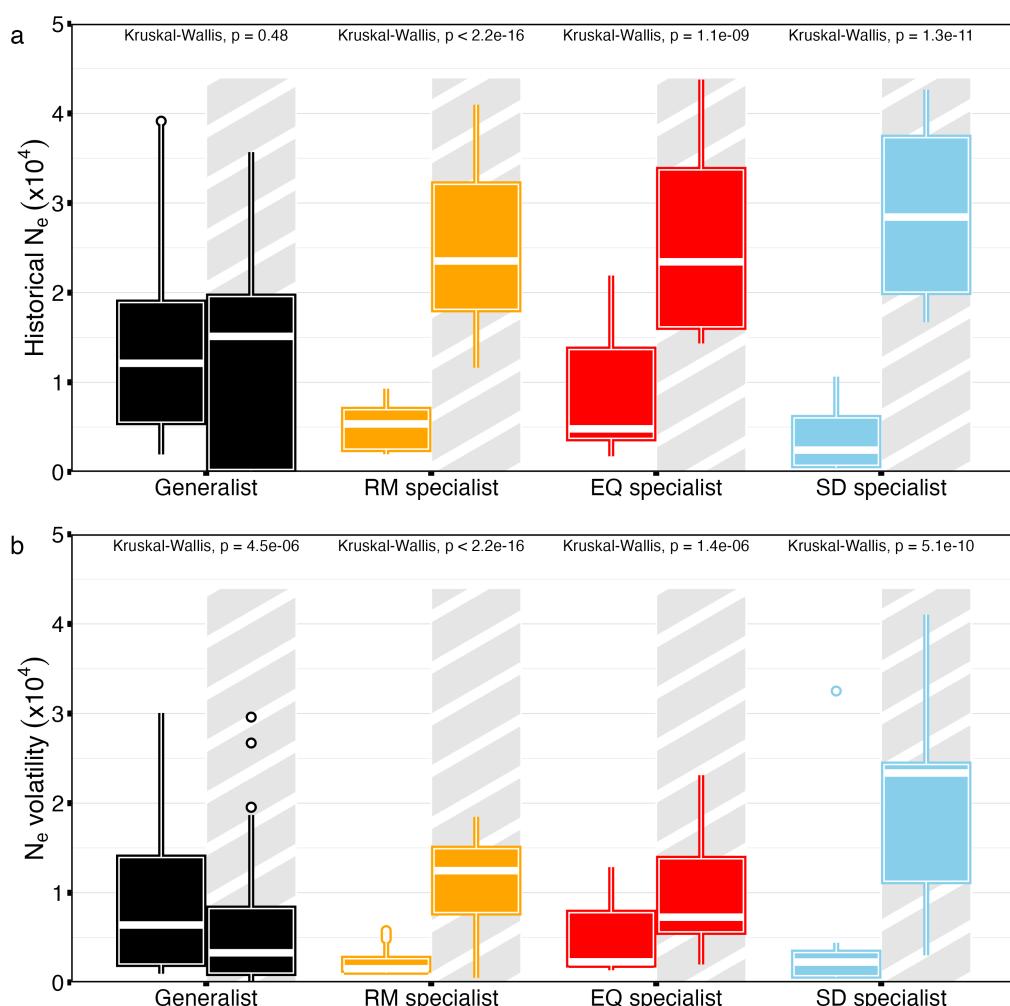
Fig. 5: Comparisons of N_e between the period of sea level declines and subsequent rise across host specialization groups. Values relative to the period of sea-level declines are differentiated by a striped grey-shadowed background. White line represents the median while colored boxes indicate the IQR. Whiskers denote the upper and lower bounds. Kruskal-Wallis statistical significance for each comparison is displayed at the top of each box plot. a) Historical N_e, calculated as the weighted harmonic mean of N_e estimates over time for each population's *msmc2* reconstruction. b) N_e volatility, calculated as the average true range (ATR) over 10,000-year intervals.

Discussion

Our demographic reconstructions and spatial analyses of clownfish populations across the Indo-Pacific reveal a complex interplay of historical, ecological, and oceanographic factors shaping their current distribution and genetic structure. The findings underscore the significant impact of Pleistocene sea-level fluctuations, host specialization, and oceanic currents on clownfish populations.

Historical Influences on Population Demography

Previous studies have hypothesized that sea-level fluctuations during the Pleistocene drove global population bottlenecks and structured populations across various marine taxa [12]. During lower sea levels, marine habitats fragmented, leading to population isolation and promoting genetic differentiation [31, 32]. As sea levels rose, previously isolated populations reconnected, facilitating gene flow and population mixing. These dynamic processes have contributed to genetic bottlenecks [13], where populations undergo drastic reductions in size, followed by expansion as environmental conditions improve.

Our estimated historical effective population sizes provide evidence of the impact of sea-level fluctuations on clownfish demographic dynamics. Specialist species—*Radianthus*, *Entacamea*, and *Stichodactyla* specialists—experienced significant declines in N_e, leading to population bottlenecks following the sea-level drop between 500,000 and 20,000 years ago, without subsequent recovery during rising sea levels. This highlights the long-lasting consequences of these bottlenecks and underscores the critical interplay between host specializations and environmental changes. In contrast, generalist species exhibited greater population stability, likely due to their ability to exploit diverse hosts, mitigating the severe consequences of habitat loss and fragmentation caused by sea-level fluctuations during this period.

