ASYMMETRIC GENE FLOW MAINTAINS RANGE EDGES IN A MARINE INVERTEBRATE

Zachary B. Hancock[*,1], Nicole E. Adams[1], Perry L. Wood Jr.[1], Gideon S. Bradburd[1]

[1]Department of Ecology & Evolutionary Biology, University of Michigan

[*]Corresponding author:
Zachary B. Hancock
Department of Ecology & Evolutionary Biology
University of Michigan
Ann Arbor, MI 48103
Email: hancockz@umich.edu

Running title: Asymmetric gene flow maintains range edges in a marine invertebrate

ABSTRACT

Species are restricted to finite ranges, and when they are in demographic equilibrium with the environment this implies a failure to adapt at the range periphery. There are several competing evolutionary hypotheses for the maintenance and persistence of range edges. One hypothesis is gene swamping of edge populations, in which migrants from the center carrying alleles adapted to the range center but maladaptive along the edge hamper local adaptation. A competing explanation is dispersal limitation, in which reduced migration to the periphery keeps edge populations small, causing drift to dominate over selection. It has proven difficult to assess these potential generative scenarios due to the inherent complexity of species ranges. In this study, we leverage range-wide genomic sequencing data to investigate the evolution of range edges in an empirical system - *Haustorius canadensis*, the Canadian shovel bug. We find increased genetic drift and inbreeding on the range edges and detect signatures of asymmetric migration towards the range center. The inferred directionality of migration is consistent with the direction of the dominant current regimes in the North Atlantic. We conclude that anisotropic dispersal towards the range center leaves progressively fewer migrants to replenish peripheral populations, consistent with the migration limitation hypothesis, and generates the appearance of an abundant center.

KEYWORDS: gene swamping, range evolution, directionality index, migration, coastal, shovel bug, isolation-by-distance

1 INTRODUCTION

All species are restricted to finite ranges. How range edges form and persist, why populations fail to adapt and continue to expand their range, and what the long-term stability of range size is, are all fundamental questions in the ecology and evolution of ranges (Kirkpatrick & Barton 1997; Sexton et al. 2009; Gaston 2003; Louthan et al. 2015; Angert et al. 2020). Answering these questions has major implications for urgent areas of research, such as predicting species' responses to climate change (Doak & William 2010) or the propensity of invasive species to rapidly expand their range (Alexander & Edwards 2010).

The simplest explanation for the persistence of a species' range limits is that they reflect a species' niche limits (Holt 2003; Hargraves et al. 2014); i.e., the species' distribution is determined by the environment (Angert et al. 2020). When range and niche limits coincide, it suggests a failure of peripheral populations to adapt to local conditions and expand both niche and range. There are several competing evolutionary hypotheses that seek to explain the persistence of range edges due to a failure to locally adapt. Haldane (1956) and Mayr (1963) were among the first to suggest that asymmetric gene flow from central populations to the edge populations might limit the latter's ability to adapt due to "gene swamping." Gene swamping is expected to occur if migrants from central populations carry with them alleles adapted to "central" conditions but, due to antagonistic pleiotropy, maladaptive at the range edge and conditions beyond. If migration rates from the center populations are high enough, any adaptive

alleles segregating in edge populations may become “swamped” by the influx of maladaptive alleles from migrants. Kirkpatrick & Barton (1997) formalized the gene swamping hypothesis by developing a quantitative genetic model across environmental clines of varying steepness, demonstrating that swamping could be a powerful mechanism maintaining species range edges. Gene swamping is a special case of the “migration load,” the reduction in fitness of a population due to the influx of migrant alleles maladapted to local conditions (Garcia-Ramos & Kirkpatrick 1997; Bolnick & Nosil 2007).

However, empirical evidence for the prevalence of gene swamping has been mixed. In a large review of the literature, Kottler et al. (2021) find little empirical evidence for gene swamping. Indeed, multiple studies indicate that increased gene flow to edge populations might have a positive effect (e.g., Sexton et al. 2011; Bontrager & Angert 2019). Range edges often have lower genetic diversity (Eckert et al. 2008; Pironon et al. 2016) and increased inbreeding depression (Arnaud-Haond et al. 2006; De Ryck et al. 2016) compared to core populations, both conditions that could be alleviated by increased migration. Therefore, Kottler et al. (2021) suggest that empirical evidence supports an alternative mechanism of range edge evolution: dispersal limitation (Sexton et al. 2009; Hargreaves et al. 2014). Under a dispersal limitation scenario, the lack of migrants to edge populations may be depriving them of the needed genetic variation to respond to strong selective pressures at the niche limit. As opposed to gene swamping, which predicts that increased migration from the center will hinder adaptation, the dispersal limitation hypothesis posits increased gene flow has the potential to rescue edge populations (Halbritter et al. 2015; Evans et al. 2014; Silva-

Arias et al. 2017). However, due to the inherent complexity of species ranges, identifying range limits (e.g., Dallas et al. 2017; Pironon et al. 2017; Soberón et al. 2018) and the underlying processes generating those limits remains challenging (Kottler et al. 2021; Fristoe et al. 2023), especially in organisms not amenable to transplant experiments. Thus, an ideal system for investigating questions of range evolution would have a relatively simple range with well-defined edges.

In the ocean, particularly for pelagic organisms, an important cause of dispersal limitation may be the direction and strength of the dominant current regimes (Siegal et al. 2003). The average dispersal distance of a marine pelagic larva, for example, is a function of its settling time and current velocity (Pringle & Wares 2007; Wares & Pringle 2008). White et al. (2010) found that incorporating larval duration and directional oceanic currents dramatically improved model predictions of population differentiation in the Kellet's whelk (*Kelletia kelletii*) over models relying on geographic distance alone. Similarly, Dambach et al. (2016), investigating a deep-sea shrimp (*Nematocarcinus lanceopes*), identified patterns of relatedness between populations consistent with current connectivity in the Southern Ocean. Thus, marine ranges can be determined by complex abiotic and biotic factors that can cause mass movement of individuals in asymmetric ways. For example, in the North Atlantic, a northward current, the Gulf Stream, emerges from the Loop Current in the Gulf of Mexico, bringing warm water from the Caribbean, and travels up the coastline to Chesapeake Bay before diverting to the mid-Atlantic (**Fig. 1A**). A second regime, the Labrador Current, travels southward along Nova Scotia towards New England, bringing with it cold Arctic water, before also diverting eastward. Hence, along the North Atlantic coastline, two currents converge

along the central latitude from opposite directions. For a marine organism distributed along the entire coastline and whose dispersal is largely dictated by the currents, this could generate a migration bias towards the center of the range. If the asymmetry in dispersal is strong enough, it might even suggest that range limits are not reflective of niche limits, but instead the lack of available migrants for successive colonization and further range expansion into otherwise habitable patches (Hargraves et al. 2014; Oldfather et al. 2019).

