Reproductive Isolation and Differential Introgression Shape the Genomic Landscape of the Red Alga Amansia glomerata in the Hawaiian Archipelago

Lauric Reynes

lreynes@hawaii.edu

University of Hawaii https://orcid.org/0000-0002-0223-4332

Alison Sherwood

University of Hawaii

James Fumo

University of Hawaii

Article

Keywords:

Posted Date: March 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9015118/v1

Additional Declarations: There is no duality of interest

Version of Record: A version of this preprint was published at Heredity on July 22nd, 2026. See the published version at https://doi.org/10.1038/s41437-026-00869-y.

Reproductive Isolation and Differential Introgression Shape the Genomic Landscape of the Red Alga Amansia glomerata in the Hawaiian Archipelago

ABSTRACT

Reproductive isolation; Genomic divergence; Introgression; Seaweed, Hawai'i

Speciation in marine environments is inherently complex, and the mechanisms underlying the evolution of reproductive isolation remain a fundamental yet understudied challenge in macroalgae. This study investigates two lineages of the marine red alga *Amansia glomerata* around O'ahu, Hawai'i, to test the hypothesis that lineage boundaries are maintained by strong reproductive barriers resulting from historical geographic isolation. Using genomic sequencing (ddRAD), fine-scale spatial transects, and demographic modeling, we characterized lineage spatial structure and genomic divergence. Our results revealed that lineages remained strongly differentiated across the genome, even where they co-occurred across extensive sympatric regions. Demographic analyses supported a scenario of allopatric divergence followed by secondary contact, likely driven by Pleistocene sea-level fluctuations within the Hawaiian Archipelago. The absence of recombinant hybrids, together with numerous loci acting as barriers to gene flow, is consistent with the evolution of postzygotic isolation. Nevertheless, genomic signatures of introgression indicate that gene flow during secondary contact was spatially structured around the island and heterogeneous across the genome, revealing differential permeability of lineage boundaries at the time of contact. While our findings highlight allopatric divergence and intrinsic reproductive barriers as major drivers of speciation in this system, the potential role of ecological differentiation remains to be explored. Overall, this study provides pioneering insights into the genomics of speciation in Hawaiian seaweeds and underscores their potential as a model system for marine speciation.

1. INTRODUCTION

Speciation is fundamentally linked to the evolution of reproductive isolation (RI), which prevents gene flow between diverging populations (Coyne & Orr, 2004). Speciation can proceed along multiple pathways, with the degree of spatial, ecological, and reproductive differentiation playing a central role in determining which pathway is followed (Dieckmann, 2004; Seehausen et al., 2014; Johannesson et al., 2024). Gaining insight into the evolution of RI is increasingly critical for evaluating how environmental and rapid climatic changes influence speciation dynamics (Hudson et al., 2021).

Speciation in marine environments has historically been considered more restricted than in terrestrial systems, due to the scarcity of physical barriers and the high dispersal potential of many marine organisms, which maintain genetic homogeneity among populations (Kay & Palumbi, 1987; Palumbi, 1992, 1994). However, physical barriers do occur in the marine environment and are widely recognized as important drivers of speciation by modulating patterns of gene flow (Mayr, 1954; Fitzpatrick et al., 2009), particularly through historical processes such as Pleistocene sea-level fluctuations (e.g., Hewitt, 2000; Ludt & Rocha, 2015; Hirschfeld et al., 2023). Notably, speciation involving geographic isolation followed by secondary contact (SC) accounts for 53% of the 116 cases of marine speciation reviewed by De Jode et al. (2023). In the absence of gene flow, reproductive isolation (RI) is predicted to be

primarily intrinsic, arising from the evolution of genetic incompatibilities between populations that have evolved independently (Dobzhansky, 1937; Muller, 1942; Orr & Turelli, 2001; Coughlan & Matute, 2020). If isolation persists long enough, RI can restrict gene flow in genomic regions associated with barrier loci when previously isolated populations experience SC. This results in a heterogeneous genomic landscape, with some regions remaining highly differentiated while others are more freely exchanged (Ravinet et al., 2017; Harrison & Larson, 2016). By contrast, speciation with gene flow mostly relies on ecological differentiation, whereby divergent selection can initiate divergence by reducing gene flow at linked loci (Rundle & Nosil, 2005; Schluter, 2009; Nosil, 2012). Although ecological differentiation can evolve secondarily through geographic isolation, when it acts as the primary driver of divergence, it is expected to produce a genomic landscape characterized by regions differentiated by selection, while the remainder of the genome remains largely homogenized (Via & West, 2008; Feder & Nosil, 2010; Feder et al., 2012). In marine systems, this process is increasingly recognized as a major driver of speciation, challenging traditional models that emphasize drift-driven differentiation (see the review by Potkamp & Fransen, 2019). A striking example is provided by the marine snail *Littorina saxatilis*, in which ecotypes evolved through divergent selection between crab-rich and low-predation environments, with patterns of differentiation consistent with divergence with gene flow (Johannesson et al., 2010).

Our understanding of marine speciation remains limited by several key knowledge gaps, including a reliance on general patterns derived from a relatively narrow range of marine taxa (Miglietta et al., 2011; Faria et al., 2021; De Jode et al., 2023). As a result, many groups, particularly marine macroalgae, remain largely underrepresented in speciation research despite their potential for investigating reproductive isolation (RI). Macroalgae exhibit a remarkable diversity of reproductive modes and life cycles, with haplodiplontic species being of particular interest for enabling fine-scale dissection of reproductive barriers (Valero et al., 1992; Heesch et al., 2021; Krueger-Hadfield, 2024). In haplodiplontic species, the two principal stages of the sexual life cycle occur as free-living, functionally independent phases, allowing reproductive isolation (RI) to be examined separately at pre- and post-zygotic stages (Valero et al., 1992), for example by investigating differential hybridization and admixture between the two phases (Montecinos et al., 2017). Despite this considerable potential, genomic studies of speciation in macroalgae remain scarce (but see Denoeud et al., 2024, for the brown alga *Ectocarpus* sp.).

The Hawaiian Archipelago, renowned for its macroalgal-diversity (Sherwood & Guiry, 2023), provides an exceptional natural laboratory for investigating reproductive isolation. Its extreme geographical isolation, more than 4,000 km from the nearest continental landmass, combined with its well-characterized geological patterns of island formation, has established the archipelago as a major evolutionary hotspot

(Rotondo et al., 1981; Wagner & Funk, 1995). Speciation in Hawai‘i is thought to result largely from rare colonization events originating outside the islands, typically on the older high islands, followed by sequential episodes of dispersal and diversification toward younger islands as they emerge (Carson & Clague, 1995; Roderick & Gillespie, 1998; Price & Clague, 2002; VanderWerf et al., 2010). Among the marine seaweeds that have diversified in the archipelago, the red alga *Amansia glomerata* C. Agardh (Ceramiales, Rhodomelaceae) comprises four described mitochondrial lineages (Sherwood et al., 2011; Fumo & Sherwood, 2023). *A. glomerata* exhibits a triphasic life cycle characterized by two free-living stages: a haploid gametophyte and a diploid tetrasporophyte. These stages are isomorphic, with thalli morphologically indistinguishable in the absence of reproductive structures. The species is relatively common in Hawai‘i and occupies diverse habitats, including eroded coral substrates, shaded overhangs, and reef flats, from the low intertidal to mesophotic zones (Abbott, 1999; Fumo & Sherwood, 2023).

Two of the four Hawaiian lineages described by Fumo & Sherwood (2023), Lineages 1 and 2, are hypothesized to form extensive sympatric populations around the island of O‘ahu; however, the extent of reproductive isolation (RI) between them and the mechanisms underlying their divergence remain unknown. In this study, we hypothesize that: (i) Lineages 1 and 2 consist of sympatric populations around O‘ahu; ii) strong barriers to gene flow maintain divergence; and (iii) historical allopatric

divergence has contributed to speciation. To test these hypotheses, we combined genome sequencing (ddRAD), fine-scale spatial transects, and demographic modeling. We first developed SNP panels to characterize lineages and their spatial distribution at multiple scales, then inferred patterns of gene flow and evaluated alternative speciation scenarios. Our results provide pioneering insights for macroalgal research, contributing to the limited number of genomic studies on marine speciation in Hawai‘i and highlighting the potential of *A. glomerata* to advance our understanding of these processes.

2. MATERIALS AND METHODS

2.1 Sample collection and identification

We collected 275 thalli of *Amansia glomerata* around the Hawaiian island of O‘ahu between March and July 2024 by sampling 5-10 cm branch tips without removing the holdfast. Sampling was conducted at nine sites via snorkeling or scuba diving. Six sites were sampled randomly within a 30 m radius, while three others were sampled along 30–40 m transects, collecting all thalli encountered. Samples were stored in individual bags and processed for epiphyte removal and morphological identification under light microscopy using diagnostic traits (Doty, 1971; Abbott, 1999). Sex determination was based on reproductive structures, with samples classified as tetrasporophytes (diploid phase) if bearing tetrasporangia, or as gametophytes (haploid phase) if bearing cystocarps. Only a few samples lacking reproductive structures were classified as

undetermined and removed from the study, along with three gametophytes. Due to the strong imbalance between gametophytes (~2%) and tetrasporophytes (~98%), investigations were limited to the diploid phase to avoid skewed representation of haploid genotypes. Individual samples were then flash-frozen and preserved for long-term storage, in the absence of diagnostic morphological traits to identify lineages, due in part to the high morphological plasticity of this species (Abbott, 1999; Phillips, 2009), lineage identity was considered undetermined until sequencing and SNP calling.

2.2 ddRAD library preparation and sequencing

Frozen tissues were manually ground in liquid nitrogen with a pestle, and genomic DNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle & Doyle, 1987). DNA extraction quality was assessed on agarose gels to confirm the presence of high-molecular-weight DNA. Double- stranded nucleic acid concentrations were quantified and standardized to 2.5 ng/µL using a Qubit 4 Fluorometer. We processed two ddRADseq libraries, each comprising 168 samples, following the protocol described by Reynes et al. (2021). The libraries included 214 samples from the present study that met DNA concentration requirements, along with 122 additional mesophotic *Amansia glomerata* specimens obtained from herbarium collections across the Hawaiian Archipelago as part of a separate project. Samples were randomized across populations, and four technical replicates (i.e., DNA from the same individual split into distinct libraries) were included. Briefly, the protocol involved