Ecological and Oceanographic Factors

While sea-level declines during the Pleistocene likely affected host sea anemones similarly [among other marine organisms 12], clownfish population dynamics might not have entirely overlapped with their hosts. Specialists may have followed the fate of their singular host sea anemones, experiencing similar declines due to habitat loss and fragmentation [33]. Conversely, generalists, with their capacity to exploit diverse hosts, likely had higher survival chances during sea-level declines [34, 35]. An upward trend in effective population size of generalist populations around 100,000 years ago could be attributed to their opportunistic behavior, which allowed them to exploit habitats vacated by specialists undergoing bottlenecks. This phenomenon, observed in other taxa in response to climate change [36], is supported by fossil records indicating that past ecosystems shifted towards dominance by generalist species during such periods [37].

Clownfish exhibit a strong attachment to their host anemones [16], coupled with short pelagic larval dispersal and robust social and reproductive structures, rendering them highly vulnerable to habitat disturbances [19]. During periods of sea level rise, oceanographic conditions would have played a pivotal role in shaping the regional recolonization dynamics of clownfish, akin to their influence on other marine species [38, 39]. These oceanographic factors likely underlie the distinctive genetic diversity and population structure observed across clownfish species. Our findings strongly indicate that Pleistocene sea-level fluctuations and associated oceanographic dynamics have profoundly influenced the present biogeography of clownfish.

These dynamics likely contributed to the allopatric speciation events within the clownfish clade during this epoch [40, 41]. The synchronous occurrence of diversification bursts and significant bottleneck events suggests that habitat fragmentation could have facilitated these speciation processes. In many other taxa, habitat fragmentation due to climatic changes or natural barriers has similarly driven allopatric speciation and diversification [e.g., 6, 7]. In the case of clownfish, Pleistocene sea-level fluctuations likely created isolated populations, potentially accelerating speciation through allopatric mechanisms. However, we cannot disregard the importance of their unique social structure and short larval dispersal in driving speciation [42], particularly after geographical expansions of the late Pliocene and early Pleistocene [27].

Dispersal Barriers and Connectivity

Genetic connectivity analyses identified key geographical barriers to clownfish dispersal, including the Gulf of Thailand, Melanesia, and the Indian Ocean Gyre. These barriers, shaped by tidal and oceanic currents, contribute to genetic isolation and differentiation among populations. Our results highlight the profound impact of oceanic currents on clownfish dispersal routes, consistent with findings in other marine taxa [43–45]. The Indonesian Through-Flow plays a vital role in facilitating genetic exchange and increasing genetic diversity among clownfish populations by linking the Indian and Pacific Oceans. Conversely, areas at the margins of clownfish distributions, such as the Western Indian Ocean and Polynesia, exhibit the lowest levels of genetic differentiation due to dispersal barriers like the Indian Ocean Gyre and the Sunda and Sahul shelves.

Integrating these findings with demographic reconstructions provides insight into how habitat fragmentation during Pleistocene sea-level declines induced reproductive isolation. This process likely restricted gene flow between previously connected populations, resulting in genetic diversity loss and compromised population resilience [8, 9, 13]. Despite reconnecting during sea-level rises, tidal and current dynamics in certain areas have perpetuated isolation among populations. This underscores the crucial role of ocean currents in mediating the dispersal of fish larvae between fragmented habitats [a notion suggested by 46–48] and highlights the importance of incorporating dispersal paths in marine conservation efforts.

The Coral Triangle as a Biodiversity Hotspot

The Coral Triangle, known for its unparalleled species richness and endemic diversity [49], boasts an abundance and complexity of reef habitats. Combined with the Indonesian Through-flow that links the Indian and Pacific Oceans, these factors have been crucial in shaping the current marine biodiversity patterns in the Indo-Pacific. This region is particularly pivotal for clownfish, representing the center of origin for the clade [40] and currently hosting the highest diversity of clownfish species [50, 51].

Beyond generating biodiversity, our findings underscore the crucial role of the Coral Triangle in maintaining genetic diversity within species. Populations in this region exhibit higher historical N_e values than those in other regions, aligning with its historical role as a critical refuge during Pleistocene sea-level fluctuations [11]. The ecological complexity and environmental heterogeneity of the reef ecosystems in the Coral Triangle have likely contributed to maintaining these higher levels of genetic diversity [52–54].

Conclusion

Our study provides crucial insights into the dynamics of clownfish populations through an integrated analysis of Pleistocene sea-level fluctuations, population dynamics, and genetic diversity across their range. We highlight how ecological specializations influence species' responses and adaptability to climate change and environmental fluctuations. Our comprehensive approach identifies regions prone to increased reproductive isolation and inbreeding, alongside areas that serve as genetic diversity reservoirs. By examining how historical climate events have shaped the current genetic landscape of clownfish species, we can better predict the impacts of ongoing and future climate change on different populations, considering their host specialization and geographic distribution.

These findings are pivotal for guiding conservation efforts, emphasizing the need to prioritize strategies that uphold genetic diversity and ecosystem resilience throughout the clownfish's habitat. Incorporating historical demographic analysis into conservation planning enhances our understanding of long-term processes affecting clownfish populations, enabling the development of effective strategies to preserve their genetic diversity and adaptive capacity in the face of ongoing environmental changes. Specifically, conservation initiatives should prioritize maintaining connectivity in historical corridors for gene flow, such as the Indonesian Through-Flow, while safeguarding isolated populations that may harbor unique genetic adaptations. This holistic approach will sustain the adaptive potential of clownfish populations amidst habitat degradation and climate change, crucial for the enduring survival of these iconic coral reef inhabitants.