Small amphipod crustaceans in the family Haustoriidae, which inhabit the exposed sandy beaches of the western North Atlantic, are an excellent system in which to test questions about range evolution. Due to their life history (see *Methods*), haustoriid amphipods are expected to be relatively poor dispersers, and this has been largely confirmed in previous studies (Jones 2005; Hancock et al. 2019). In the western North Atlantic, *Haustorius canadensis* (Bousfield 1962) can be found on beaches from central Florida to southern Maine with some relict, isolated populations reported in the Gulf of Saint Lawrence (Bousfield 1962; 1965). Haustoriid amphipods have no prior common name; here we introduce the term “shovel bugs” (with reference to their modified shovel limbs) to describe haustoriids, and “Canadian shovel bug” to refer to *H. canadensis*. A curious feature of their southernmost distribution is the abundance of suitable sandy beaches farther south than they are found, with no Canadian shovel bug recorded south of Melbourne, Florida. The northern range edge, alternatively, might represent a hard environmental transition, as most of coastal Maine is composed of rocky outcrops with few fine sand beaches.

Given their relatively constrained niche (the intertidal of wave-exposed sandy beaches), the ease of identifying range boundaries, and the one-dimensional nature of their range, the Canadian shovel bug is an ideal organism for evaluating hypotheses of the evolutionary mechanisms determining range limits. If gene swamping is maintaining range edges in this species, we should detect a signature of asymmetric gene flow from the range center towards the edge. Furthermore, we expect edge populations to harbor the genetic signature of population sinks – lower genetic diversity relative to more central populations, and the genetic diversity they do possess should be a subset of that harbored in the center. Alternatively, range edges might be maintained by a failure to colonize habitable patches due to asymmetries in the dispersal medium. Because shovel bugs passively disperse via the currents, successively more distant patches in the opposite direction of the current flow may have fewer available migrants to colonize them, ultimately determining range limits. If dispersal limitation due to asymmetries in current regimes are driving the persistence of range edges, we expect to find patterns of asymmetric migration towards the range center. In addition, we would expect to identify population-clustering associated with the two major currents, and a breakdown of this clustering pattern at central latitudes where the currents converge.

In this study, we first characterize the population genetic structure of the Canadian shovel bug across its entire range in the North Atlantic. Next, we estimate genetic diversity and inbreeding across the range to test for patterns of reduced abundance along the range edges. Finally, we determine whether patterns of migration and colonization are consistent with either gene swamping or dispersal limitation as processes maintaining range edges. We find evidence for directional colonization from

the periphery of the range toward the center, which we hypothesize is due to current regimes in the North Atlantic and consistent with the migration limitation hypothesis.

2 METHODS

2.1 Natural history of shovel bugs

Deemed “the most interesting group of amphipods,” shovel bugs are haustoriid amphipods that have highly derived morphological adaptations, including expanded articles on the 6th and 7th pereopods, extensive setal elaboration, and spiny, cup-like structures on their 3rd and 4th pereopods that they utilize as shovels (**Fig. 1A**) (Barnard 1969; Hancock & Wicksten 2018).

Shovel bugs are blind and entirely fossorial, living in the first 1–10 cm of sand and almost exclusively occupying the narrow swash zone (Croker & Hatfield 1980). They lack a pelagic larva; instead, females brood juveniles through the first few molts in their marsupium. Adult shovel bugs range in body size from 2–18 mm, sometimes within the same species, and past work has found strong correlations between body size and latitude (Hancock et al. 2021), though there may also be an effect of sand grain size (LeCroy et al. 2002; Hancock & Wicksten 2018). Several species may occupy the same beach, but there is evidence that they partition the beach by sex and by species to limit competition for space (Croker & Hatfield 1980).

The relationship of shovel bugs to the broader Amphipoda has been shrouded in mystery, with some hypothesizing a close kinship with other fossorial amphipods (Lowry & Myers 2017). However, recent molecular phylogenies have cast doubt on the taxonomic affinities between the fossorial amphipods, with multiple studies indicating

181 shovel bugs are gammarids and that modifications to the fossorial lifestyle found across
182 several families arose convergently (Verheye et al. 2016; Copilas-Ciocianu et al. 2020;
183 Hancock et al. 2022).
184
185 *2.2 Field collecting*
186 Shovel bugs were collected from 18 sites from Cocoa Beach, Florida, USA to Crescent
187 Beach, Maine, USA in May of 2023 (**Fig. 1A**; *Table S1*). Collections began in West
188 Palm Beach, Florida, USA to identify the southernmost established population; traveling
189 north from West Palm Beach, no specimens were located until Cocoa Beach, Florida.
190 Similarly, sampling continued beyond Crescent Beach, Maine, but no samples were
191 located further north (last attempted location was Popham Beach, ~93 km north of
192 Crescent Beach). Specimens were sifted from sand using a 435 μm sieve plate, plucked
193 from the plate with forceps, and preserved in 95% ethanol. Individuals were identified to
194 species in the lab using a dissecting microscope and key from Hancock & Wicksten
195 (2018), and then stored at -20°C prior to DNA extraction.
196
197 *2.3 Molecular techniques*
198 Whole genomic DNA was extracted from either the whole individual or from pereopods
199 4–7 using a Qiagen DNeasy Blood & Tissue Kit following manufacturer's protocols.
200 Library prep was performed using Restriction site Associated DNA sequencing (2RAD)
201 following Bayona-Vásquez et al. (2019), with slight modifications. All 96 samples were
202 quantified using a Qubit™ 1X dsDNA High Sensitivity (HS) kit on a Qubit 2.0
203 fluorometer. Samples were then normalized to have 100ng per sample (5 μL) on a 96