152 double-digesting 100 ng of DNA with *PstI* and *HhaI*, ligating adaptors containing 14
153 unique *PstI*-specific i5 barcodes, cleaning with AMPure beads, and amplifying
154 fragments by PCR to add Illumina adaptors with 12 unique i7 indices. Individual libraries
155 were checked on an agarose gel and pooled at equimolar concentrations (i.e., by
156 increasing the volume of low-concentration samples in the final pool). Libraries were
157 quality-checked using a Bioanalyzer at the Microbial Genomics and Analytical Lab
158 (MGAL, University of Hawai‘i) and size-selected to 300–500 bp using a BluePippin at
159 Azenta Life Sciences (Azenta Corporation, South Plainfield, NJ, USA). Because the two
160 libraries used the same set of i7 indices, sequencing was performed on two separate
161 lanes of a NovaSeq flow cell (2 × 150 bp), yielding approximately 750 Gb of data at
162 Azenta Life Sciences. After quality control, we obtained a total of 2.14 billion paired-
163 end reads (excluding 1.26 billion reads from a separate project), corresponding to an
164 average of 9.87 ± 3.75 million reads per sample (see Figure S1 for the distribution of
165 per-individual read counts).
166
167 2.3 Bioinformatic analyses
168 Individual libraries were quality-controlled, and reads were assembled into loci using
169 the de novo Stacks v.2.52 pipeline (Catchen et al., 2011; Rochette et al., 2019). To
170 account for intraspecific genetic variation, de novo assembly was performed on
171 populations from O‘ahu together with specimens from other Hawaiian Islands.
172 However, the catalog of loci was filtered to retain only polymorphisms segregating

within O‘ahu populations. Prior to de novo assembly, index-demultiplexed libraries were quality-checked with FASTQC v.0.11.7 (Andrews et al., 2010) and adapter-trimmed using Trimmomatic (Bolger et al., 2014). Individual-level demultiplexing was performed for each i7 index file using process_radtags, allowing up to one mismatch in the i5 barcode. Reads beginning with the enzyme cut-site motif were retained, and forward and reverse reads were resynchronized using Pairfq and trimmed to 137 bp. Stacks de novo assembly parameters were fine-tuned using a subset of 77 samples, representative of overall sample diversity (i.e., up to five samples per location), with RADStackhelpR (DeRaad, 2021) and vcfR (Knaus & Grünwald, 2017). This consisted of optimizing ustacks parameters by setting a minimum depth of five identical reads (m = 5) to form a stack and allowing up to five mismatches between stacks (M = 5) for them to be considered alleles of the same locus (see Figures S2-S3). The maximum number of mismatches allowed between homozygous individuals when constructing the catalog of homologous RAD loci (the n parameter) was set to n = 7, based on evaluation of lineage divergence and population clustering across a range of n values (see Figure S4). All individuals, except 42 with low sequencing output (i.e., <100,000 paired-end reads), were matched against the resulting catalog using sstacks, and genotypes were exported in VCF format using the populations module. We then post-filtered SNPs to retain those genotyped in at least 50% of individuals and with a maximum per-SNP heterozygosity of <50% across all samples (Table 1). SNPs segregating within the O‘ahu population were retrieved from this dataset using VCFtools v0.1.16 (Danecek et al.,

194 2011) in combination with the dDocent pipeline (Puritz et al., 2014) (see Table 1 for
195 details). Linkage disequilibrium and Hardy–Weinberg equilibrium analyses were
196 performed within lineages, based on lineage clustering observed from exploratory PCA
197 (Figure S4). The dataset comprised 58,563 SNPs across 174 individuals. Individual
198 missingness averaged 10.9% (±6.8%), and mean sequencing depth per individual was
199 46.9× (±21×). SNP calling was highly consistent across three replicate pairs, with a mean
200 SNP error rate of 0.008 (range: 0.003–0.017), estimated following the approach of
201 Mastretta-Yanes et al. (2015).
202
203 2.4 Lineage differentiation and hybrid detection
204 We employed principal component analysis (PCA) with the R package ADEGENET
205 (Jombart, 2008) to differentiate lineage and admixed genotypes. Lineage
206 differentiation was also investigated using the sNMF (sparse Non-Negative Matrix
207 Factorization) algorithm implemented in the R package LEA v.2.8 (Frichot et al., 2014;
208 Frichot & François, 2015). sNMF was run for 10,000 iterations and 10 replicates, with K
209 ranging from 1 to 10. We further defined lineage-specific SNPs as loci exhibiting high
210 differentiation (F_{ST} > 0.99) between parental lineages, characterized by individuals
211 showing nearly no admixed ancestry at K = 2 (Q_1 or Q_2 > 0.99). A total of 1,722 SNPs
212 were obtained and used to conduct fine-scale hybrid analyses with the R package
213 INTROGRESS v1.2.3 (Gompert & Buerkle, 2010). We also employed this SNP set in
214 HYBRIDLAB 1.0 (Nielsen et al., 2006) to simulate offspring genotypes (F1, F2, and up to

two generations of backcrosses) from parental lineages and to evaluate their power to differentiate among hybrid classes. We simulated 1,000 hybrid offspring and performed triangle plot analyses with INTROGRESS. Putative F_1 hybrids were defined as individuals with a hybrid index between 0.45 and 0.55 and interspecific heterozygosity greater than 0.75. Slightly admixed individuals were defined as those with a hybrid index deviating by more than 0.05 from either parental ancestry ($H_{index} > 0.05$ and $H_{index} < 0.95$) and interspecific heterozygosity higher than 0.05. We assigned lineage identity (Lineages 1 and 2), following Fumo & Sherwood (2023), by clustering the two reference samples from their study that were incorporated into our sequencing libraries.

2.5 Lineage spatial distribution and fine-scale differentiation

Lineage proportions were determined for each location to distinguish sympatric populations, in which both lineages co-occur, from parapatric populations, in which only a single lineage is present. Isolation by distance (IBD) was assessed by calculating pairwise F_{ST} (Weir & Cockerham, 1984) using StAMPP v1.6.3 (Pembleton et al., 2013) and by measuring least-cost coastal distances constrained by land using marmap (Pante & Simon-Bouhet, 2013). Pairwise F_{ST} was estimated within and between lineages, and the significance of IBD was assessed using Mantel tests with 1,000 permutations, implemented in the R package vegan (Dixon, 2003). Fine-scale differentiation patterns were examined along transects by calculating pairwise spatial distances (m) between individuals of the same lineage and between individuals of different lineages. Pairwise

236 genetic distances were calculated using the dissimilarity measure implemented in
237 poppr (Kamvar et al., 2014), and correlations with spatial distances were assessed for
238 each transect using Mantel tests with 1,000 permutations. Finally, we tested the null
239 hypothesis of no genetic structure within lineages by performing per-transect spatial
240 autocorrelation analyses using dartR v2.9 (Gruber et al., 2018). The autocorrelation
241 coefficient (r) was calculated for each pair of individuals across four distance classes
242 (m).
243
244 2.6 Demographic modeling
245 We employed the Demographic Inferences with Linked Selection (DILS) pipeline
246 (Csilléry et al., 2012; Pudlo et al., 2016; Fraïsse et al., 2021) to infer speciation scenarios
247 and test for heterogeneous gene flow across the genome. DILS, like other ABC
248 approaches, simulates genetic data under various demographic models and compares
249 the simulated data with the observed data using summary statistics. Specifically, in the
250 two-population mode, DILS uses a hierarchical model-selection procedure by
251 comparing models with current isolation (SI: strict isolation; AM: ancient migration)
252 versus ongoing migration (IM: isolation with migration; SC: secondary contact). DILS
253 offers several advantages over other methods by explicitly accounting for linked
254 selection and the resulting heterogeneity in introgression rates (Fraïsse et al., 2021).
255 We ran DILS five times, each using 1,000 loci randomly sampled without replacement
256 from the filtered catalog (see Table 1). Diploid sequences were extracted in FASTA

format using the STACKS populations module. Inferences using DILS were limited to the sympatric distribution, excluding F1 hybrids. Consistent priors and parameter settings were used across all runs (Table S1). We evaluated DILS performance by visualizing PCA results based on summary statistics from both simulations and the observed data to assess how well the simulations match the observed data (Figure S5). Parameters were then estimated from the optimized posterior of the best-supported model using both an ABC neural network and a random forest, as implemented in DILS (Fraïsse et al., 2021).

2.7 Introgression and barriers to gene flow

Barriers to gene flow, characterized by higher differentiation than the genomic background and, in contrast to introgressed regions, consistent with an excess of shared polymorphism, can be detected using DILS locus-specific models when heterogeneous gene flow (hetNm) is retained (Fraïsse et al., 2021). We used these locus-specific inferences to identify candidate SNPs associated with isolation or migration, retaining only those with a posterior probability ≥ 0.7; SNPs below this threshold were considered undetermined. We used metrics such as net divergence (D_a) and F_{ST}, estimated with DILS, to characterize genetic differentiation between the two sets of loci. SNPs associated with these two sets of loci were extracted to generate separate VCF files, representing loci associated with isolation versus migration. We then estimated the hybrid index for each set (migration-associated vs. isolation-

associated SNPs) using INTROGRESS. To investigate potential introgression between lineages, we examined allele frequencies and individual genotypes by extracting genotype matrices with vcfR (Knaus & Grünwald, 2017) for a subset of migration-associated SNPs selected for shared polymorphism in sympatry and high differentiation (F_{ST} > 0.95) between parapatric populations. This approach assumes that introgression is restricted to the sympatric range, with allopatric populations serving as a conservative control for detecting introgressed loci (see Gompert & Buerkle, 2010).

3. RESULTS

3.1 Lineage determination and admixture

Principal component analysis (PCA) of 58, 563 SNPs revealed two well-differentiated clusters, designated Lineages 1 and 2, based on the clustering of two individuals with known mitochondrial haplotypes (Fumo & Sherwood, 2023). The first principal component (PC 1) explained 20.87% of the total genotypic variance, captured lineage differentiation, and highlighted an intermediate individual from Mākaha (West Oʻahu) with high admixture (Figure 1A). Additional PCA axes (PC 2, PC 3, PC 4), which accounted for 5.51%, 3.05%, and 2.92% of the total variation, respectively, tended to differentiate individuals by geography (Figure S6). SNMF analyses identified two well-differentiated ancestry groups (Q_1 and Q_2; Figure 1B), consistent with the clusters highlighted by PCA, and detected a single intermediate individual with roughly equal proportions of Q_1 and Q_2 ancestry. We further identified 1,772 lineage-diagnostic SNPs (F_{ST} > 0.99; see

299 Methods), which are highly differentiated between parental lineages, to infer
300 hybridization and test whether intermediate individuals are consistent with F1 hybrids.
301 Across these SNPs, hybrid classes were well differentiated by interspecific
302 heterozygosity and hybrid index (Figure 2A), confirming that the observed intermediate
303 individual most likely resulted from an F1 cross (Figure 2B). Lineage-diagnostic SNPs
304 further indicated that individuals showing slight deviations from one parental ancestry
305 (i.e., hybrid index up to 0.12 and interspecific heterozygosity up to 0.17) were not
306 consistent with first- or second-backcross generations. In total, we detected 86
307 individuals from Lineage 1, including three that deviated from this threshold and were
308 identified as slightly admixed, and 87 individuals from Lineage 2, including seven
309 individuals that showed slight deviations from this threshold and one F1 hybrid (Table
310 S2). The F1 hybrid was excluded from demographic modeling, and slightly admixed
311 individuals were retained.
312
313 3.2 Spatial distribution and fine-scale differentiation
314 Analyses of spatial assemblage revealed that Lineages 1 and 2 are mostly sympatric
315 around O‘ahu, co-occurring at six of nine sampled sites (Figure 3A). Two parapatric
316 populations were also identified: one from Lineage 1 in the southeastern part of O‘ahu
317 (Popoia; Flat Island) and the other from Lineage 2 in the northern part of O‘ahu (Ke‘iki).
318 Mantel tests revealed significant isolation-by-distance (IBD) within lineages around the
319 island (Lineage 1: $R = 0.67$, $p < 0.001$; Lineage 2: $R = 0.71$, $p < 0.001$), but not between

lineages (Lineage 1-Lineage 2: R = –0.20, p = n.s.). The absence of IBD, when lineage differentiation is not accounted for, reflects consistently high pairwise F_{ST} values (0.6-0.8), regardless of whether populations were sampled from the same site or separated by more than 75 km of coastline (Figure 3B; Figure S7 for details of pairwise F_{ST} among populations). Fine-scale analyses along six transects showed that lineages co-occur within sites. The two lineages were generally separated by an average distance of 12.7 m (Figure 4A), although they were sometimes spatially close (< 1 m). Fine-scale IBD analyses indicated that intra-lineage differentiation was homogeneous at distances up to 40 m (Figure 4B, Table S3). Pairwise genetic distances indicated low intra-lineage differentiation (genetic distances < 0.10 for both lineages), higher differentiation between lineages (genetic distances > 0.30), and intermediate values for pairwise comparisons involving the F1 hybrid (pairwise genetic distances ≈ 0.20; Figure 4B). The absence of fine-scale IBD is supported by the absence of spatial autocorrelation, as independently tested for each transect and lineages (Figure S8).