Future research should integrate these findings with contemporary ecological data and direct observations to offer a comprehensive understanding of clownfish population dynamics. Additionally, investigating the current and potential future impacts of climate change on clownfish and their host anemones is imperative for predicting and mitigating challenges facing these emblematic reef fishes.

Methods

Sampling and Data Collection

A total of 382 fin clips were collected from 10 clownfish species, representing distinct host categories within the clade as classified by [27]. The samples were systematically collected by capturing fish within their host anemones during SCUBA diving expeditions (see Table S1). Samples were preserved either in 70% ethanol (ETOH) or DESS buffer [55] and stored at 6°C immediately after retrieval. Collection efforts spanned multiple locations over an 11-year period starting in 2012 (see Supplementary Material).

Host Association

Species were categorized based on their host preferences following [27]. Two species, *Amphiprion akallopisos* and *A. perideraion*, classified as RM specialists, have adapted to inhabit and reproduce within *Radianthus magnifica*, a solitary sea anemone species with moderate-length tentacles found in coral reef habitats with moderate water flow. *A. melanopus* and *A. ephippium*, classified as EQ specialists, associate with *Entacmaea quadricolor*, a colonial sea anemone with long tentacles inhabiting crevices and rocky substrates in shallow coral reefs. Finally, *A. sandaracinos* specialized in *Stichodactyla mertensii* and *A. polymnus* specialized in *Stichodactyla haddoni* are classified as SD specialists, as both species belong to the *Stichodactyla* genus, characterized by short-tentacle solitary anemones forming a carpet-like extension over sandy areas of coral reefs. The remaining four species examined (*A. latifasciatus*, *A. akindynos*, *A. chrysopterus*, and *A. clarkii*) were categorized as Generalists, showing no clear preferences for any specific host sea anemone species.

Environmental and Geographical Data Collection

Sea level dynamics over the last million years were obtained from [56], and current climate data were sourced from GMED [57]. Environmental variables included current velocity, surface current, tide average, temperature, bottom temperature, sea surface temperature range, chlorophyll concentration, phytoplankton concentration, nitrate concentration, salinity, dissolved oxygen range, and bottom dissolved oxygen. Environmental data were standardized to a resolution of 0.083° x 0.083°, equivalent to a grid size of $100\,km^2$ near the equator, and adjusted to the Indo-Pacific study area. Geographically defined marine regions were obtained from Marine Ecoregions Of the World [MEOW; 58] and shallow reef habitats locations in the Indo-Pacific Ocean from UNEP-WCMC [59]

to implement geographical constrains into species distribution models (SDMs). Oceanic current data and Indonesian Through-Flow (ITF) polygons were derived from [60] and [61], respectively, for visualization and interpretation alongside genomic results.

Species Distributions

We estimated distributions of clownfish and their host sea anemones using 3,340 occurrences of 28 clownfish species (average 115 ± 150 per species) and 1,211 occurrences of the 10 sea anemones species hosting clownfishes (average 121 ± 58 per species) sourced from GBIF [62]. Species distribution models were conducted using the *ENMTML* R package [63] to delineate the geographic ranges and distributions of each species. See Supplementary Material for details on data and modeling parameters. Additionally, we inferred clownfish and sea anemone species richness distributions by aggregating presence/absence distribution maps across all clownfishes and all sea anemones, respectively.

Molecular Data Collection

DNA extraction, library preparation and sequencing

DNA was extracted using DNeasy blood and Tissue Kit (Qiagen GmbH, Hilden, Germany) and quantified using Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, USA). The integrity of the samples was assessed by electrophoresis. Libraries for samples collected before 2021 were prepared from 100 ng of DNA using the TruSeq Nano DNA library prep standard protocol, while libraries from samples collected after 2021 were prepared using the Nextera DNA Flex Library Preparation Kit following manufacturer's instructions. Fragment length distribution of the libraries was validated using a Bioanalyzer (Agilent Technologies, Santa Clara, USA). Libraries generated in 2017 and 2018 were sequenced on Illumina HiSeq 2500 platform with 100 paired-end lanes, while libraries prepared in 2019 were sequenced on the Illumina 4000 HiSeq platform with 150 paired-end lanes. Libraries prepared in 2022 and 2023 were sequenced on the NovaSeq 6000 Sequencing System (Illumina, San Diego, California) with 150 paired-end lanes. Sequencing was performed at the Genomic Technologies Facility (GTF) of the University of Lausanne, Switzerland.

Reads processing, mapping and SNP calling

Low-quality raw reads were trimmed, and potential adapter contamination was removed using *Trimmomatic* (parameters:–qual-threshold 20 –length-threshold 50, v.0.39; [64]). The quality of the processed reads was assessed with *FastQC* v.0.11.9 [65] and *MultiQC* [66]. Reads were mapped against *Amphiprion clarkii* reference genome (GenBank; ID: JALBFV000000000; [67]) using *BWA* v0.7.17 [68] and *Samtools* v1.15.1 [69] was used to process and sort the mapped reads, generating mapping statistics with *Bamtools* v2.5.2 [70]. The mapping results were processed to reduce potential SNPs calling errors [71]. Mapped reads were filtered with *Samtools* v1.15.1 [69] to retain only primary alignments and proper pairs (i.e., at the expected insert size and orientation) with a mapping quality higher than 30. Redundant sequencing data originated from the overlap of paired reads was removed using the `mergeReads` task from *ATLAS* v0.9.9 [72]. We used *ATLAS* pipeline for the SNP calling as it was shown to be more reliable than *GATK* in non-model species, maintaining high accuracy in variant calling for moderately divergent species [73]. Genotype likelihoods (GL) were calculated at each position using the `GLF` task from *ATLAS* for each sample and the reference genome. Then, genotype likelihoods were used to infer the major and the minor alleles for each position running the `majorMinor` task, obtaining the individuals genotype. SNP calling pipeline was carried out with all individuals from all species together generating an all-species dataset, and for each species independently, using all individuals from multiple populations, generating species-specific datasets.