204 well plate, and then digested in a 15 μL solution with 1.5 μL 10X CutSmart Buffer (New
205 England Biolabs, USA), with 10 units of EcoRI (0.5 μL), 10 units MseI (0.5 μL), 3.5 μL
206 Invitrogen™ UltraPure™ DNase/RNase-Free Distilled Water, 2.5 μL of EcoRI
207 compatible adapter (2 μL), and 0.25 μM MseI compatible adapter (μL) on a Eppendorf
208 Mastercycler PRO S 6325 Thermal Cycler thermocycler at 37°C for one hour with the
209 heated lid turned off. Immediately after digestion we added 5 μL of ligation mastermix
210 composed of 2.75 μL of nuclease free water, 0.5 μL 10X ligation buffer (NEB), 1.5 μL 10
211 mM rATP (Promega), and 0.25 μL T4 DNA Ligase 400,000 U/ml to each sample. The
212 ligation reaction was implemented on the same thermocycler as above for two cycles of
213 22°C for 20 min, 37°C for 10 min, followed by 80°C for 20 min to stop the enzyme
214 activity. We then pulled 5 μL of each sample and split these into two microcenterfuge
215 tubes and performed a 1.25X SpeedBead cleanup (Rohland and Reich, 2012).
216 To reduce PCR duplicates and add the iTru5–8N barcode (which has a segment
217 of eight random base pairs) we implemented a single PCR cycle with six tubes each
218 with a 50 μL reaction. Each reaction contained 10 μL nuclease free water, 25 μL KAPA
219 HiFi hot start ready mix (Roche, USA), 5 μL iTru5–8N barcode (5 μM), and 10 μL of
220 library and were added using a single cycle of 98˚C for 1 min, 60˚C for 30s, and 72˚C
221 for 6 min. All six PCR reactions were then pooled and a 1.5X SpeedBead cleanup was
222 performed and resuspended in 30 μL of nuclease free water. Next, we split the library
223 into three 55 μL reactions with three unique iTru7 adapters. Each reaction had 10 μL of
224 nuclease free water, 25 μL KAPA HiFi hot start ready mix (Roche, USA), 5 μL P5 primer
225 (5 μM) and 5 μL of each iTru7 (5 μM) added separately. The PCR was implemented on
226 the thermocycler mentioned above under the following conditions: 95˚C for 3 mins, 8

227 cycles of 98˚C 20s, 60˚ 15s, 72˚C 30s, and a final 72˚C for 5 mins. All PCR reactions
228 were pooled and then cleaned using a 1.5X SpeedBead cleanup and resuspended in 30
229 μL of nuclease free water.
230 We performed size selection on the pooled library using a Pippin Prep (Sage
231 Science, USA) with internal standards on a 1.5% dye free cassette, target 350–480 bp
232 range (mean= 415 bp). After size selection to enrich the library, we performed a PCR
233 under the same conditions for the iTru7 reaction above, except with only four 50 μL
234 reactions and 6 cycles, using both the P5 and P7 primers (5 μM). Finally, we performed
235 a 1X SpeedBead clean up and quantified the library on the Qubit 2.0 to ensure that our
236 library had a high enough concentration for sequencing. The library was then
237 sequenced on 12.5% percent of a NovaSeq X lane at the University of Michigan
238 Advanced Genomics Core.
239 Raw sequence reads were assembled and filtered using the de novo pipeline in
240 STACKS v.2.67 (Catchen et al. 2013; Rochette et al. 2019). A custom *process_radtags*
241 function was provided to us by J. Catchen to handle the unique FASTQ files of 2RAD,
242 which include an additional adapter sequence. Reads were assembled using the de
243 novo pipeline (Catchen et al. 2011); briefly, stacks were built in *ustacks* with a maximum
244 nucleotide distance allowed of 2 and minimum number of reads needed to seed a stack
245 of 3. Stacks were merged allowing a maximum of 2 gaps and 6 mismatches. Next, we
246 used the *populations* function, filtering out alleles that did not occur in at least 10
247 populations and keeping those that were present in 75% of individuals per population.
248 Lastly, we randomly selected one SNP per RAD locus (--write-random-snp).
249 Sequences were output as two VCF files, one that included only variant sites and a

second that contained both variant and monomorphic sites. VCF files were additionally filtered using *vcftools* v.1.16 (Danecek et al. 2011); first, individuals with >80% missing data were removed. Next, we filtered out loci with more than two alleles (--max-alleles 2) and with a minor allele frequency of less than 0.05 (--maf 0.05). We then filtered out loci with a minimum mean depth of less than 20 (--min-meanDP 20) and removed all indels (--remove-indels). After filtering, we evaluated the relationship between mean depth, estimated heterozygosity, and distance from the range center via linear regression to determine if mean depth might be a confounding factor.

2.4 Population genetic analysis

To evaluate potential patterns in genetic diversity, we estimated individual-level heterozygosity using *pixy* (Korunes & Samuk 2021) on the VCF that included invariant sites. The inbreeding coefficient (*F*; Wright, 1931), estimated as the difference between the expected and observed homozygosity per individual, was estimated using *vcftools* (--het).

To test for geographic structure, we first used *pixy* to estimate population-level pairwise F_{ST} (Weir & Cockerham 1984), which we then linearized and regressed against geographic distance to test for isolation-by-distance (Wright 1943; Rousset 1997). Next, we performed a principal component analysis (PCA) to identify signs of genetic clustering using *VCF2PCACluster* (He et al. 2024). We then performed two cluster analyses. First, we used *fastSTRUCTURE* (Raj et al. 2014) with *k* (the number of discrete, ancestral clusters) from 2–5. Next, to account for continuous spatial differentiation, we ran *conStruct* (Bradburd et al. 2018), varying *k* from 1–4. For each

273 value of *k*, we performed 5 independent MCMC chains each with 2000 iterations. We
274 then reported results from the chain with the highest mean posterior probability. As an
275 additional test of whether identified clusters and admixed individuals were the result of
276 increasing gene flow towards the range center or continuous isolation-by-distance, we
277 also performed a triangle plot analysis using *triangulaR* (Wiens & Colella 2024a). In the
278 presence of directional gene flow, individuals with a hybrid index approaching 0.5
279 should have concomitantly high interclass heterozygosity. However, a continuous
280 population characterized solely by isolation-by-distance will display a high hybrid index
281 but no increase in interclass heterozygosity (Wiens & Colella 2024b).