3.3 Scenarios of speciation

We subsequently employed DILS to infer the most likely speciation scenario across five independent runs, each comprising 1,000 loci (including invariant positions) randomly sampled from the Stacks catalog. Analyses were restricted to the sympatric range, excluding individuals from Ke'iki and Flat Island (Popoia) as well as the F1 hybrid. DILS summary statistics indicated advanced lineage divergence, with D_a values of 0.013-

0.014, F_{ST} of 0.49-0.50, and 57% of loci containing at least one fixed difference (Table 2). Consistent differentiation patterns across these locus sets further supported genome-wide divergence rather than differentiation restricted to specific regions. We assessed model fit using the goodness-of-fit output generated by DILS. For each run, the best-supported model adequately fit the data. In the PCA summarizing goodness-of-fit for the selected model, the observed dataset clustered near the center of the optimized posterior distribution along PC1 and PC2 (Figure S5). This indicated that the summary statistics simulated under the best-fitting model closely matched those of the observed dataset. DILS analyses provided strong support for models of ongoing migration (posterior probability = 0.95-0.97), with secondary contact (SC) receiving higher posterior support (0.76-0.86) than isolation-with-migration (IM). DILS further indicated that migration is heterogeneous among loci, with high posterior probabilities for heterogeneous migration (hetNm, 0.91-0.93). Divergence times were estimated at 138,000-398,000 generations ago (T_{split}), with substantial variability in the timing of secondary contact (T_{sc}) across runs (Table 3). In run 3, which yielded the highest posterior model probability, the inferred T_{split} coincided with demographic change: Lineage 1 experienced an approximately seven-fold population expansion, reaching an effective population size of ~207,000 individuals (95% CI: 141,046-283,182), whereas Lineage 2 underwent a 0.55-fold reduction, declining to ~30,000 individuals (95% CI: 20,235-48,026) relative to the ancestral population.

3.4 Barriers and migration-associated loci

The retained heterogeneous migration model (hetNm) classified 37% of loci as isolation, 24% as migration, and 39% as ambiguous, using a posterior probability threshold of 0.7 in locus-specific models (Figure 5). Isolation-associated loci showed approximately 2.35-fold higher F_{ST}, 3.3-fold higher net divergence (D_a), and were associated with a 7.3-fold higher fraction of sites with fixed differences (S_f) relative to migration-associated loci (Figure 6). Introgression rates were predominantly asymmetrical, from Lineage 1 to Lineage 2 ($M_1 \rightarrow_2$) rather than in the opposite direction ($M_2 \rightarrow_1$). However, introgression from $M_2 \rightarrow_1$ was heterogeneous across loci, with a lower barrier to gene flow ($nBar_{2 \rightarrow 1}$; Table 3). We then estimated the hybrid index (H_{index}) for 2,525 isolation-associated SNPs and 920 migration-associated SNPs to test for significant differences in individual admixture (ΔH_{index}) between the two sets of loci. Significant differences were observed among Lineage 1 individuals, with isolation-associated SNPs showing lower admixture than migration-associated SNPs (ΔH_{index} = 0.040 [95% CI: 0.035-0.050]) but not among Lineage 2 individuals (ΔH_{index} = -0.004 [95% CI: -0.007 -0.002]) (Figure S9). Furthermore, ΔH_{index} was geographically determined, with significant differences among Lineage 1 populations ($\chi^2 = 58.05$, df = 8, $p < 0.001$), with western populations showing higher ΔH_{index}, except for the F_1 hybrid, which showed high consistency in admixture between these two sets of loci (ΔH_{index} = 0.002) (Figure S9).

382 Genomic signatures of introgression were assessed using a subset of candidate
383 migration-associated SNPs, selected for their excess of shared polymorphism in
384 sympatric populations and for high differentiation (F_{ST} > 0.95) between parapatric
385 populations. This analysis identified 51 candidate SNPs, with allele frequencies and
386 individual genotypes revealing a long tail of Lineage 2-diagnostic alleles introgressed
387 into Lineage 1 (Figure 7). Introgressed alleles exhibited a geographic cline, being most
388 frequent in western populations and progressively declining toward the eastern part of
389 the island. Their genotypic composition was also geographically determined, with
390 Lineage 2-specific alleles occurring predominantly in the homozygous state (1/1, blue)
391 in western populations, progressively shifting to heterozygous genotypes (0/1, yellow)
392 in southern populations, and maintaining at low frequency or even disappearing in the
393 eastern locations, where only Lineage 1-specific homozygotes (0/0, orange) were
394 detected (Figure 7).
395
396 **4. DISCUSSION**
397 By conducting one of the few genomic investigations of speciation in marine seaweeds,
398 we found evidence that lineage diversification in the red alga *Amansia glomerata*, is
399 driven by the evolution of strong reproductive isolation between lineages, resulting in
400 the persistence of sympatric populations. Below, we discuss speciation pathways,
401 emphasizing the key role of allopatric divergence and secondary episodes of gene flow
402 in shaping the genomic landscape.

403 Allopatric divergence in Hawaiian marine seaweeds

404 Elucidating the temporal dynamics of gene flow is key to understanding both
405 contemporary genomic landscapes and the mechanisms underlying reproductive
406 isolation (Butlin et al., 2008; Fitzpatrick et al., 2009; Ravinet et al., 2017). Here, we
407 found that Lineages 1 and 2 of *A. glomerata* are substantially differentiated (D_a = 1.3-
408 1.4%) despite their extensive sympatric distribution around O‘ahu, with demographic
409 modeling providing the strongest support for allopatric divergence followed by
410 secondary contact. Allopatric divergence was also supported by genome-wide analyses
411 of randomly sampled loci, which showed that divergence is relatively widespread rather
412 than concentrated in localized clusters (e.g., Martin et al., 2013; Carimán et al., 2025),
413 although the chromosomal context of these differentiated loci remains unknown in the
414 absence of a reference genome. Speciation in *A. glomerata* aligns with observations in
415 other marine species, in which allopatric divergence has been associated with the
416 evolution of strong reproductive isolation (e.g., Roux et al., 2013; Johannesson et al.,
417 2020a; De Jode et al., 2023). Demographic modeling indicated that the ancestral split
418 in *A. glomerata* occurred approximately 100,000–400,000 generations ago, consistent
419 with Pleistocene sea-level fluctuations during the Quaternary (2.6 million to 11,700
420 years ago), which shaped population differentiation in tropical marine taxa (Hewitt,
421 2000; Ludt & Rocha, 2015). Although precise knowledge of the species’ generation time
422 is lacking, limiting the exact dating of divergence and secondary contact, Pleistocene
423 sea-level fluctuations have generated cyclic inter-island connectivity in oceanic

archipelagos such as Hawai‘i (Ludt & Rocha, 2015; Weigelt et al., 2016; Hirschfeld et al., 2023). Composite Pleistocene landmasses, such as the O‘ahu Nui complex (O‘ahu and West Moloka‘i) and the Maui Nui complex (Carson & Clague, 1995), emerged during this period, creating land bridges and facilitating episodes of gene flow among previously isolated islands (Holland & Cowie, 2007; Jones & Jordan, 2015). In marine species with low dispersal potential, such as *A. glomerata*, as evidenced by significant isolation-by-distance (IBD) along the coastline, allopatric Pleistocene refuges have likely contributed to population divergence, although empirical evidence remains limited. Nevertheless, drift-driven differentiation is unlikely to be the sole mechanism driving speciation. On Hawai‘i Island (Big Island), Lineages 1 and 2 form parapatric populations, geographically defined and spatially separated from Lineage 4, which is restricted to the island's southern tip (Fumo & Sherwood, 2023). Diversification of Lineage 4 appears relatively recent, in line with the island's young substrate age (<0.1 Ma; Clague, 1996), and differentiation in allopatry seems unlikely given the absence of clear geographic barriers. These observations underscore the need for further studies to assess whether ecological differentiation has emerged alongside geographic isolation in *A. glomerata*.

Nature and evolution of reproductive isolation

Barriers to gene flow are most effectively studied where divergent populations are in contact (Johannesson et al., 2020b; Simon et al., 2021). Spatial distribution analyses indicated that Lineage 1 and 2 form extensive sympatric populations around O‘ahu,

445 while fine-scale transects show that lineages are inter-mixed within sites. This spatial
446 assemblage is consistent with what has been described as mosaic sympatry and pure
447 sympatry, depending on whether lineages specialize on different resources or whether
448 resources are otherwise perfectly mixed, respectively (see Figure 2 in Mallet et al.,
449 2009; Harrison & Larson, 2016). The absence of fine-scale IBD implies that lineage
450 differentiation is maintained despite the potential for dispersal and gene flow.
451 Furthermore, the detection of a single F1 hybrid and the absence of F2 individuals or
452 backcrosses are consistent with the evolution of strong reproductive isolation between
453 the parental forms, a pattern observed in other taxa at advanced stages of speciation
454 (e.g., Roux et al., 2013; Mikkelsen & Irwin, 2021). This interpretation is further
455 supported by the observation that approximately one-third of loci were inferred to be
456 isolated by DILS, indicating that a substantial portion of the genome contributes to
457 reproductive barriers. In *Amansia glomerata*, the near absence of F1 hybrids, despite
458 extensive sympatry, and the lack of backcrosses or recombinant hybrids suggest that
459 pre-zygotic barriers may favor assortative mating, potentially reinforced by
460 post-zygotic barriers, as observed in brown algae *Fucus* spp. (e.g., Monteiro et al., 2012;
461 Hoarau et al., 2015). The preliminary observation that isolated loci did not result in
462 limited admixture in the F1 hybrid is consistent with the dominance theory of
463 postzygotic isolation (Turelli et al., 2001), where Dobzhansky-Muller incompatibilities
464 (DMIs, Dobzhansky, 1937; Muller, 1942) did not affect the F1 diploid phase, with the
465 expression of recessive incompatibilities in subsequent generations (F2 or backcrosses)

466 preventing cross-species gene flow (e.g., Bierne et al., 2006; Roux et al., 2013). The
467 effect of DMIs is predicted to be stronger in haplodiploid species, where DMI alleles are
468 predicted to be strongly counterselected in the haploid phase (Nouhaud et al., 2020).
469 DMI alleles are thought to arise from the independent evolution of allopatric
470 populations (Orr & Turelli, 2001) and have been shown to evolve rapidly in marine
471 systems. For instance, marine mussel species experiencing relatively short periods of
472 geographic isolation (1-2 Ma) can accumulate DMIs quickly (Bierne et al., 2006).
473 However, a more detailed characterization of the barriers will require additional work,
474 including cross-experiments and admixture analyses in gametophytes.
475
476 Differential introgression
477 In the absence of evidence for ongoing gene flow, shared polymorphisms may reflect
478 either the retention of ancestral variation (Clark, 1997) or differential introgression
479 following secondary contact. The identification of migration-associated SNPs with
480 alleles shared among sympatric populations and highly differentiated in parapatric
481 populations provides a powerful framework for distinguishing these processes. Analysis
482 of these SNPs revealed a long tail of Lineage 2-diagnostic alleles introgressed into
483 Lineage 1, suggesting that secondary introgression, rather than ancestral
484 polymorphism, accounts for this pattern. Additional support came from the geographic
485 structure of the introgression: it occurred at high frequencies in western populations,
486 shifted progressively to heterozygous genotypes in southern populations, and

remained at low frequency or even disappeared in eastern populations. Such a geographic and genetic cline of introgressed alleles has been observed in marine populations resulting from secondary contact (e.g., Roux et al., 2013; Simon et al., 2021), with several non-mutually exclusive mechanisms, including asymmetric mating opportunities, hybrid zone movement, or adaptive introgression (Borge et al., 2005; Whitney et al., 2010; Yang et al., 2018) involved to explain its asymmetric nature. DILS analyses showed that demographic dynamics may contribute to this pattern, with Lineage 1 experiencing an approximately seven-fold increase in Ne, compared to a 0.55-fold reduction for Lineage 2 relative to its ancestor. This demographic dynamic, differing between lineages, may have caused shifts in population ranges and brought the expanding lineage into contact with the declining one. As a result, a demographic wave front of the advancing lineage could have moved into the retreating lineage's range, carrying selectively neutral alleles from the retreating lineage forward and gradually increasing their frequency behind the advancing front. This process is predicted to produce asymmetrical introgression (Currat et al., 2008). Nonetheless, caution is warranted when interpreting introgression, as loci assumed to be diagnostic may lose their diagnostic power when genotypes from a broader geographic context are considered (e.g., Simon et al., 2021).