Variant filtering and data processing

Reliability of the SNP dataset was improved by filtering invariant or poor confidence positions with *vcftools* v1.15.1 [69]. We filtered out low-confidence positions (variant quality score lower than 30), positions with coverage higher than 50, and with more than 10% of missing data (parameters: –minQ 30 –max-missing 0.9 –minDP 5 –max-meanDP 30 –maxDP 50). The VCF files were split by species populations to estimate individuals

heterozygosities (–het) and minor allele frequencies (–maf) using *vcftools* v1.15.1 [69]. With this information, we removed rare variants suspected to be private alleles or sequencing errors using a MAF threshold of 0.02. Further, we merged species population VCF files for each species, keeping only common variants using `merge` task from *bcftools* v1.15.1 [69]. Finally, VCF files were LD pruned and converted to BED files using *Plink* [74].

Population Structure Analysis

We assessed species population structure using Admixture [75] and Principal Component Analysis (PCA) with *Plink* v1.9 [74], aiming to identify genetic clustering patterns between and within species. We carried out Admixture and PCA on each species BED file to determine population identity and assess population structure, crucial for Sequential Markov Coalescence model interpretations [76]. Admixture was ran with K from 1 to the number of assumed populations of each species plus three, with 10-fold cross-validation and 2,000 bootstraps. Following [77], the best K value was selected upon the lowest cross-validation error. We also carried out PCA on all-species dataset to determine species identities. Three *A. polymnus* samples were clustering with *A. clarkii* showing to be potentially misidentified or mislabelled. Consequently, these three samples were removed for demographic reconstruction and *eems* analysis. Additionally, we reconstructed phylogenetic relationships of all samples using *PhyML* [78] with a GTR + Γ and 100 bootstraps to explore the evolutionary relationships among species, complementing population structure analyses.

Spatial Connectivity and Effective Migration Rates

We first evaluated Isolation-by-distance patterns by fitting a linear regression between geographical distances and genetic distances. Geographical distances were calculated as the geodesic distance between population centroids using the `geoDist` function from the *oce* R package [79]. We estimated population pairwise Fst for each species using *vcftools* v1.15.1 [69] to quantify the genetic distances between populations. The positive association between geographical and genetic distances indicated a clear Isolation-by-distance signal, thereby validating the use of the *eems* method [29], which is based on Isolation-by-distance expectations.

We used the Estimated Effective Migration Surfaces *eems* [29] to assess genetic connectivity patterns within and between populations. *eems* estimates effective migration rates and genetic diversity using genetic data, modeling spatial gene flow patterns. Employing Markov Chain Monte Carlo (MCMC) simulations, *eems* infers these rates between demes, providing insights into genetic structure, barriers to gene flow, and dispersal dynamics. It also calculates a posterior probability for each estimate, offering a probabilistic assessment of whether rates are higher or lower than expected under the Isolation-by-Distance assumption. This Bayesian approach facilitates a nuanced understanding of genetic connectivity and the impact of spatial factors on population dynamics and evolutionary processes.

Genetic dissimilarities between pairs were computed using the `bed2diffs` program, available within the *eems* toolkit. All analyses were executed with 2,000,000 iterations following a burn-in phase of 1,000,000 iterations for the Markov Chain Monte Carlo. For each species, we conducted analyses with 50, 200, and 500 demes, employing three chains for each configuration. This extensive exploration allowed us to compare results and fine-tune parameters, following the developers recommendations. The outcomes of these three runs were visualized using the *rEEMSplots* R package [29], demonstrating consistent results across runs. Species *A. latifasciatus* and *A. ephippium* were excluded from these analyses due to having only one population of each.

Spatial Statistical Analyses

We examined the relationship between effective migration and genetic diversity rates derived from *eems* analysis and environmental variables. Initially, we gauge the significance of environmental variables in explaining effective migration surfaces within generalized linear models (GLMs). Then, we assessed spatial auto-correlation using Moran's on each GLM. Upon confirming spatial auto-correlation's significance, we applied Rao's score tests using `lm.RStests` from the *spdep* R package (Bivand et al., 2011) to determine suitable spatially dependent linear models. Further analysis involved fitting Lagrange Multiplier tests for Spatial Autoregressive (SAR) models using `lagsarlm` from the *spatialreg* R package [80],

independently testing each environmental variable alongside clownfish and sea anemone richness. Additionally, we explored various combinations of variables based on biological expectations (see Supplementary Material for detailed model descriptions).

Demographic history reconstruction

We employed *msmc2* [30] to reconstruct the historical dynamics of effective population size (N_e) across all populations of the ten clownfish species, following guidelines from *msmc-tools* (https://github.com/stschiff/msmc-tools). From individual BAM files, we used `mpileup` and `call` tasks in *bcftools* v1.15.1 [69] followed by `bamCaller.py` from *msmc-tools* to generate chromosome-level individual VCF files. Average coverage by chromosome was estimated using `depth` task in *Samtools* v1.15.1 [69]. VCF files were phased with *Whatshap* [81], and consensus and individual masks were generated using `makeMappabilitymask.py`. Then, *msmc2* input files were generated with `generate_multihetsep.py`.