282 We evaluated potential barriers to migration using the program *FEEMS* (Marcus
283 et al. 2021), which estimates an effective migration surface using edge-specific
284 parameter estimation in a pseudo-likelihood framework. Lastly, we tested for a signature
285 of range expansion along the periphery of the population using Peter & Slatkin's (2013)
286 directionality index, ψ, as implemented in *rangeExpansion*
287 (https://github.com/BenjaminPeter/rangeexpansion). The directionality index is
288 calculated by comparing the difference in allele frequencies between pairs of
289 populations, conditioned on the allele being present in both populations. Large source
290 populations typically contain many rare alleles. When a new population is founded from
291 this larger source population, these same alleles, if they have not been lost due to drift,
292 are likely to drift to higher frequencies in the new population. Thus, successively more
293 recently founded populations in an expanding range will have increasingly positive
294 values of ψ, permitting the identification of both the origin and the direction of the range
295 expansion (Peter & Slatkin 2013; 2015). Following Peter & Slatkin (2015), the

directionality of range expansion is inferred by summing across the columns of the matrix of pairwise estimates of ψ for each population (by row), where the value of ψ is the strength of the signal and the sign is the direction. In populations at migration-drift equilibrium in continuous space, simulations show that ψ should be zero (Peter & Slatkin 2013). Therefore, if ψ is zero, there is no marked departure from migration-drift equilibrium and hence no signature of reduced fitness or recurrent colonization in edge populations. However, if gene swamping is driving edge dynamics, we expect ψ should be positive in the direction of the edge. Alternatively, if dispersal limitation, driven by directional currents, is the primary cause of range limits, then ψ should be negative at the edges and increasingly positive toward the range center.

2.5 Simulations

Because the expansion statistic ψ is influenced by relative rates of drift, the inferred direction of population expansion, as given by ψ, may be influenced by spatial clines in population density even under scenarios with isotropic dispersal. To determine whether the results of our ψ analyses are robust to spatial variation in population density (see Results/Discussion), we conducted forward-time simulations using SLiM v.4.0.1 (Haller & Messer 2023). We established a 10-deme stepping-stone model with four different range scenarios: 1) equal population size per deme with symmetrical migration; 2) equal population size per deme with asymmetrical migration from the edge populations towards the center; 3) an abundant center model in which the central demes had a fivefold higher census population size than the edge, with symmetric migration; and 4) abundant center with asymmetric migration from the edges inward. For the asymmetric

319 migration scenarios, migration rates from the edge towards the center were twice that of
320 center to edge, and for scenarios with an abundant population center, population size
321 decayed linearly towards the edge. In addition, we performed simulations with
322 increasing differences in population sizes between the central and edge populations,
323 with equal migration, to evaluate how these relative differences in genetic drift might
324 drive patterns in ψ. We simulated ratios of population size between the center and edge
325 of 5:1, 25:1, and 50:1 and performed 10 replicates for each scenario. For each ratio,
326 population size decreased linearly, but more precipitously at higher ratios to balance
327 maintaining constant deme number and computational efficiency. Individuals were
328 diploid and hermaphroditic, and we simulated a 10 Mb chromosome with a
329 recombination rate of 1×10^{-8} per base pair. Simulations were run for 100,000 ticks
330 (>10*N*) to reach demographic equilibrium, and tree sequences were recorded for
331 subsequent processing using *pyslim* v.1.0.1 (Kelleher et al. 2018) and *tskit* v.0.5.3
332 (Kelleher et al. 2018). For tree sequences with more than one root, we used recapitation
333 to simulate a neutral coalescent history until only a single root remained; mutations
334 were added at a rate of 1×10^{-7} per base-pair to the trees using *msprime* v.1.2.0
335 (Kelleher et al. 2016). Next, we randomly sampled five individuals per deme and output
336 their sequences in VCF format. From here, we trimmed the VCFs to only include
337 biallelic SNPs, and converted them to BED files using *plink* v.1.9 (Purcell et al. 2007).
338 To use the *rangeExpansion* package, we converted the individual coordinates from
339 SLiM, which were in discrete demes, into Earth coordinates as required by
340 *rangeExpansion*; thus, we chose 10 coordinates 5° apart in latitude. Because the units
341 of distance are arbitrary in SLiM, we needed only to ensure that deme 1 is closest to

342 deme 2, etc., to adequately capture the SLiM environment in an Earth coordinate
343 system. We then calculated ψ for each simulation to see how the different range
344 scenarios might impact inferences made from ψ.
345
346 **3 RESULTS**
347
348 *3.1 ddRAD assembly*
349 A total of 2,605,149 loci passed filtering in STACKS; the mean effective per-sample
350 coverage was 23.0 with a standard deviation of 13.6, a minimum of 4.0, and a maximum
351 of 67.4. The mean number of sites (bp) per locus was 220.1. After filtering for biallelic
352 SNPs and missingness in *vcftools*, the final dataset included 7,405 SNPs. We found a
353 significant positive relationship between mean depth per individual and our estimates of
354 heterozygosity from *pixy* ($P < 5.3e–7$, $R^2 = 0.27$, $\beta = 8.77e–6$; *Table S2*); in addition, we
355 found a significant negative relationship between mean depth and the distance from the
356 center of the range (taken as 36° latitude; $P < 0.002$, $R^2 = 0.10$, $\beta = –0.051$) (see *Table*
357 *S2*). To account for the potential confounding between measures of heterozygosity,
358 mean depth, and their relationship with the distance from the range center, all
359 subsequent multivariate linear regressions were performed with mean depth as an
360 additional predictor. To determine if a few individuals with high mean read depth were
361 driving the observed patterns, we re-tested for a significant correlation between depth
362 and heterozygosity after dropping all individuals with a mean depth greater than 30.
363 While a significant correlation between heterozygosity and mean depth remained ($P <$

364 2.7e–5; R^2 = 0.38, β = 3.1e–5), it was no longer significantly correlated with distance

365 from the center (P = 0.57; R^2 = –0.02, β = –2.03) (*Fig. S1*).

366

367 *3.2 Genetic structure*

368 The PCA revealed relatively continuous structure, with a possible break between

369 populations north and south of the Chesapeake (**Fig. 1B**), but there were no obvious

370 discrete breaks identified in patterns of isolation-by-distance when regressing linearized

371 F_{ST} against geographic distance (**Fig. 1C**). Results of the *fastSTRUCTURE* analysis

372 supported three geographically restricted population clusters (k = 3) as the model that

373 maximized the marginal likelihood – one from Florida to South Carolina, a second from

374 North Carolina to Virginia, and a third that included the remaining sample sites from

375 northern Virginia to Maine (*Fig. S2*). However, *conStruct*, which allows for continuous

376 structure to be accommodated within discrete populations, supported k = 2. Identified

377 clusters corresponded to the dominant current regimes in the Atlantic with increasing

378 admixture towards the center of the range (**Fig. 2**). This result was echoed by the

379 triangle plot analysis of the *conStruct* results, which found increasingly elevated

380 interclass heterozygosity at central latitudes (*Fig. S3*).