In this pioneering genomic investigation of speciation in Hawaiian seaweeds, we show that lineage differentiation in *Amansia glomerata* is maintained by strong reproductive

isolation, which limits ongoing gene flow and allows the persistence of sympatric populations around O‘ahu. Despite limited knowledge of the species’ biology and the absence of genomic resources, ddRADseq and *de novo* assembly provided sufficient resolution in the genomic landscape to support allopatric divergence and secondary contact, while revealing signatures of differential introgression. We argue that *A. glomerata* represents a promising system for studying speciation, with the Hawaiian Archipelago offering an exceptional natural setting for such investigations. Future work will be needed to test the role of ecological differentiation and to determine whether sympatric populations differ in habitat use or resource specialization. These advances will require developing refined genomic resources and broader population sampling across the archipelago, including both ploidy phases. Overall, our study highlights the substantial potential of marine seaweeds to advance our understanding of speciation.

Acknowledgements

This work was supported by the U.S. National Science Foundation (DEB-2242142). Molecular analyses were additionally supported by the Microbial Genomics and Analytical Laboratory (MGAL) through NIH/NIGMS grant P20 GM125508. We thank all collectors who assisted with field work, especially Edoardo Sena, Noah Doeden, Callie Stephenson, Pearl Thompson, and Madison Tracy. We thank Kazumi Allsopp, Sophie Paradis, and Allexa Burger for providing reliable lab guidance and support with sample

processing. We thank the University of Hawai‘i System High Performance Computing for providing support for data storage and analysis.

Author Contributions

L.R. conceptualized the study, performed laboratory work and analyses, generated the figures, and drafted the manuscript. J.T.F. and L.R. conducted fieldwork. A.R.S. secured funding for the project and contributed to study conceptualization and manuscript preparation. J.T.F. provided feedback on the analyses. All authors contributed to writing and editing the manuscript and approved the final version.

Competing Interests

The authors declare no competing financial interests.

Data Archiving

Individual raw reads are available in FASTQ.gz format in NCBI (BioProject PRJNA1425114). Metadata are provided in CSV format within the same BioProject, with one row corresponding to each sample and include the following columns: sample ID, laboratory identification number (ARS), sequencing index barcode, barcode sequence, sampling site and coordinates, sampling date, transect information, DNA extraction date, and DNA concentration measured using Qubit. Additionally, results of downstream analyses, including lineage identity, sNMF ancestry coefficients, and

546 hybrid index from INTROGRESS, are provided in the same table. A VCF file, consisting
547 of 58,563 SNPs employed for speciation inferences (step 7, Table 1), FASTA files
548 containing the locus datasets used for demographic modeling with DILS, with the R
549 script used to generate them, will be made accessible in a Dryad Repository (XXX).

550

551 References

552 Abbott IA (1999). *Marine Red Algae of the Hawaiian Islands*. Bishop Museum Pr:
553 Honolulu, Hawai'i.
554 Andrews S, Krueger F, Segonds-Pichon A, Biggins L, Krueger C, Wingett S (2010). FastQC.
555 *A quality control tool for high throughput sequence data* **370**.
556 Bierne N, Bonhomme F, Boudry P, Szulkin M, David P (2006). Fitness landscapes support
557 the dominance theory of post-zygotic isolation in the mussels Mytilus edulis
558 and M. galloprovincialis. *Proceedings of the Royal Society B: Biological Sciences*
559 **273**: 1253–1260.
560 Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina
561 sequence data. *Bioinformatics* **30**: 2114–2120.
562 Borge T, Lindroos K, Nádvorník P, Syvänen A-C, Sætre G -P. (2005). Amount of
563 introgression in flycatcher hybrid zones reflects regional differences in pre and
564 post-zygotic barriers to gene exchange. *j evol Biol* **18**: 1416–1424.
565 Butlin RK, Galindo J, Grahame JW (2008). Sympatric, Parapatric or Allopatric: The Most
566 Important Way to Classify Speciation? *Philosophical Transactions: Biological*
567 *Sciences* **363**: 2997–3007.
568 Carimán P, Guillemin M-L, Giles EC, Narváez G, Suescún AV, Sáenz-Agudelo P (2025).
569 Exploring evolutionary mechanisms of genomic divergence in marine intertidal
570 limpets. *Heredity* **134**: 500–511.
571 Carson HL, Clague DA (1995). Geology and biogeography. Hawaiian biogeography:
572 evolution on a hotspot archipelago. *Hawaiian biogeography: evolution on a*
573 *hotspot archipelago*: 14–29.
574 Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011). Stacks: building
575 and genotyping Loci de novo from short-read sequences. *G3 (Bethesda)* **1**: 171–
576 182.
577 Clague DA (1996). The growth and subsidence of the Hawaiian-Emperor volcanic chain.
578 *The origin and evolution of Pacific island biotas, New Guinea to eastern*

579 *Polynesia: patterns and processes*: 35–50.
580 Clark AG (1997). Neutral behavior of shared polymorphism. *Proceedings of the National*
581 *Academy of Sciences* **94**: 7730–7734.
582 Coughlan JM, Matute DR (2020). The importance of intrinsic postzygotic barriers
583 throughout the speciation process. *Philosophical Transactions of the Royal*
584 *Society B: Biological Sciences* **375**: 20190533.
585 Coyne JA, Orr HA (2004). Speciation Sinauer Associates. *Sunderland, MA* **276**: 854.
586 Csilléry K, François O, Blum MGB (2012). abc: an R package for approximate Bayesian
587 computation (ABC). *Methods in Ecology and Evolution* **3**: 475–479.
588 Currat M, Ruedi M, Petit RJ, Excoffier L (2008). The hidden side of invasions: massive
589 introgression by local genes. *Evol* **62**: 1908–1920.
590 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, *et al.* (2011). The
591 variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.
592 De Jode A, Le Moan A, Johannesson K, Faria R, Stankowski S, Westram AM, *et al.* (2023).
593 Ten years of demographic modelling of divergence and speciation in the sea.
594 *Evolutionary Applications* **16**: 542–559.
595 Denoeud F, Godfroy O, Cruaud C, Heesch S, Nehr Z, Tadrent N, *et al.* (2024).
596 Evolutionary genomics of the emergence of brown algae as key components of
597 coastal ecosystems. *Cell* **187**: 6943-6965.e39.
598 DeRaad DA (2021). RADStackshelpR.
599 Dieckmann U (2004). *Adaptive Speciation*. Cambridge University Press.
600 Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of*
601 *Vegetation Science* **14**: 927.
602 Dobzhansky T (1937). *Genetics and the Origin of Species*. Columbia University Press.
603 Doty MS (1971). *The productivity of benthic frondose algae at Waikiki Beach, 1967-*
604 *1968*. University of Hawaii.
605 Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh
606 leaf tissue. *PHYTOCHEMICAL BULLETIN*.
607 Faria R, Johannesson K, Stankowski S (2021). Speciation in marine environments: Diving
608 under the surface. *j evol Biol* **34**: 4–15.
609 Feder JL, Nosil P (2010). The efficacy of divergence hitchhiking in generating genomic
610 islands during ecological speciation. *Evolution* **64**: 1729–1747.
611 Feder JL, Egan SP, Nosil P (2012). The genomics of speciation-with-gene-flow. *Trends in*
612 *Genetics* **28**: 342–350.
613 Fitzpatrick BM, Fordyce JA, Gavrilets S (2009). Pattern, process and geographic modes
614 of speciation. *j evol Biol* **22**: 2342–2347.
615 Fraïsse C, Popovic I, Mazoyer C, Spataro B, Delmotte S, Romiguier J, *et al.* (2021). DILS:
616 Demographic inferences with linked selection by using ABC. *Molecular Ecology*
617 *Resources* **21**: 2629–2644.
618 Frichot E, Mathieu F, Trouillon T, Bouchard G, François O (2014). Fast and Efficient
619 Estimation of Individual Ancestry Coefficients. *Genetics* **196**: 973–983.
620

621 Frichot E, François O (2015). LEA: An R package for landscape and ecological association
622 studies. *Methods in Ecology and Evolution* **6**: 925–929.
623 Fumo JT, Sherwood AR (2023). Phylogeography of *Amansia glomerata* C.Agardh
624 (Ceramiales, Rhodomelaceae) in Hawai'i: A Single Species with High
625 Divergence. *crya* **44**: 85–100.
626 Gompert Z, Buerkle AC (2010). Introgress: a software package for mapping components
627 of isolation in hybrids. *Molecular Ecology Resources* **10**: 378–384.
628 Gruber B, Unmack PJ, Berry OF, Georges A (2018). dartr: An r package to facilitate
629 analysis of SNP data generated from reduced representation genome
630 sequencing. *Molecular Ecology Resources* **18**: 691–699.
631 Harrison RG, Larson EL (2016). Heterogeneous genome divergence, differential
632 introgression, and the origin and structure of hybrid zones. *Molecular Ecology*
633 **25**: 2454–2466.
634 Heesch S, Serrano-Serrano M, Barrera-Redondo J, Luthringer R, Peters AF, Destombe C,
635 *et al.* (2021). Evolution of life cycles and reproductive traits: Insights from the
636 brown algae. *Journal of Evolutionary Biology* **34**: 992–1009.
637 Hewitt G (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.
638 Hirschfeld M, Barnett A, Sheaves M, Dudgeon C (2023). What Darwin could not see:
639 island formation and historical sea levels shape genetic divergence and island
640 biogeography in a coastal marine species. *Heredity* **131**: 189–200.
641 Hoarau G, Coyer JA, Giesbers MCWG, Jueterbock A, Olsen JL (2015). Pre-zygotic
642 isolation in the macroalgal genus Fucus from four contact zones spanning 100–
643 10 000 years: a tale of reinforcement? *Royal Society Open Science* **2**: 140538.
644 Holland BS, Cowie RH (2007). A geographic mosaic of passive dispersal: population
645 structure in the endemic Hawaiian amber snail Succinea caduca (Mighels,
646 1845). *Molecular Ecology* **16**: 2422–2435.
647 Hudson J, McQuaid CD, Rius M (2021). Contemporary climate change hinders hybrid
648 performance of ecologically dominant marine invertebrates. *j evol Biol* **34**: 60–
649 72.
650 Johannesson K, Panova M, Kemppainen P, André C, Rolán-Alvarez E, Butlin RK (2010).
651 Repeated evolution of reproductive isolation in a marine snail: unveiling
652 mechanisms of speciation. *Philos Trans R Soc Lond B Biol Sci* **365**: 1735–1747.
653 Johannesson K, Butlin RK, Panova M, Westram AM (2020a). Mechanisms of Adaptive
654 Divergence and Speciation in Littorina saxatilis: Integrating Knowledge from
655 Ecology and Genetics with New Data Emerging from Genomic Studies. In:
656 Oleksiak MF, Rajora OP (eds) *Population Genomics: Marine Organisms*,
657 Springer International Publishing: Cham, pp 277–301.
658 Johannesson K, Moan AL, Perini S, André C (2020b). A Darwinian Laboratory of Multiple
659 Contact Zones. *Trends in Ecology & Evolution* **35**: 1021–1036.
660 Johannesson K, Faria R, Moan AL, Rafajlović M, Westram AM, Butlin RK, *et al.* (2024).
661 Diverse pathways to speciation revealed by marine snails. *Trends in Genetics*
662 **40**: 337–351.