For each species, we conducted ten population reconstructions, each time selecting three random individuals per population. The time-segmented pattern used was $1*2+16*1+1*2$, recommended for shorter genomes [30], with a maximum of 100 iterations. Additionally, we performed 50 bootstraps using `multihetsep_bootstrap.py` to assess uncertainty levels at each time step. Historical time periods and N_e were inferred using a fixed per-generation mutation rate of 4×10^{-8} [82, 83] and generation time of 5 years [84], consistent with previous clownfish studies [85, 86].

Population N_e reconstructions for each species were averaged to derive overall species trajectories. Historical N_e values were computed using the weighted harmonic mean, which adjusts for the relative time intervals to account for dataset unevenness. Additionally, we assessed the volatility of species N_e using the Average True Range (ATR), a statistical measure commonly used in financial markets to gauge volatility. N_e volatility was calculated as the average of the true ranges over successive 10,000-year intervals. This metric provides insights into the fluctuation of N_e over short evolutionary timescales, helping to assess the stability or volatility of population sizes across evolutionary periods and enhancing our understanding of population dynamics and their responses to environmental changes.

Supplementary information

Declarations

Conflict of interest statement. The authors declare no competing interests in the publication of this work.

Biological Sampling Permits. The research permits obtained for this study are the following:

Country	Permit Number
Mayotte	06/UTM/2016
New Caledonia	N°60912-895-2017/JJC
Kenya	NCST/RRI/12/1/BS/250
Indonesia	55/PSTK/UH/XI/04
Thailand	0401/16454
Australia	
Queensland Department of Primary Industries	118636, 150981
QLD General Fisheries Permit:	180731
Great Barrier Marine Park (GBRMPA)	G17/38160.1
Queensland Parks and Wildlife Marine Parks	G08/26733.1
	G08/28114.1
	G09/31678.1
	G10/33597.1
	G11/34452.1
	G11/34640.1
University of Queensland, Animal Ethics	QBI/304/16
New Caledonia	
Environment Direction of the South Province	N°60912-895-2017/JJC
Environment Direction of the North Province	N°1291-2017/ARR/DENV
Maldives	OTHR30-D/INDIV/2016/538
Solomon Islands	Permit to S. Albert
Papua New Guinea	Research Visa: 10350008304
	Permit to Export Wildlife: 011318
Taiwan	03/F/0390B
Philippines	
Bureau of Fisheries and Aquatic Resources	Commodity Clearance 2016-20091
Palawan Council for Sustainable Development (PCSD)	GP 2016-03, Wildlife Transport Permit no. 2016-05-000062-DMO – Calamianes
Fisheries inspection and quarantine clearance	009391

Benefits Generated.   The sharing of our data and results on public databases ensures transparency and facilitates further research in the field. All fieldwork conducted during this study adhered to local regulations and was carried out in collaboration with local entities, including the University of Mayotte (Mayotte), University of New Caledonia (New Caledonia), Kenya Marine and Fisheries Research Institute (Kenya), National Research Institute (Papua New Guinea), Universitas Hasanuddin (Indonesia), University of Ramkhamhaeng (Thailand), University of Toliara and the Ministry of Marine Resources and Fisheries (Madagascar), Lizard Island Research Station and University of Queensland (Australia). Our research activities were conducted in accordance with the "Access and Benefit Sharing" (ABS) principles of the Nagoya Protocol established by the Convention on Biological Diversity (CBD).

Samples from Indonesia (2013), Maldives (2017), New Caledonia (2017), Thailand (2022), Madagascar (2022) and Australia (2023) were obtained by Nicolas Salamin's research group. Samples from Solomon Islands, Papua New Guinea and Lizard Island (2018) were obtained by Cynthia Riginos' research group. Sampling in the Solomon Islands was via the Australian Government's Pacific Strategy Assistance Program and with the assistance of the Roviana Conservation Foundation. Samples from Taiwan were obtained by Filip Huyghe. Samples from Philippines were obtained by Jimmy O'Donnell.

Data availability.   Metadata associated with our study will be uploaded to GEOME, and genomic data deposited in the National Center for Biotechnology Information (NCBI) under a BioProject ID. Access to these data will also be provided upon acceptance of the manuscript.

Code availability.   The scripts used for the analyses will be deposited on Dryad upon acceptance of the manuscript, ensuring the reproducibility and accessibility of our research findings.

Author contribution.   AGJ and NS designed the study. AGJ, TG, LF, SH, AM, SS, JB, GL, LM, BF, FC, AS, CS, PR, PR, TW, SP and TY participated in the fieldwork expeditions and sample collection. AGJ conducted the research and analyses and wrote the manuscript. TG and NS contributed to the interpretation of the analyses, and the writing of the manuscript. All authors participated in reviewing, correcting, and approving the final version of the manuscript.

References

[1] Crooks, K. R. *et al.* Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. *Proceedings of the national Academy of Sciences* **114** (29), 7635–7640 (2017) .

[2] Fahrig, L. Effects of habitat fragmentation on biodiversity. *Annual review of ecology, evolution, and systematics* **34** (1), 487–515 (2003) .

[3] Crespo-Miguel, R., Jarillo, J. & Cao-García, F. J. Scaling of population resilience with dispersal length and habitat size. *Journal of Statistical Mechanics: Theory and Experiment* **2022** (2), 023501 (2022) .

[4] Ibáñez, I., Katz, D. S., Peltier, D., Wolf, S. M. & Connor Barrie, B. T. Assessing the integrated effects of landscape fragmentation on plants and plant communities: the challenge of multiprocess–multiresponse dynamics. *Journal of Ecology* **102** (4), 882–895 (2014) .