381

382 *3.3 Patterns of diversity*

383 After incorporating sequencing depth as a possible confounding effect, measures of

384 individual heterozygosity were significantly predicted by the distance from the center of

385 the range (P < 3.79e–16, β = –1.5e–6; mean depth P = 5.1e–5, β = 5.38e–6). This

386 relationship remained significant after dropping individuals with mean depth greater than

387 30 ($P < 0.0009$; $R^2 = 0.25$, $\beta = -1.2e-6$) (*Fig. S4*). Similarly, measures of inbreeding
388 were the highest at the range edges, declining as a function of distance from the center
389 (**Fig. 3B**). These patterns are consistent with field observations: the ease of collecting
390 and observed highest abundance was in the central range locations and shovel bugs at
391 the edge sample locations were rarer and thus more difficult to collect. However, we
392 note that while distance is a much stronger predictor of patterns of genetic diversity and
393 inbreeding, mean sequencing depth was still a significant predictor in each model
394 (*Table S2*).
395 The effective migration surface estimated from *FEEMS* suggested elevated
396 levels of dispersal from northern Florida to Virginia along the path of the Gulf Stream
397 and reduced effective migration towards the edges (**Fig. 2B**). In accordance with results
398 from the PCA, the migration surface indicated a strong depression in effective migration
399 across the Chesapeake, which suggests it might act as a partial barrier to gene flow.
400
401 *3.4 Directionality index & simulations*
402 The inferred origin of the range expansion of the entire sample of individuals, as
403 indicated by the lowest value of ψ, was 29.9°N in northern Florida (**Fig. 4A**). When the
404 data were subset into two regions, South and North (based on the *conStruct* analysis),
405 the inferred origins were at 29.9° and 43.6° latitude, respectively. The sum of ψ for each
406 population resulted in a pattern of increasing ψ towards the range center, mirroring the
407 increase in heterozygosity (**Fig. 4A**). This pattern of ψ superficially suggests a range
408 expansion from the edges toward the center of the range that follows the dominant
409 current regimes of the Gulf Stream running south to north and the Labrador Current

410 running north to south, both converging in the center of the range (**Fig. 1A, 4A**). This
411 pattern of expansion is in opposition to expectations from the gene swamping
412 hypothesis and suggests that asymmetric migration due to the currents might be
413 maintaining range edges, consistent with the migration limitation hypothesis.

414 Simulation results from SLiM supported this interpretation. Across replicates, the
415 strongest divergences from ψ = 0, which is expected at migration-drift equilibrium,
416 occurred in simulations with asymmetric migration from the edges toward the center,
417 with the most striking inflation in the scenario in which there is both an abundant center
418 and asymmetric migration (**Fig. 4B**). While there was some deviation from zero along
419 the edges in scenarios with symmetric migration in both the equal and center-abundant
420 scenarios, they did not systematically produce a positive ψ in the center as in the case
421 of asymmetric migration (**Fig. 4B**). Furthermore, increasing the population size
422 discrepancy between the center and the edge did not consistently lead to an increase in
423 central estimates of ψ (*Fig. S5*), indicating that asymmetric migration, not increased drift
424 along the edges, is the dominant force shaping the observed dynamics.

425

426 **4 DISCUSSION**

427 The persistence of range edges despite the potential for local adaptation to marginal
428 conditions and resulting continued expansion is an enduring evolutionary mystery.
429 Several hypotheses have been proposed to explain the finite nature of species ranges,
430 principle among these gene swamping (Mayr 1963; Kirkpatrick & Barton 1996) and
431 dispersal limitation (Sexton et al. 2009; Hargraves et al. 2014). In this study, we
432 investigated the suitability of these hypotheses to explain range evolution of the

433 Canadian shovel bug (*Haustorius canadensis*; Bousfield 1962), which is distributed
434 along the Eastern Seaboard of North America.
435 Past work suggests that shovel bugs have the potential for large amounts of
436 cryptic diversity and are sensitive to environmental barriers (Jones 2005; Hancock et al.
437 2019). Thus, our first aim was to characterize the population genetic structure of the
438 Canadian shovel bug. While we found that patterns of linearized F_{ST} were consistent
439 with a single continuous population, the PCA and *conStruct* results suggested the
440 presence of two populations associated with each current regime. In addition, we found
441 evidence for resistance to migration across the Chesapeake Bay. Shovel bugs of the
442 genus *Haustorius* specialize on open, wave-exposed beaches in which the sand is
443 highly oxygenated (Bousfield 1970), and rarely occur in low salinity, brackish bays.
444 While the Chesapeake Bay was formed by an impact crater during the Eocene (Ivany et
445 al. 2000), and this event has been postulated to have contributed to haustoriid
446 diversification in North America (Hancock et al. 2022), its recent form as a large bay is
447 more recent (around the end of the Last Glacial Maximum ~10,000 years ago (Colman
448 & Mixon 1988)). Thus, the bay has the potential to be acting as a partial barrier to
449 dispersal and gene flow. However, because the isolation-by-distance regression did not
450 support a discrete break (**Fig. 1C**; *Fig. S3*), the barrier may be weak or has occurred
451 recently enough that it has not yet influenced genome-wide patterns of allele frequency
452 divergence, which can lag behind other population summary statistics (Barton et al.
453 2013; Ringbauer et al. 2018).
454 Once the genetic structure of shovel bugs was characterized, our next aim was
455 to determine if range edges have reduced genetic diversity and increased inbreeding

relative to central populations. Past theoretical (Polechová & Barton 2015) and empirical (Eckert et al. 2008; Pironon et al. 2016) work suggest that genetic diversity generally decreases as a function of distance from the range center when the environment changes gradually across space. Consistent with this past work, we found genetic diversity and inbreeding strongly correlated with distance from the range center (**Fig. 3**); edge populations harbor the lowest genetic diversity and highest rates of inbreeding. This pattern is broadly consistent with the “abundant center hypothesis” (Grinnell 1922; Hengeveld & Haeck 1982; Brown 1984), which predicts that population abundance is highest in the core of the range and decreases monotonically towards the edge. However, the generality of the abundant center hypothesis remains hotly contested, as there is little empirical evidence that population abundance consistently declines from the core to the edge (Dallas et al. 2020; Santini et al. 2019; Sagarin et al. 2006). While we observed higher densities of shovel bugs in the central sample sites, anecdotally consistent with the abundant center hypothesis, future work would benefit from quantitative estimates of census abundance across the Canadian shovel bug’s range.