663 Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic
664 markers. *Bioinformatics* **24**: 1403–1405.
665 Jones BR, Jordan S (2015). Genetic Consequences of Pleistocene Sea-Level Change on
666 Hawaiian Megalagrion Damselflies. *J Hered* **106**: 618–627.
667 Kamvar ZN, Tabima JF, Grünwald NJ (2014). Poppr: an R package for genetic analysis of
668 populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**:
669 e281.
670 Kay EA, Palumbi SR (1987). Endemism and evolution in Hawaiian marine invertebrates.
671 *Trends in Ecology & Evolution* **2**: 183–186.
672 Knaus BJ, Grünwald NJ (2017). VCFR : a package to manipulate and visualize variant call
673 format data in R. *Molecular Ecology Resources* **17**: 44–53.
674 Krueger-Hadfield SA (2024). Let's talk about sex: Why reproductive systems matter for
675 understanding algae. *Journal of Phycology* **60**: 581–597.
676 Ludt WB, Rocha LA (2015). Shifting seas: the impacts of Pleistocene sea-level
677 fluctuations on the evolution of tropical marine taxa. *Journal of Biogeography*
678 **42**: 25–38.
679 Mallet J, Meyer A, Nosil P, Feder JL (2009). Space, sympatry and speciation. *j evol Biol*
680 **22**: 2332–2341.
681 Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, *et al.* (2013).
682 Genome-wide evidence for speciation with gene flow in Heliconius butterflies.
683 *Genome Res* **23**: 1817–1828.
684 Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015).
685 Restriction site-associated DNA sequencing, genotyping error estimation and
686 *de novo* assembly optimization for population genetic inference. *Molecular*
687 *Ecology Resources* **15**: 28–41.
688 Mayr E (1954). Geographic Speciation in Tropical Echinoids. *Evolution* **8**: 1–18.
689 Miglietta MP, Faucci A, Santini F (2011). Speciation in the Sea: Overview of the
690 Symposium and Discussion of Future Directions. *Integr Comp Biol* **51**: 449–455.
691 Mikkelsen EK, Irwin D (2021). Ongoing production of low-fitness hybrids limits range
692 overlap between divergent cryptic species. *Molecular Ecology* **30**: 4090–4102.
693 Montecinos AE, Guillemin M, Couceiro L, Peters AF, Stoeckel S, Valero M (2017).
694 Hybridization between two cryptic filamentous brown seaweeds along the
695 shore: analysing pre- and postzygotic barriers in populations of individuals with
696 varying ploidy levels. *Molecular Ecology* **26**: 3497–3512.
697 Monteiro CA, Serrão EA, Pearson GA (2012). Prezygotic barriers to hybridization in
698 marine broadcast spawners: reproductive timing and mating system variation.
699 *PLoS One* **7**: e35978.
700 Muller HJ (1942). *Isolating mechanisms, evolution, and temperature.* Biol. Symp.
701 Nielsen EE, Bach LA, Kotlicki P (2006). hybridlab (version 1.0): a program for generating
702 simulated hybrids from population samples. *Molecular Ecology Notes* **6**: 971–
703 973.
704 Nosil P (2012). *Ecological Speciation*. Oxford University Press.

Nouhaud P, Blanckaert A, Bank C, Kulmuni J (2020). Understanding Admixture: Haplodiploidy to the Rescue. *Trends in Ecology & Evolution* **35**: 34–42.

Orr HA, Turelli M (2001). The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**: 1085–1094.

Palumbi SR (1992). Marine speciation on a small planet. *Trends in Ecology & Evolution* **7**: 114–118.

Palumbi SR (1994). Genetic Divergence, Reproductive Isolation, and Marine Speciation. *Annual Review of Ecology and Systematics* **25**: 547–572.

Pante E, Simon-Bouhet B (2013). marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. *PLOS ONE* **8**: e73051.

Pembleton LW, Cogan NOI, Forster JW (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* **13**: 946–952.

Phillips L (2009). *The taxonomy of selected tribes and genera of the Rhodomelaceae*. Lambert Academic Publishing.

Potkamp G, Fransen CHJM (2019). Speciation with gene flow in marine systems. *Contributions to Zoology* **88**: 133–172.

Price JP, Clague DA (2002). How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. *Proceedings of the Royal Society of London Series B: Biological Sciences* **269**: 2429–2435.

Pudlo P, Marin J-M, Estoup A, Cornuet J-M, Gautier M, Robert CP (2016). Reliable ABC model choice via random forests. *Bioinformatics* **32**: 859–866.

Puritz JB, Hollenbeck CM, Gold JR (2014). dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* **2**: e431.

Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, *et al.* (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30**: 1450–1477.

Reynes L, Aurelle D, Chevalier C, Pinazo C, Valero M, Mauger S, *et al.* (2021). Population Genomics and Lagrangian Modeling Shed Light on Dispersal Events in the Mediterranean Endemic Ericaria zosteroides (=Cystoseira zosteroides) (Fucales). *Front Mar Sci* **8**.

Rochette NC, Rivera-Colón AG, Catchen JM (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology* **28**: 4737–4754.

Roderick GK, Gillespie RG (1998). Speciation and phylogeography of Hawaiian terrestrial arthropods. *Molecular Ecology* **7**: 519–531.

Rotondo GM, Springer VG, Scott GAJ, Schlanger SO (1981). Plate Movement and Island Integration—A Possible Mechanism in the Formation of Endemic Biotas, with Special Reference to the Hawaiian Islands. *Syst Biol* **30**: 12–21.

Roux C, Tsagkogeorga G, Bierne N, Galtier N (2013). Crossing the Species Barrier: Genomic Hotspots of Introgression between Two Highly Divergent Ciona intestinalis Species. *Mol Biol Evol* **30**: 1574–1587.

747 Rundle HD, Nosil P (2005). Ecological speciation. *Ecology Letters* **8**: 336–352.
748 Schluter D (2009). Evidence for Ecological Speciation and Its Alternative. *Science* **323**:
749 737–741.
750 Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, *et al.* (2014).
751 Genomics and the origin of species. *Nat Rev Genet* **15**: 176–192.
752 Sherwood AR, Kurihara A, Conklin KY (2011). Molecular diversity of Amansieae
753 (Ceramiales, Rhodophyta) from the Hawaiian Islands: a multi-marker
754 assessment reveals high diversity within Amansia glomerata. *Phycological*
755 *Research* **59**: 16–23.
756 Sherwood AR, Guiry MD (2023). Inventory of the Seaweeds and Seagrasses of the
757 Hawaiian Islands. *Biology* **12**: 215.
758 Simon A, Fraïsse C, El Ayari T, Liautard-Haag C, Strelkov P, Welch JJ, *et al.* (2021). How
759 do species barriers decay? Concordance and local introgression in mosaic
760 hybrid zones of mussels. *j evol Biol* **34**: 208–223.
761 Turelli M, Barton NH, Coyne JA (2001). Theory and speciation. *Trends in Ecology &*
762 *Evolution* **16**: 330–343.
763 Valero M, Richerd S, Perrot V, Destombe C (1992). Evolution of alternation of haploid
764 and diploid phases in life cycles. *Trends in Ecology & Evolution* **7**: 25–29.
765 VanderWerf EA, Young LC, Yeung NW, Carlon DB (2010). Stepping stone speciation in
766 Hawaii's flycatchers: molecular divergence supports new island endemics
767 within the elepaio. *Conserv Genet* **11**: 1283–1298.
768 Via S, West J (2008). The genetic mosaic suggests a new role for hitchhiking in ecological
769 speciation. *Molecular Ecology* **17**: 4334–4345.
770 Wagner WL, Funk VA (1995). *Hawaiian biogeography : evolution on a hot spot*
771 *archipelago*. Smithsonian Institution Press: Washington.
772 Weigelt P, Steinbauer MJ, Cabral JS, Kreft H (2016). Late Quaternary climate change
773 shapes island biodiversity. *Nature* **532**: 99–102.
774 Weir BS, Cockerham CC (1984). Estimating F-Statistics for the Analysis of Population
775 Structure. *Evolution* **38**: 1358–1370.
776 Whitney KD, Randell RA, Rieseberg LH (2010). Adaptive introgression of abiotic
777 tolerance traits in the sunflower Helianthus annuus. *New Phytologist* **187**: 230–
778 239.
779 Yang W, While GM, Laakkonen H, Sacchi R, Zuffi MAL, Scali S, *et al.* (2018). Genomic
780 evidence for asymmetric introgression by sexual selection in the common wall
781 lizard. *Molecular Ecology* **27**: 4213–4224.
782

783

784

Figure Legends

Figure 1. Lineage differentiation across 58,563 SNPs and 174 selected individuals. (A) Principal Component Analysis (PCA) and (B) individual ancestry coefficients at K = 2 using sNMF. Each point on the PCA represents a single individual, and the colors of the dots correspond to their sNMF ancestry coefficient (Q_1).

Figure 2. Triangle plots, showing the relationship between hybrid index and interspecific heterozygosity across 1,772 lineage-specific SNPs. (A) Triangle plot generated with 1,000 simulated hybrid offspring (F1, F2, backcrosses) and (B) observed genotypes. In panel A, colors indicate the two parental lineages and hybrid classes, as shown in the legend on the right, while in panel B, colors correspond to the sNMF ancestry coefficient (Q_1).

Figure 3. (A) Lineage distributions around O‘ahu, with (B) correlations between F_{ST} and least-cost distance (km) for Lineage 1 (orange), Lineage 2 (blue), and between populations of different lineages (gray).

Figure 4. Fine-scale analyses along transect. (A) Distribution of pairwise spatial distances, distinguishing within- and between-lineage comparisons, and (B) correlations between spatial distance (m) and genetic distances for Lineage 1 (orange), Lineage 2 (blue), pairwise comparisons with the F1 hybrid (green), and comparisons between individuals of different lineages (grey).

804 Figure 5. DILS locus-specific models showing a positive correlation between F_{ST} and D_{XY}.
805 Loci are classified as migration-associated and isolation-associated using a posterior
806 probability threshold of 0.7.

807 Figure 6. Genomic landscape of differentiation summarized with per-locus distribution
808 of F_{ST}, net divergence (D_a), and the fraction of sites with fixed differences (S_f) using DILS
809 summary statistics. Panels A-C show the per-locus raw distributions, while panels D-F
810 distinguish loci classified as isolation (red) or migration (blue) by DILS locus-specific
811 models. Loci with posterior probability <0.7 were considered ambiguous and are
812 excluded from panels D-F for clarity.

813 Figure 7. Genomic signatures of differential introgression across 51 candidate SNPs. (A)
814 Individual genotypes ordered by lineage and geographic location. (B) Allele frequencies
815 showing geographic clinal patterns and strong asymmetry from Lineage 2 into Lineage
816 1. Individual samples are first grouped by lineage (orange = Lineage 1, blue = Lineage 2)
817 and then by location. Colors indicate genotype classes: orange = homozygote for
818 Lineage 1, blue = homozygote for Lineage 2, and yellow = heterozygote.

819

820

821

822 **Table Legends**

823 Table 1. Summary of SNP filtering steps, software, and criteria used to generate the
824 final dataset.

825

826 Table 2. Best demographic models assessed in DILS across five datasets, each consisting
827 of 1,000 loci randomly sampled without replacement. DILS summary statistics (D_a and
828 F_{ST}) are reported.

829

830 Table 3. Parameter estimates from DILS for each dataset, showing the estimated values
831 and 95% confidence intervals (CIs).