[5] Fahrig, L. Ecological responses to habitat fragmentation per se. *Annual review of ecology, evolution, and systematics* **48** (1), 1–23 (2017) .

[6] Tolley, K. A. *et al.* Ancient forest fragmentation or recent radiation? testing refugial speciation models in chameleons within an african biodiversity hotspot. *Journal of Biogeography* **38** (9), 1748–1760 (2011) .

[7] Dias, M. S., Cornu, J.-F., Oberdorff, T., Lasso, C. A. & Tedesco, P. A. Natural fragmentation in river networks as a driver of speciation for freshwater fishes. *Ecography* **36** (6), 683–689 (2013) .

[8] Habel, J. C. & Zachos, F. E. Habitat fragmentation versus fragmented habitats. *Biodiversity and Conservation* **21** (11), 2987–2990 (2012) .

[9] Young, E. F. *et al.* Stepping stones to isolation: Impacts of a changing climate on the connectivity of fragmented fish populations. *Evolutionary Applications* **11** (6), 978–994 (2018) .

[10] Haddad, N. M. *et al.* Habitat fragmentation and its lasting impact on earth's ecosystems. *Science advances* **1** (2), e1500052 (2015) .

[11] Pellissier, L. *et al.* Quaternary coral reef refugia preserved fish diversity. *Science* **344** (6187), 1016–1019 (2014) .

[12] Ludt, W. B. & Rocha, L. A. Shifting seas: The impacts of pleistocene sea-level fluctuations on the evolution of tropical marine taxa. *Journal of Biogeography* **42** (1), 25–38 (2015) .

[13] Fauvelot, C., Bernardi, G. & Planes, S. Reductions in the mitochondrial dna diversity of coral reef fish provide evidence of population bottlenecks resulting from holocene sea-level change. *Evolution* **57** (7), 1571–1583 (2003) .

[14] McLaughlin, B. *et al.* Conservation strategies for the climate crisis: An update on three decades of biodiversity management recommendations from science. *Biological Conservation* **268**, 109497 (2022) .

[15] Laudet, V. & Ravasi, T. (eds) *Evolution, Development and Ecology of Anemonefishes: Model Organisms for Marine Science* (CRC Press, 2023).

[16] Fautin, D. The anemonefish symbiosis what is known and what is not. *Symbiosis* (1991) .

[17] Elliott, J., Elliott, J. & Mariscal, R. Host selection, location, and association behaviors of anemonefishes in field settlement experiments. *Marine Biology* **122**, 377–389 (1995) .

[18] Buston, P. Mortality is associated with social rank in the clown anemonefish (amphiprion percula). *Marine Biology* **143**, 811–815 (2003) .

[19] Salles, O. C. *et al.* Strong habitat and weak genetic effects shape the lifetime reproductive success in a wild clownfish population. *Ecology letters* **23** (2), 265–273 (2020) .

[20] Wellington, G. & Victor, B. Planktonic larval duration of one hundred species of pacific and atlantic damselfishes (pomacentridae). *Marine Biology* **101**, 557–567 (1989) .

[21] Berumen, M. L. *et al.* Persistence of self-recruitment and patterns of larval connectivity in a marine protected area network. *Ecology and evolution* **2** (2), 444–452 (2012) .

[22] Pinsky, M. L., Montes, H. R., Jr & Palumbi, S. R. Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution* **64** (9), 2688–2700 (2010) .

[23] Fricke, H. Individual partner recognition in fish: field studies on amphiprion bicinctus. *Naturwissenschaften* **60** (4), 204–204 (1973) .

[24] Ricciardi, F., Boyer, M. & Ollerton, J. Assemblage and interaction structure of the anemonefish-anemone mutualism across the manado region of sulawesi, indonesia. *Environmental Biology of Fishes* **87**, 333–347 (2010) .

[25] Dixon, A. K., McVay, M. J. & Chadwick, N. E. Demographic modelling of giant sea anemones: population stability and effects of mutualistic anemonefish in the jordanian red sea. *Marine and Freshwater Research* **68** (11), 2145–2155 (2017) .

[26] Litsios, G. *et al.* Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evolutionary Biology* **12** (2012). https://doi.org/10.1186/1471-2148-12-212 .

[27] Gaboriau, T. *et al.* Host-use drives convergent evolution in clownfish and disentangles the mystery of an iconic adaptive radiation. In prep.

[28] Elliott, J. K. & Mariscal, R. Coexistence of nine anemonefish species: differential host and habitat utilization, size and recruitment. *Marine Biology* **138**, 23–36 (2001) .

[29] Petkova, D., Novembre, J. & Stephens, M. Visualizing spatial population structure with estimated effective migration surfaces. *Nature genetics* **48** (1), 94–100 (2016) .

[30] Schiffels, S. & Wang, K. in *Msmc and msmc2: The multiple sequentially markovian coalescent* (ed.Dutheil, J. Y.) *Statistical Population Genomics*, Vol. 2090 of *Methods in Molecular Biology* 147–166 (Springer US, New York, NY, 2020).

[31] Blanco Gonzalez, E., Knutsen, H. & Jorde, P. E. Habitat discontinuities separate genetically divergent populations of a rocky shore marine fish. *PloS one* **11** (10), e0163052 (2016) .

[32] Rönkä, N. *et al.* Genetic differentiation in an endangered and strongly philopatric, migrant shorebird. *BMC Ecology and Evolution* **21** (1), 125 (2021) .

[33] Chomicki, G., Weber, M., Antonelli, A., Bascompte, J. & Kiers, E. T. The impact of mutualisms on species richness. *Trends in Ecology & Evolution* **34** (8), 698–711 (2019) .