Reduced genetic diversity and elevated rates of inbreeding in peripheral populations might indicate a reduced adaptive capacity (Angert et al. 2020; Kottler et al. 2023). During a range expansion, this can be driven by allele “surfing,” which describes the process by which genetic drift is increased along an expanding front due to reduced population densities, allowing alleles to reach high frequencies (Excoffier & Ray 2008). When a range size is stable, reduced adaptive capacity might result from gene swamping or dispersal limitation (Angert et al. 2020); both are expected to reduce population density on the range periphery. If range limits are being determined by gene

479 swamping, we expected to infer asymmetric migration from the center toward the edge,
480 with edge populations harboring a subset of the genetic diversity of central populations.
481 Furthermore, we predicted that the shared alleles with central populations would be at
482 higher frequency on the edge, consistent with repeated founder events at increasingly
483 peripheral populations. Conversely, if range limits are being determined by dispersal
484 limitation, we expected to infer a net direction of movement towards the range center
485 from the edges, which might be reducing dispersal to peripheral populations. This is due
486 to the dominant current regimes in the North Atlantic flowing in opposite directions,
487 converging at a central latitude in the Chesapeake Bay region (**Fig. 1A**). In addition, if
488 dispersal limitation is driving the persistence of range edges, we expected increasing
489 admixture proportions towards the center of the range where the Gulf Stream and
490 Labrador Current converge.

491 Our results strongly support the finding that dispersal limitation is the dominant
492 mechanism shaping the persistence of range edges in shovel bugs. Our spatial
493 clustering analysis indicated there are two populations of the Canadian shovel bug
494 associated with each of the current regimes in the North Atlantic, with admixture
495 proportions increasing toward the range center (**Fig. 2**). We also identified reduced
496 effective migration on the range periphery (**Fig. 2B**). Interestingly, we observed that
497 effective migration is generally higher south of the Chesapeake than north, which is
498 consistent with the greater velocity of the Gulf Stream relative to the much weaker
499 Labrador Current (**Fig. 1**). Furthermore, we detected a linear increase in ψ towards the
500 range center from each edge. The standard interpretation of this index would imply a
501 range expansion from the periphery towards the center. However, this is contradictory

502 to patterns of inbreeding and genetic diversity, which clearly indicated that drift is likely
503 more dominant on the range edge (**Fig. 3**). A possible explanation for these conflicting
504 patterns is that the ψ index is influenced by asymmetric dispersal, and, if so, the
505 observed directionality would support the dispersal limitation hypothesis.

506 To evaluate whether inferences made on the basis of the ψ index could be
507 affected by heterogeneity in density as well as migration, we conducted forward-time,
508 individual-based simulations intended to tease apart the impacts of differences in
509 population size and migration (**Fig. 3B**; *Fig. 5*). Consistent with our explanation above,
510 we found that ψ was highly sensitive to asymmetric migration and would infer a
511 population expansion in the direction of the range center even in equilibrium (i.e., non-
512 expansion) scenarios. In addition, this bias was strongest when under the combined
513 influence of asymmetric migration and abundant center dynamics (**Fig. 3B**), but,
514 importantly, did not occur in the presence of differences in population size alone, at
515 least not at the ratios we tested (*Fig. S5*). This suggests that asymmetric migration may
516 be playing the dominant role in shaping patterns in the SFS that are being identified by
517 the ψ index.

518 The ψ index is a popular statistic for testing for range expansions and has been
519 applied to a diverse array of different organisms (Manuel et al. 2016; Zhan et al. 2014;
520 Fifer et al. 2022). In this study, we uncovered a potential confounding influence on ψ
521 that is likely to impact any organism with high asymmetric migration. While asymmetric
522 gene flow has been identified in terrestrial organisms (Fedorka et al. 2012; Oswald et al.
523 2017), the taxa most likely to display directional dispersal are aquatic organisms with
524 passive dispersal via river flow or oceanic currents (Siegal et al. 2003; Pringle & Wares

525 2007; Morrissey & de Kerckhove 2009). Past work has shown that directional dispersal
526 in the ocean can mislead many population genetic analyses, including the inference of
527 random dispersal via patterns of isolation-by-distance and the probability of
528 coalescence (Nagylaki 1977; Lenormand 2002; Paz-Vinas et al. 2013). Using a
529 simulation approach, we uncovered that asymmetric migration can also bias the
530 detection of a range expansion, potentially generating a false positive signature when,
531 in fact, the population is in demographic equilibrium.

532 In addition to the impact asymmetric migration has on the inference of a range
533 expansion, certain biased dispersal patterns may themselves be capable of generating
534 an abundant center pattern. Specifically, asymmetric migration from the edges towards
535 the range center can maintain high abundance in the center, causing it to act like a
536 population sink, while colonization towards increasingly distant patches becomes less
537 likely as the number of available migrants dispersing away from the center becomes
538 increasingly small. Our findings in this system suggest that this may be an
539 underappreciated form of migration limitation (Sexton et al. 2009; Hargraves et al.
540 2013), in which the rate of extinction in peripheral patches exceeds that of colonization;
541 the cause of this decreased migration is asymmetries in the probability of migration due
542 to the dispersal medium itself.