832

833

834

835

836

837

838

839

840

841

842

Table 1

Step	Software	Criteria	n (samples)	SNPs
1	Stacks: population	Initial filtering to retain informative loci in the Stacks catalog: loci genotyped in ≥ 50% of samples *(--R*); maximum heterozygosity per SNP < 50% *(--Het*).	291	731,288
2	VCFtools and AWK	Remove samples with high missing rates up to 80% *(--missing-indv*) before filtering SNPs	226	731,288
3	dDocent, pop_missing_filter.sh	Building the O'ahu dataset: SNPs with a call rate ≥ 0.90 within each region (*percent_missing_per_pop*) and present in all four O'ahu regions (E, S, W, N; *number_of_pops_for_cutoff*).	188	314,319
4	VCFtools	Keep only sites where the mean depth (DP) is ≥ 10 *(--min-meanDP*) and ≥ 100 (*--max-meanDP*)	188	231,314
5	VCFtools and AWK	Setting maximal individual missingness to ≤30% *(--missing-indv*) and minor allele frequency (MAF) ≥0.01 (*--maf*)	174	99,755
6	Plink and AWK	LD pruning: r^2 estimated within lineages, before excluding SNPs with $r^2 \geq 0.3$ in the two lineages *(--indep-pairwise 10 5 0.3*)	174	59,083
7	dDocent, filter_hwe_by_pop.pl	Remove SNPs showing significant deviation from HWE (P < 0.01; *--hwe*) within each lineage in ≥25% of populations (*--cutoff*)	174	58,563

Table 2

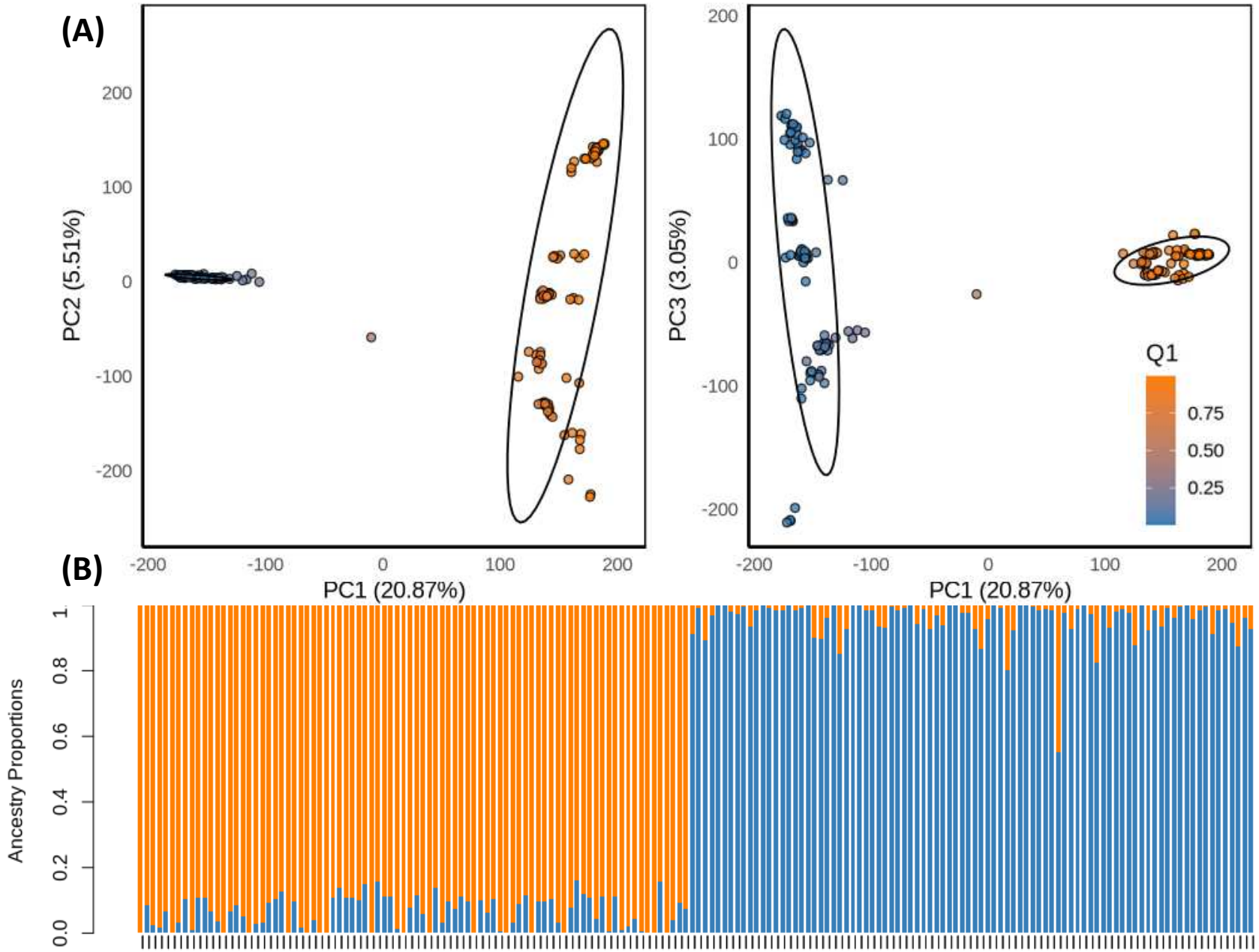
Dataset	D_a	F_{ST}	General model	Submodel	M-homo vs M-hetero	N-homo vs N-hetero
1	0.013	0.49	Migration (0.97)	SC (0.76)	M-hetero (0.91)	N-hetero (0.74)
2	0.014	0.5	Migration (0.97)	SC (0.79)	M-hetero (0.95)	N-hetero (0.77)
3	0.013	0.49	Migration (0.97)	SC (0.82)	M-hetero (0.96)	N-hetero (0.75)
4	0.013	0.49	Migration (0.96)	SC (0.76)	M-hetero (0.94)	N-hetero (0.66)
5	0.014	0.5	Migration (0.95)	SC (0.86)	M-hetero (0.93)	N-hetero (0.87)

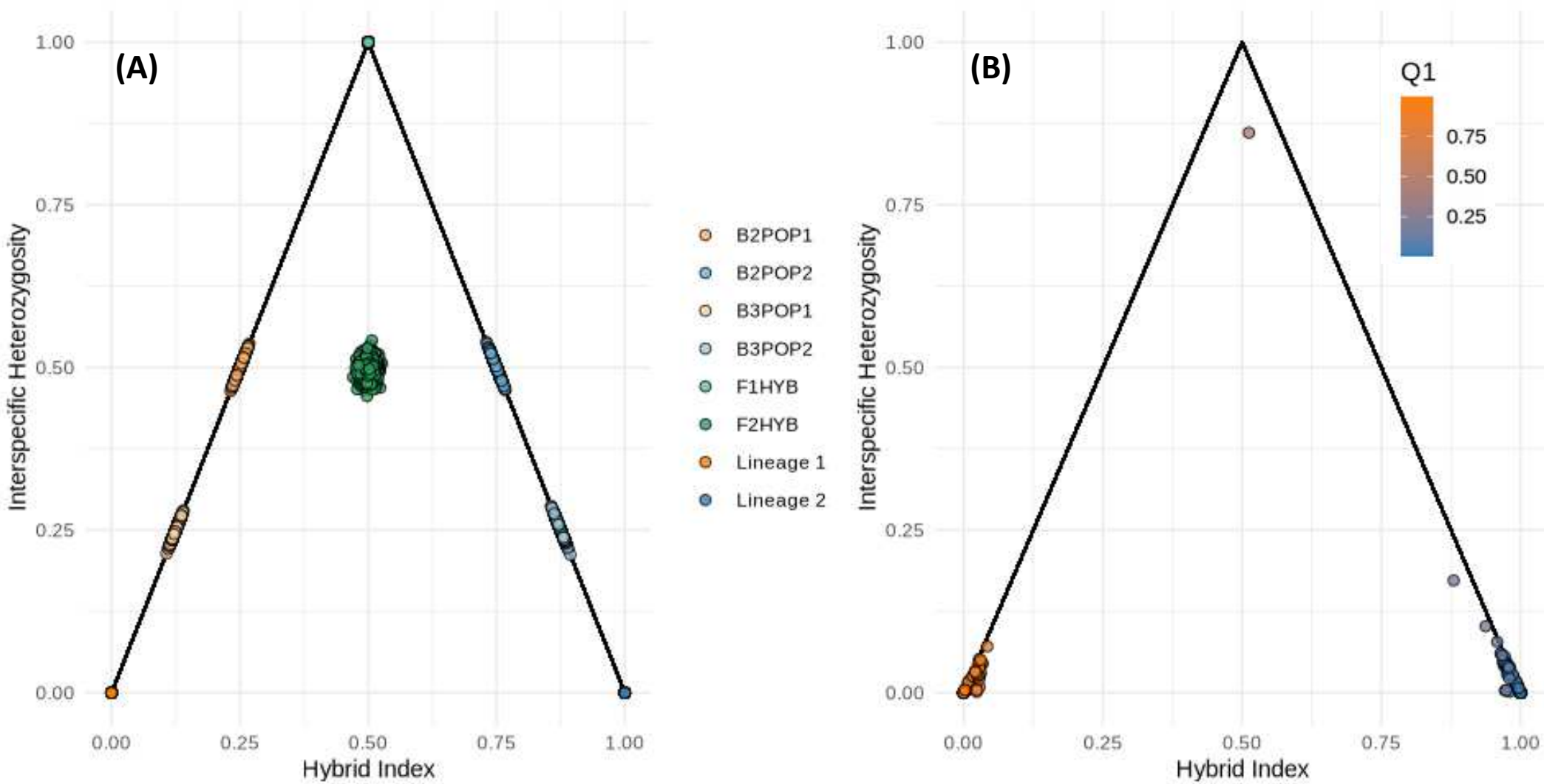
The number in brackets represents the posterior probability for each model, ranging from 0 to 1. N-homo/N-hetero: heterogeneous or homogenous population effective size. M-homo vs M-hetero: heterogeneous or homogenous migration rate. SC: Secondary Contact.