[34] Kiers, E. T., Palmer, T. M., Ives, A. R., Bruno, J. F. & Bronstein, J. L. Mutualisms in a changing world: an evolutionary perspective. *Ecology letters* **13** (12), 1459–1474 (2010) .

[35] Fontúrbel, F. E., Nespolo, R. F., Amico, G. C. & Watson, D. M. Climate change can disrupt ecological interactions in mysterious ways: Using ecological generalists to forecast community-wide effects. *Climate Change Ecology* **2**, 100044 (2021) .

[36] Weiskopf, S. R. *et al.* Climate change effects on biodiversity, ecosystems, ecosystem services, and natural resource management in the united states. *Science of the Total Environment* **733**, 137782 (2020) .

[37] Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C. & Finnegan, S. Climate change and the past, present, and future of biotic interactions. *Science* **341** (6145), 499–504 (2013) .

[38] Underwood, J. N. Genetic diversity and divergence among coastal and offshore reefs in a hard coral depend on geographic discontinuity and oceanic currents. *Evolutionary Applications* **2** (2), 222–233 (2009) .

[39] Bradbury, I. R. *et al.* Landscape structure and climatic variation determine atlantic salmon genetic connectivity in the northwest atlantic. *Canadian Journal of Fisheries and Aquatic Sciences* **71** (2), 246–258 (2014) .

[40] Litsios, G., Pearman, P. B., Lanterbecq, D., Tolou, N. & Salamin, N. The radiation of the clownfishes has two geographical replicates. *Journal of Biogeography* **41**, 2140–2149 (2014). https://doi.org/10.1111/jbi.

12370 .

[41] Gaboriau, T., Leprieur, F., Mouillot, D. & Hubert, N. Influence of the geography of speciation on current patterns of coral reef fish biodiversity across the indo-pacific. *Ecography* **41** (8), 1295–1306 (2018) .

[42] Jones, G. P., Planes, S. & Thorrold, S. R. Coral reef fish larvae settle close to home. *Current biology* **15** (14), 1314–1318 (2005) .

[43] White, C. *et al.* Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences* **277** (1688), 1685–1694 (2010) .

[44] Miller, A. D., Versace, V. L., Matthews, T. G., Montgomery, S. & Bowie, K. C. Ocean currents influence the genetic structure of an intertidal mollusc in southeastern australia–implications for predicting the movement of passive dispersers across a marine biogeographic barrier. *Ecology and evolution* **3** (5), 1248–1261 (2013) .

[45] Snead, A. A. *et al.* Out to sea: Ocean currents and patterns of asymmetric gene flow in an intertidal fish species. *Frontiers in Genetics* **14**, 1206543 (2023) .

[46] Treml, E. A., Halpin, P. N., Urban, D. L. & Pratson, L. F. Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology* **23**, 19–36 (2008) .

[47] Cowen, R. K. & Sponaugle, S. Larval dispersal and marine population connectivity. *Annual review of marine science* **1**, 443–466 (2009) .

[48] Morrison, R. A. & Sandin, S. A. Biogeography and population connectivity of coral reef fishes. *Changing diversity in changing environment* 299–322 (2011) .

[49] Bellwood, D. R., Renema, W. & Rosen, B. R. Biodiversity hotspots, evolution and coral reef biogeography. *Biotic evolution and environmental change in Southeast Asia* **216** (2012) .

[50] Camp, E. F., Hobbs, J.-P. A., De Brauwer, M., Dumbrell, A. J. & Smith, D. J. Cohabitation promotes high diversity of clownfishes in the coral triangle. *Proceedings of the Royal Society B: Biological Sciences* **283** (1827), 20160277 (2016) .

[51] García Jiménez, A., Broennimann, O., Guisan, A., Gaboriau, T. & Salamin, N. Implementation of biotic interactions in niche analyses unravels the patterns underneath community composition in clownfishes. *bioRxiv* (2023). https://doi.org/10.1101/2023.03.30.534900 .

[52] Pamilo, P. Genetic variation in heterogeneous environments. *Annales Zoologici Fennici* **25** (1), 99–106 (1988) .

[53] Temunović, M. *et al.* Environmental heterogeneity explains the genetic structure of continental and mediterranean populations of fraxinus angustifolia vahl. *PloS one* (2012) .

[54] Abdul-Rahman, F., Tranchina, D. & Gresham, D. Fluctuating environments maintain genetic diversity through neutral fitness effects and balancing selection. *Molecular Biology and Evolution* **38** (10), 4362–4375 (2021) .

[55] Oosting, T., Hilario, E., Wellenreuther, M. & Ritchie, P. A. Dna degradation in fish: Practical solutions and guidelines to improve dna preservation for genomic research. *Ecology and Evolution* **10**, 8643–8651 (2020). https://doi.org/10.1002/ece3.6558 .

[56] Müller, R. D., Sdrolias, M., Gaina, C., Steinberger, B. & Heine, C. Long-term sea-level fluctuations driven by ocean basin dynamics. *science* **319** (5868), 1357–1362 (2008) .

[57] Basher, Z., Bowden, D. A. & Costello, M. J. Gmed: Global marine environment datasets for environment visualisation and species distribution modelling. *Earth System Science Data Discussions* **2018**, 1–62 (2018) .

[58] Spalding, M. D. *et al.* Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience* **57** (7), 573–583 (2007) .