543 The evolutionary processes that maintain range edges, especially in systems in
544 which the environmental change is clinal, are still subjects of active debate (Kottler et al.
545 2023). We uncovered strong support for the hypothesis that range edges in this system
546 are maintained by dispersal limitation, itself driven by a combination of passive dispersal
547 and directional currents reducing the number of available migrants to colonize

548 peripheral patches. We suggest this mechanism for generating an apparent abundant
549 center might be widespread in coastal marine taxa, especially those with passive
550 dispersal, as oceanic currents regularly converge within species ranges. Furthermore,
551 focusing on asymmetric migration via directional water or wind currents may help clarify
552 what maintains range edges in cases where the fundamental niche appears to exceed
553 the range limit.
554

555 **DATA AVAILABILITY**

556 Bioinformatic and analysis pipelines, including filtered VCF data objects, are available at
557 https://github.com/zachbhancock/asym_gene_flow_proj. Raw sequences were
558 deposited to GENBANK (ascension PRJNA1183394).
559

560 **AUTHOR CONTRIBUTIONS**

561 ZBH and GSB conceptualized and designed the study. ZBH carried out the field
562 collections. ZBH and PLW performed the laboratory work. ZBH and NEA carried out the
563 bioinformatics and analyzed the data. All authors wrote and edited the manuscript.
564

565 **ACKNOWLEDGEMENTS**

566 Thanks to members of the Bradburd lab who provided helpful comments on this
567 manuscript. Thanks also to Kristen Wacker, Teresa Pegan, and Benjamin Peter for
568 thoughtful discussion and feedback. We are indebted to Leonard Jones who coined the
569 common name “shovel bug” for haustoriid amphipods, and proudly use it in print for the
570 first time here. May it catch on. Lastly, special thanks to Daniel Stern Cardinale and

Robert Joel Duff for their hospitality and conversations while ZBH was collecting. Research reported was supported by the National Institute of General Medical Sciences of the National Institutes of Health (NIH) under award R35GM137919 (awarded to GSB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

Alexander, J. M., & Edwards, P. J. (2010). Limits to the niche and range margins of alien species. *Oikos*, *119*(9), 1377-1386.

Angert, A. L., Bontrager, M. G., & Ågren, J. (2020). What do we really know about adaptation at range edges?. *Annual Review of Ecology, Evolution, and Systematics*, *51*(1), 341-361.

Arnaud-Haond, S., Teixeira, S., Massa, S. I., Billot, C., Saenger, P., Coupland, G., Duarte, C.M., & Serrao, E. A. (2006). Genetic structure at range edge: low diversity and high inbreeding in Southeast Asian mangrove (Avicennia marina) populations. *Molecular Ecology*, *15*(12), 3515-3525.

Barnard, J.L. (1969) *The Families and Genera of Marine Gammaridean Amphipoda*. Smithsonian Institution Press, Washington, D.C., 535 pp.

Barton, N. H., Etheridge, A. M., Kelleher, J., & Véber, A. (2013). Genetic hitchhiking in spatially extended populations. *Theoretical population biology*, *87*, 75-89.

Bayona-Vásquez, N.J., Glenn, T.C., Kieran, T.J., Pierson, T.W., Hoffberg, S.L., Scott, P.A., Bentley, K.E., Finger, J.W., Louha, S., Troendle, N., Diaz-Jaimes, P., Mauricio, R., Faircloth, B.C. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, 7:e7724 http://doi.org/10.7717/peerj.7724.

Beerli, P., Mashayekhi, S., Sadeghi, M., Khodaei, M., & Shaw, K. (2019). Population genetic inference with MIGRATE. *Current protocols in bioinformatics*, *68*(1), e87.

Bolnick, D. I., & Nosil, P. (2007). Natural selection in populations subject to a migration load. *Evolution*, *61*(9), 2229-2243.

Bousfield, E. L. (1962). *New Haustoriid Amphipods from the Canadian Atlantic Region*. Canada Department of Northern Affairs and National Resources.

Bousfield, E. L. (1965). Haustoriidae of New England (Crustacea: Amphipoda). *Proceedings of the United States National Museum*.

Bousfield, E. L. (1970). Adaptive radiation in sand-burrowing amphipod crustaceans. *Chesapeake Science*, *11*(3), 143-154.

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

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*(2), 255-279.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*(3), 171-182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124-3140.

Colman, S. M., & Mixon, R. B. (1988). The record of major Quaternary sea-level changes in a large coastal plain estuary, Chesapeake Bay, eastern United States. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *68*(2-4), 99-116.

Copilaş-Ciocianu, D., Borko, Š., & Fišer, C. (2020). The late blooming amphipods: Global change promoted post-Jurassic ecological radiation despite Palaeozoic origin. *Molecular Phylogenetics and Evolution*, *143*, 106664.

Croker, R. A., & Hatfield, E. B. (1980). Space partitioning and interactions in an intertidal sand-burrowing amphipod guild. *Marine Biology*, *61*, 79-88.

Dallas, T., Decker, R. R., & Hastings, A. (2017). Species are not most abundant in the centre of their geographic range or climatic niche. *Ecology Letters*, *20*(12), 1526-1533.

Dallas, T. A., Santini, L., Decker, R., & Hastings, A. (2020). Weighing the evidence for the abundant-centre hypothesis. *Biodiversity Informatics*, *15*(3), 81-91.

Dambach, J., Raupach, M. J., Leese, F., Schwarzer, J., & Engler, J. O. (2016). Ocean currents determine functional connectivity in an Antarctic deep-sea shrimp. *Marine Ecology*, *37*(6), 1336-1344.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.

De Ryck, D. J., Koedam, N., Van der Stocken, T., van der Ven, R. M., Adams, J., & Triest, L. (2016). Dispersal limitation of the mangrove Avicennia marina at its South African range limit in strong contrast to connectivity in its core East African region. *Marine Ecology Progress Series*, *545*, 123-134.

Dixon, A. L., Herlihy, C. R., & Busch, J. W. (2013). Demographic and population-genetic tests provide mixed support for the abundant centre hypothesis in the endemic plant Leavenworthia stylosa. *Molecular ecology*, *22*(7), 1777-1791.

Doak, D. F., & Morris, W. F. (2010). Demographic compensation and tipping points in climate-induced range shifts. *Nature*, *467*(7318), 959-962.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, *17*(5), 1170-1188.

Evans, S. M., Sinclair, E. A., Poore, A. G., Steinberg, P. D., Kendrick, G. A., & Vergés, A. (2014). Genetic diversity in threatened Posidonia australis seagrass meadows. *Conservation Genetics*, *15*, 717-728.

Excoffier, L., & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, *23*(7), 347-351.

Fedorka, K. M., Winterhalter, W. E., Shaw, K. L., Brogan, W. R., & Mousseau, T. A. (2012). The role of gene flow asymmetry along an environmental gradient in constraining local adaptation and range expansion. *Journal of Evolutionary Biology*, *25*(8), 1676-1685.