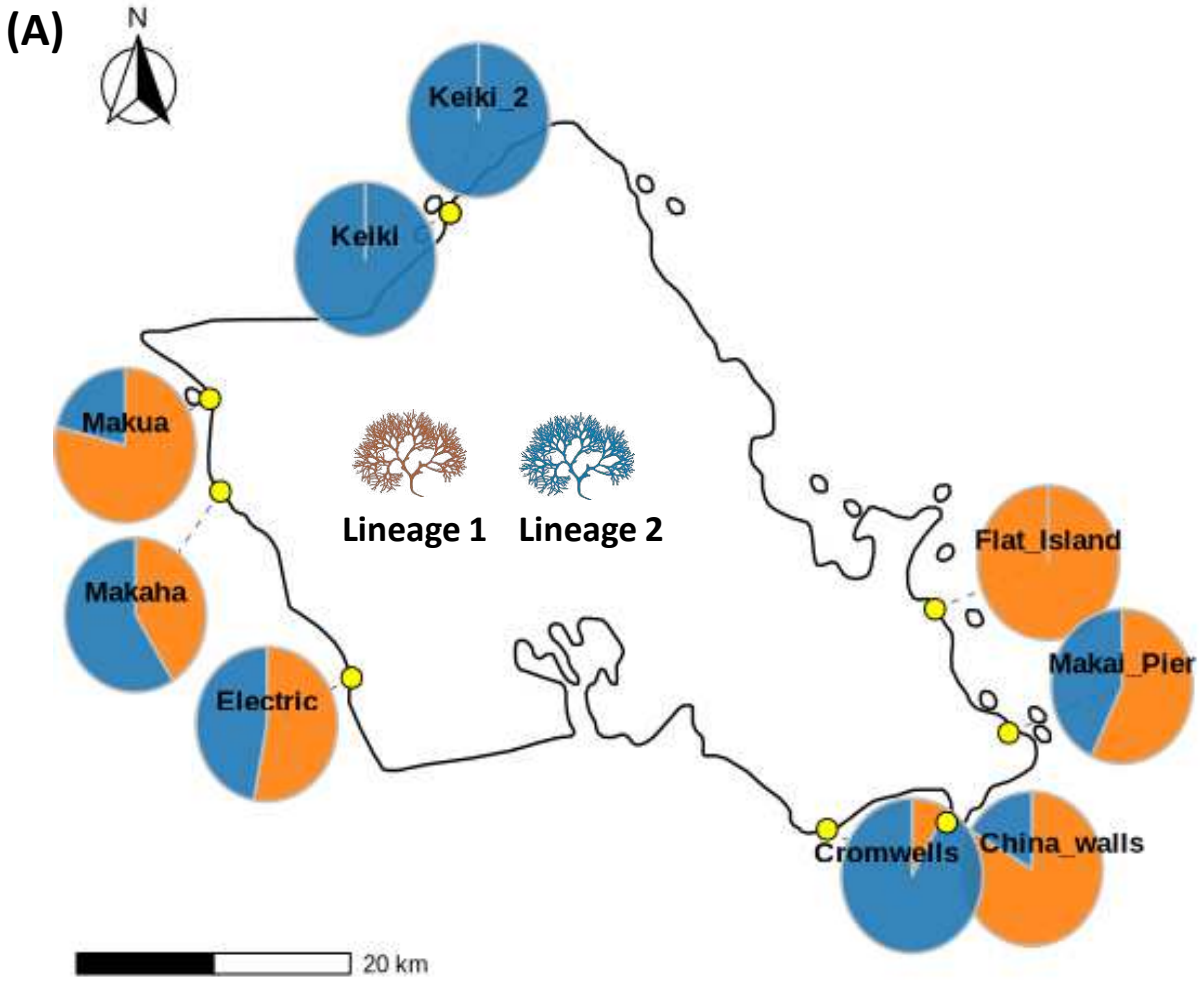
Table 3

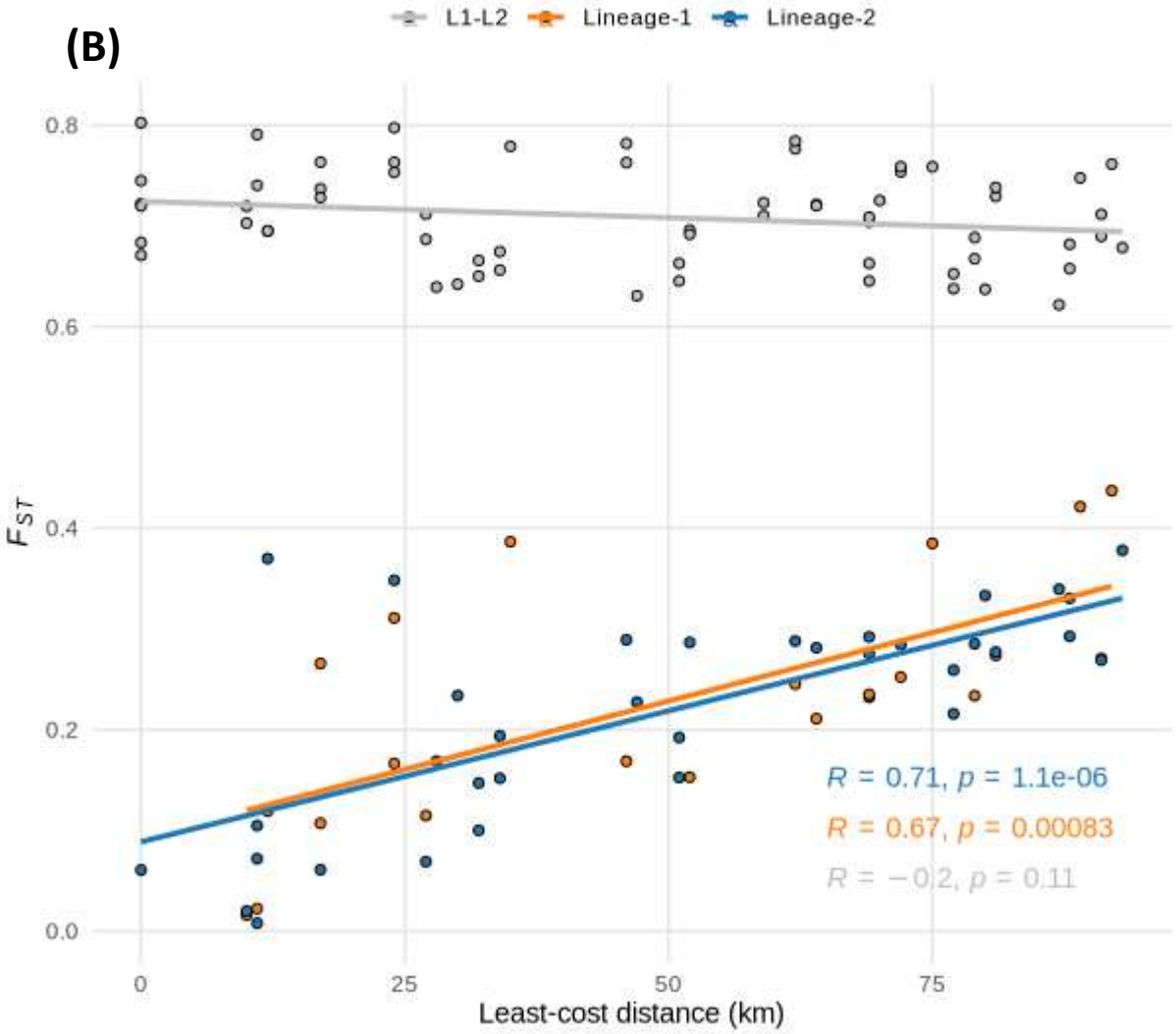
Dataset	T_{split}	T_{sc}	$M_1 \rightarrow_2$	$M_2 \rightarrow_1$	$n_{Bar1 \rightarrow 2}$	$n_{Bar2 \rightarrow 1}$
1	198,026 [144,370-255,306]	71,151 [48,114-95,972]	33.29 [25.26-44.64]	2.24 [1.48-2.97]	916 [648-966]	610 [485-717]
2	124,540 [100,500-157,630]	134 [92-173]	12.26 [8.07-16.55]	28.74 [21.26-36.24]	405 [258-557]	364 [230-513]
3	259,076 [199,440-322,951]	8,157 [6,868-9,505]	17.54 [13.34-22.64]	2.36 [1.81-2.97]	973 [916-984]	86 [46-127]
4	398,240 [331,190-470,597]	93,442 [67,534-111,936]	21.32[16.03-24.82]	1.97 [1.34-2.65]	809 [704-887]	493 [344-648]
5	138,450 [107,575-181,382]	2,612 [1,929-3,338]	17.31 [13.62-21.15]	3.81 [2.56-5.09]	847 [719-923]	425 [307-564]

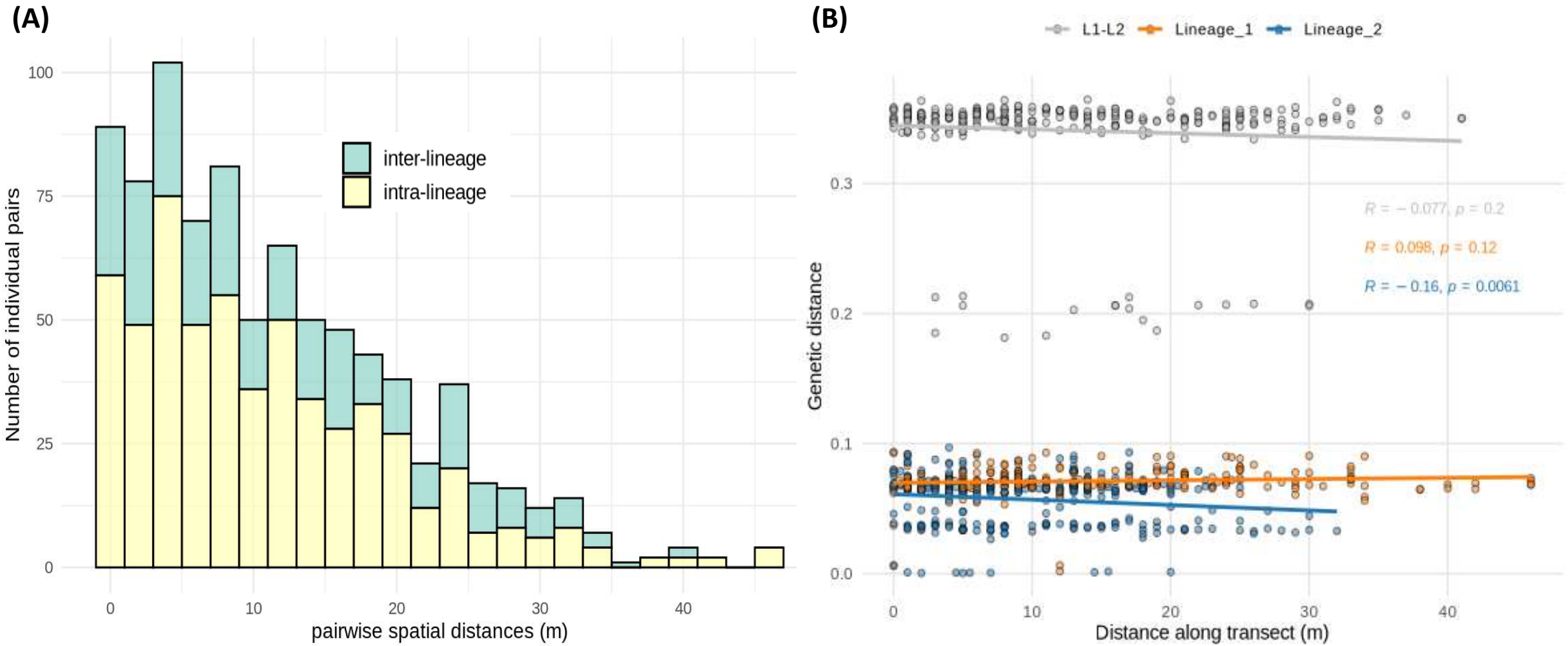
T_{split}: Time from the ancestry split (in number of generations), Tsc: Time from secondary contact (in number of generations), $M_1 \rightarrow_2$ or $M_2 \rightarrow_1$: introgression rate from lineage 1 to 2 or lineage 2 to 1; n_{Bar}: the number of barrier loci detected from lineage 1 to 2 or lineage 2 to 1.