[59] Unep-Wcmc, T. & Wri, T. Global distribution of warm-water coral reefs, compiled from multiple sources including the millennium coral reef mapping project. *UNEP World Conservation Monitoring Centre Cambridge, UK* (2010) .

[60] Brown, e. a. Ocean Currents. National Snow and Ice Data Center; NOAA; National Weather Service; US Army; IBTrACS v03r07; Maps.com (2007). Online resource.

[61] Mitsuguchi, T., Dang, P. X., Kitagawa, H., Yoneda, M. & Shibata, Y. Tropical south china sea surface 14c record in an annually-banded coral. *Radiocarbon* **49** (2), 905–914 (2007) .

[62] GBIF.Org User. Occurrence download (2023). URL https://www.gbif.org/occurrence/download/0132153-230224095556074.

[63] de Andrade, A. F. A., Velazco, S. J. E. & Júnior, P. D. M. Enmtml: An r package for a straightforward construction of complex ecological niche models. *Environmental Modelling & Software* **125**, 104615 (2020) .

[64] Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence

data. *Bioinformatics* **30** (15), 2114–2120 (2014). URL https://doi.org/10.1093/bioinformatics/btu170. https://doi.org/10.1093/bioinformatics/btu170, https://academic.oup.com/bioinformatics/article-pdf/30/15/2114/48924714/bioinformatics_30_15_2114.pdf .

[65] Andrews, S. *et al.* FastQC. Babraham Institute (2012).

[66] Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32** (19), 3047–3048 (2016). URL https://doi.org/10.1093/bioinformatics/btw354. https://doi.org/10.1093/bioinformatics/btw354, https://academic.oup.com/bioinformatics/article-pdf/32/19/3047/49021359/bioinformatics_32_19_3047.pdf .

[67] Moore, B. *et al.* The chromosome-scale genome assembly of the yellowtail clownfish amphiprion clarkii provides insights into the melanic pigmentation of anemonefish. *G3 GenesGenomesGenetics* **13** (3), jkad002 (2023). URL https://doi.org/10.1093/g3journal/jkad002. https://doi.org/10.1093/g3journal/jkad002, https://academic.oup.com/g3journal/article-pdf/13/3/jkad002/49479873/jkad002.pdf .

[68] Li, H. & Durbin, R. Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics* **25** (14), 1754–1760 (2009). URL https://doi.org/10.1093/bioinformatics/btp324. https://doi.org/10.1093/bioinformatics/btp324, https://academic.oup.com/bioinformatics/article-pdf/25/14/1754/48994219/bioinformatics_25_14_1754.pdf .

[69] Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10** (2) (2021). URL https://doi.org/10.1093/gigascience/giab008. https://doi.org/10.1093/gigascience/giab008, giab008, https://academic.oup.com/gigascience/article-pdf/10/2/giab008/36332246/giab008.pdf .

[70] Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P. & Marth, G. T. Bamtools: a c++ api and toolkit for analyzing and managing bam files. *Bioinformatics* **27** (12), 1691–1692 (2011). URL https://doi.org/10.1093/bioinformatics/btr174. https://doi.org/10.1093/bioinformatics/btr174 .

[71] Altmann, A. *et al.* A beginners guide to snp calling from high-throughput dna-sequencing data. *Human Genetics* **131**, 1541–1554 (2012). https://doi.org/10.1007/s00439-012-1213-z .

[72] Link, V. *et al.* Atlas: analysis tools for low-depth and ancient samples. *BioRxiv* 105346 (2017) .

[73] Duchen, P. & Salamin, N. A cautionary note on the use of genotype callers in phylogenomics. *Systematic biology* **70** (4), 844–854 (2021) .

[74] Purcell, S. *et al.* Plink: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–575 (2007). https://doi.org/10.1086/519795 .

[75] Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome research* **19** (9), 1655–1664 (2009) .

[76] Mather, N., Traves, S. M. & Ho, S. Y. A practical introduction to sequentially markovian coalescent methods for estimating demographic history from genomic data. *Ecology and evolution* **10** (1), 579–589 (2020) .

[77] Bradburd, G. S., Coop, G. M. & Ralph, P. L. Inferring continuous and discrete population genetic structure across space. *Genetics* **210** (1), 33–52 (2018) .

[78] Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of phyml 3.0. *Systematic biology* **59** (3), 307–321 (2010) .

[79] Kelley, D. E. & Kelley, D. E. The oce package. *Oceanographic analysis with r* 91–101 (2018) .

[80] Bivand, R. *et al.* spatialreg: Spatial regression analysis. *R package version* 1–1 (2019) .

[81] Martin, M. *et al.* Whatshap: fast and accurate read-based phasing. *BioRxiv* 085050 (2016) .

[82] Delrieu-Trottin, E. *et al.* Population expansions dominate demographic histories of endemic and widespread pacific reef fishes. *Scientific reports* **7** (1), 40519 (2017) .

[83] Rolland, J., Silvestro, D., Litsios, G., Faye, L. & Salamin, N. Clownfishes evolution below and above the species level. *Proceedings of the Royal Society B: Biological Sciences* **285** (1873), 20171796 (2018) .

[84] Buston, P. M. & García, M. An extraordinary life span estimate for the clown anemonefish amphiprion percula. *Journal of Fish Biology* **70** (6), 1710–1719 (2007) .

[85] Schmid, S. *et al.* Gene flow throughout the evolutionary history of a colour polymorphic and generalist clownfish. *Molecular Ecology* e17436 (2024) .

[86] Marcionetti, A. *et al.* Recurrent gene flow events occurred during the diversification of clownfishes of the skunk complex. *Molecular Ecology* **33** (11), e17347 (2024) .