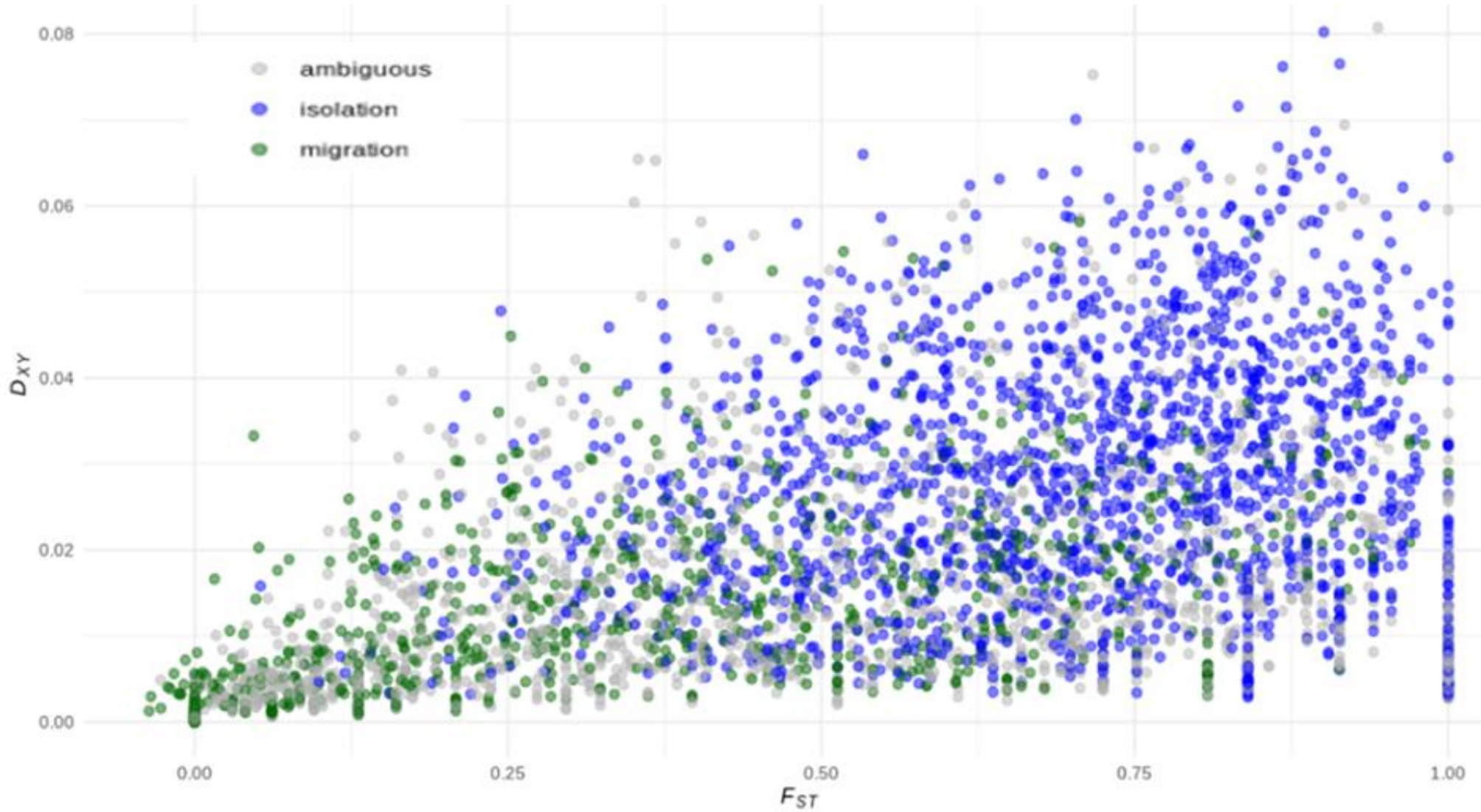

(A)
PC2 (5.51%)
PC1 (20.87%)
PC3 (3.05%)
PC1 (20.87%)
Q1
0.75
0.50
0.25
(B)
Ancestry Proportions

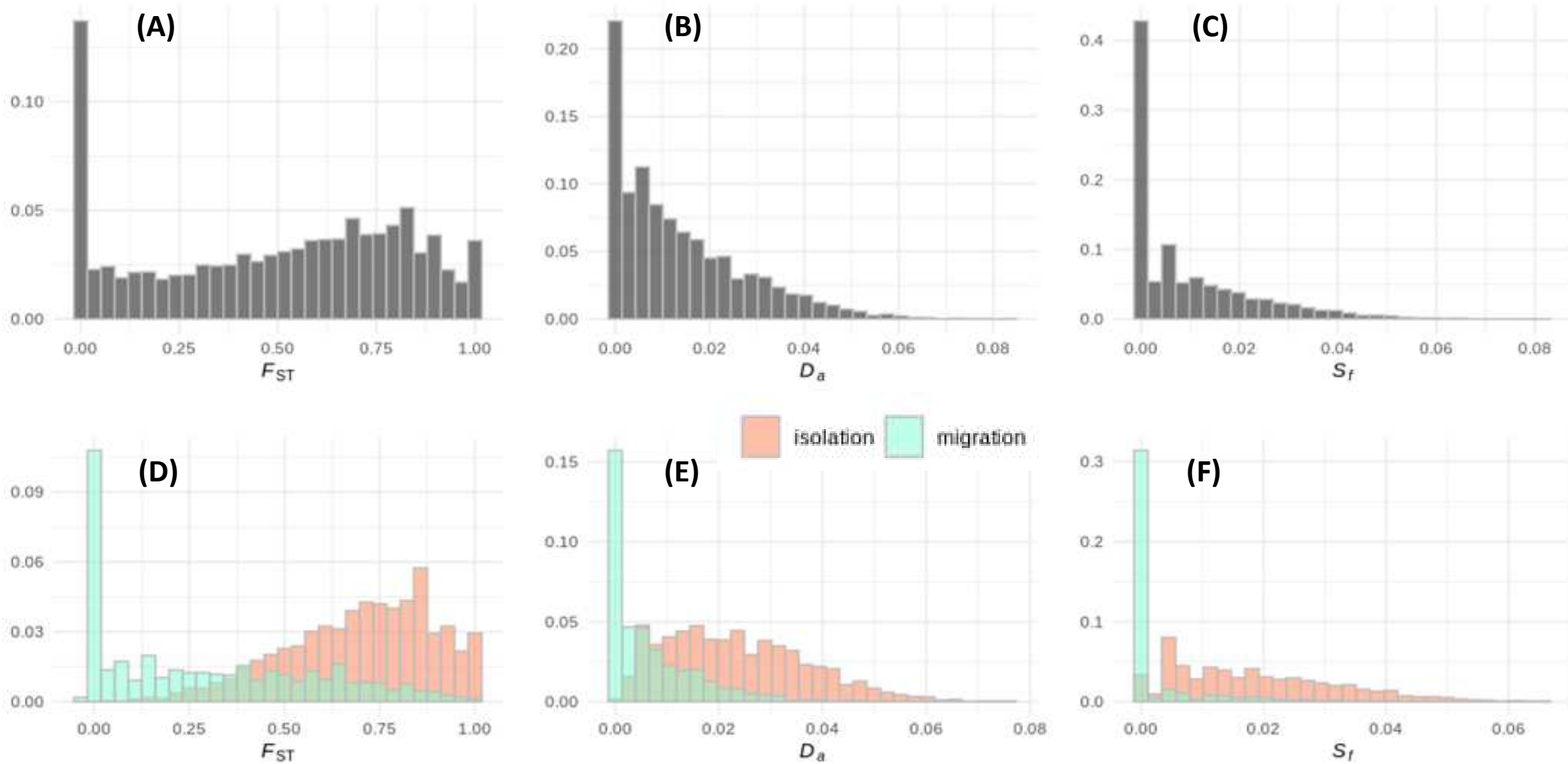

(A)
(B)
Interspecific Heterozygosity
Hybrid Index
1.00
0.75
0.50
0.25
0.00
B2POP1
B2POP2
B3POP1
B3POP2
F1HYB
F2HYB
Lineage 1
Lineage 2
Q1
0.75
0.50
0.25

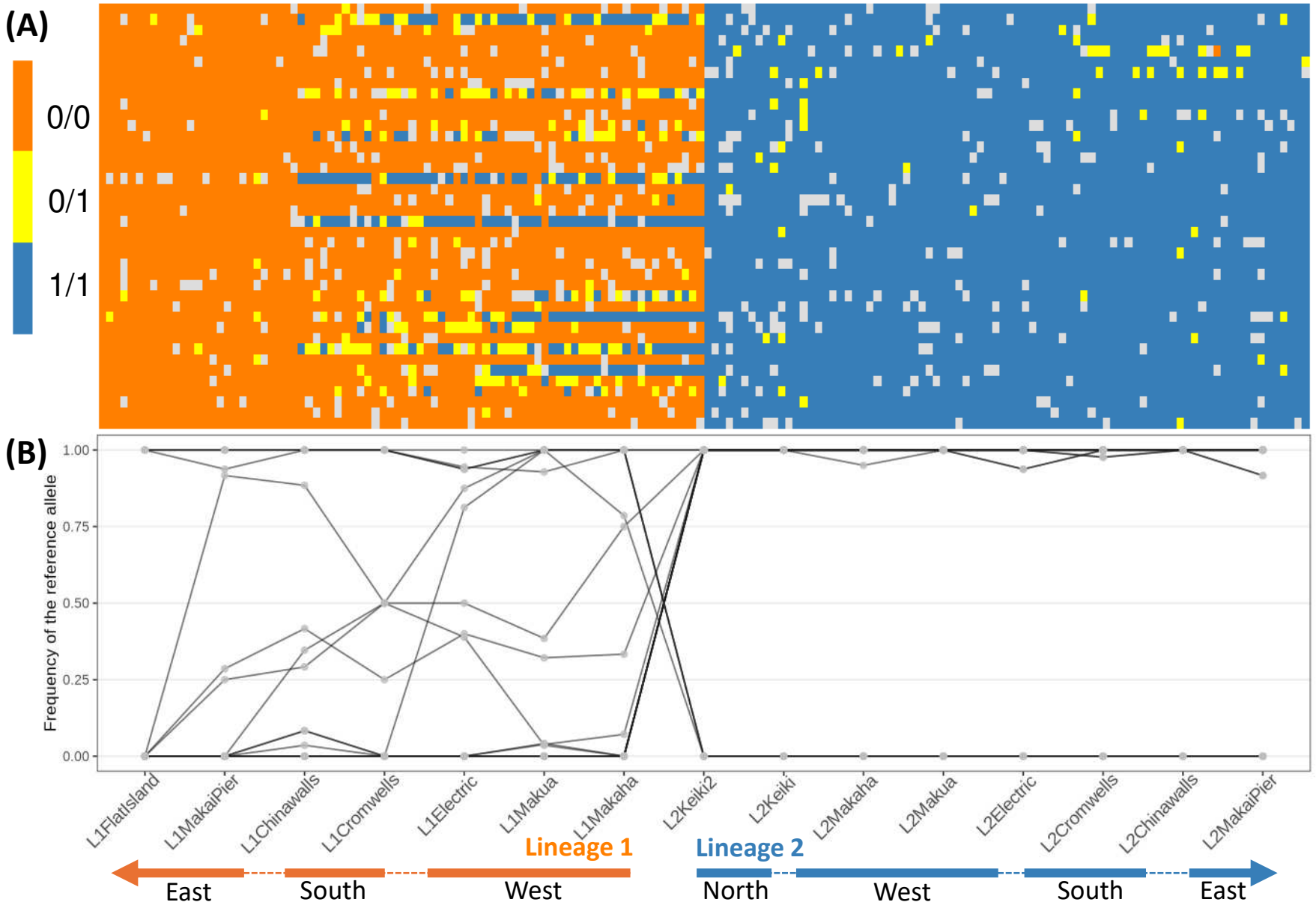

(A)
N
Keiki_2
Keiki
Makua
Makaha
Electric
Lineage 1
Lineage 2
Flat_Island
Makai_Pier
Cromwells
China_walls
20 km


(B)
L1-L2
Lineage-1
Lineage-2
F_{ST}
0.8
0.6
0.4
0.2
0.0
0
25
50
75
Least-cost distance (km)
R = 0.71, p = 1.1e-06
R = 0.67, p = 0.00083
R = −0.2, p = 0.11


(A)
100
75
50
25
0
Number of individual pairs
inter-lineage
intra-lineage
0
10
20
30
40
pairwise spatial distances (m)
(B)
L1-L2
Lineage_1
Lineage_2
0.3
0.2
0.1
0.0
Genetic distance
R = − 0.077, p = 0.2
R = 0.098, p = 0.12
R = − 0.16, p = 0.0061
0
10
20
30
40
Distance along transect (m)


ambiguous
isolation
migration
0.08
0.06
0.04
0.02
0.00
D_{XY}
0.00
0.25
0.50
0.75
1.00
F_{ST}


(A)
(B)
(C)
(D)
(E)
(F)
isolation
migration
F_{ST}
D_a
S_f


(A)
0/0
0/1
1/1
(B)
Frequency of the reference allele
1.00
0.75
0.50
0.25
0.00
L1FlatIsland
L1MakaiPier
L1Chinawalls
L1Cromwells
L1Electric
L1Makua
L1Makaha
L2Keiki2
L2Keiki
L2Makaha
L2Makua
L2Electric
L2Cromwells
L2Chinawalls
L2MakaiPier
Lineage 1
Lineage 2
East
South
West
North
West
South
East

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SUPPLEMENTARYMS.pdf
- Graphicalabstract.pdf