Fifer, J. E., Yasuda, N., Yamakita, T., Bove, C. B., & Davies, S. W. (2022). Genetic divergence and range expansion in a western North Pacific coral. *Science of the Total Environment*, *813*, 152423.

Fristoe, T. S., Vilela, B., Brown, J. H., & Botero, C. A. (2023). Abundant-core thinking clarifies exceptions to the abundant-center distribution pattern. *Ecography*, *2023*(2), e06365.

García-Ramos, G., & Kirkpatrick, M. (1997). Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, *51*(1), 21-28.

Gaston, K. J. (2009). Geographic range limits of species. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1661), 1391-1393.

Grinnell, J. (1922). The role of the" accidental". *The Auk*, *39*(3), 373-380.

Halbritter, A. H., Billeter, R., Edwards, P. J., & Alexander, J. M. (2015). Local adaptation at range edges: comparing elevation and latitudinal gradients. *Journal of Evolutionary Biology*, *28*(10), 1849-1860.

Haldane, J. B. S. (1956). The relation between density regulation and natural selection. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, *145*(920), 306-308.

Haller, B. C., & Messer, P. W. (2023). SLiM 4: multispecies eco-evolutionary modeling. *The American Naturalist*, *201*(5), E127-E139.

Hancock, Z. B., & Wicksten, M. K. (2018). Two new species of sand-burrowing amphipods of the genus Haustorius Müller, 1775 (Amphipoda: Haustoriidae) from the northwestern Gulf of Mexico. *Zootaxa*, *4459*(1), 101-127.

Hancock, Z. B., Hardin, F. O., & Light, J. E. (2019). Phylogeography of sand-burrowing amphipods (Haustoriidae) supports an ancient suture zone in the Gulf of Mexico. *Journal of Biogeography*, *46*(11), 2532-2547.

Hancock, Z. B., Hardin, F. O., Murthy, A., Hillhouse, A., & Johnston, J. S. (2021). Rapid genomic expansion and purging associated with habitat transitions in a clade of beach crustaceans (Amphipoda: Haustoriidae). *Journal of Crustacean Biology*, *41*(3), ruab042.

Hancock, Z. B., Ogawa, H., Light, J. E., & Wicksten, M. K. (2022). Origin and evolution of the Haustoriidae (Amphipoda): a eulogy for the Haustoriidira. *Zoological Journal of the Linnean Society*, *194*(4), 1252-1267.

He, W., Xu, L., Wang, J., Yue, Z., Jing, Y., Tai, S., Yang,J., & Fang, X. (2024). VCF2PCACluster: a simple, fast and memory-efficient tool for principal component analysis of tens of millions of SNPs. *BMC Bioinformatics*, *25*(1), 173.

Hengeveld, R., & Haeck, J. (1982). The distribution of abundance. I. Measurements. *Journal of Biogeography*, 303-316.

Hutchinson, G.E. (1957). Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology. 22 (2): 415–427.

Hutchinson, G.E. (1965). *The Ecological Theater and Evolutionary Play*. New Haven: Yale University Press.

Ivany, L. C., Patterson, W. P., & Lohmann, K. C. (2000). Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary. *Nature*, *407*(6806), 887-890.

Jin, P. Y., Sun, J. T., Chen, L., Xue, X. F., & Hong, X. Y. (2020). Geography alone cannot explain *Tetranychus truncatus* (Acari: Tetranychidae) population abundance and genetic diversity in the context of the center–periphery hypothesis. *Heredity*, *124*(2), 383-396.

Jones, M.E. (2005) *Comparative Phylogeography, Population Structure, and Extended Maternal Care in Three Intertidal Haustoriid Amphipods from the Western North Atlantic*. Unpublished Master's thesis, College of Charleston.

Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, *12*(5), e1004842.

Kelleher, J., Thornton, K. R., Ashander, J., & Ralph, P. L. (2018). Efficient pedigree recording for fast population genetics simulation. *PLoS Computational Biology*, *14*(11), e1006581.

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. *The American Naturalist*, *150*(1), 1-23.

Kottler, E. J., Dickman, E. E., Sexton, J. P., Emery, N. C., & Franks, S. J. (2021). Draining the swamping hypothesis: little evidence that gene flow reduces fitness at range edges. *Trends in Ecology & Evolution*, *36*(6), 533-544.

Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, *21*(4), 1359-1368.

LeCroy, S. E., Richardson, J. S., & Cobb, D. (2000). *An illustrated identification guide to the nearshore marine and estuarine gammaridean Amphipoda of Florida* (Vol. 1, p. 2000). Florida Department of Environmental Protection, Division of Resource Assessment and Management, Bureau of Laboratories.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183-189.

Louthan, A. M., Doak, D. F., & Angert, A. L. (2015). Where and when do species interactions set range limits?. *Trends in Ecology & Evolution*, *30*(12), 780-792.

Lowry, J. K., & Myers, A. A. (2017). A phylogeny and classification of the Amphipoda with the establishment of the new order Ingolfiellida (Crustacea: Peracarida). *Zootaxa*, *4265*(1), 1-89.

De Manuel, M., Kuhlwilm, M., Frandsen, P., Sousa, V. C., Desai, T., Prado-Martinez, J., et al. (2016). Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science*, *354*(6311), 477-481.

Marcus, J., Ha, W., Barber, R. F., & Novembre, J. (2021). Fast and flexible estimation of effective migration surfaces. *Elife*, *10*, e61927.

Mayr, E. (1963). *Animal species and evolution*. Harvard University Press.

Morrissey, M. B., & de Kerckhove, D. T. (2009). The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. *The American Naturalist*, *174*(6), 875-889.

Nagylaki, T. (1977). Decay of genetic variability in geographically structured populations. *Proceedings of the National Academy of Sciences*, *74*(6), 2523-2525.

Ntuli, N. N., Nicastro, K. R., Zardi, G. I., Assis, J., McQuaid, C. D., & Teske, P. R. (2020). Rejection of the genetic implications of the “Abundant Centre Hypothesis” in marine mussels. *Scientific Reports*, *10*(1), 604.

Oldfather, M. F., Kling, M. M., Sheth, S. N., Emery, N. C., & Ackerly, D. D. (2020). Range edges in heterogeneous landscapes: Integrating geographic scale and climate complexity into range dynamics. *Global Change Biology*, *26*(3), 1055-1067.

Oswald, J. A., Overcast, I., Mauck III, W. M., Andersen, M. J., & Smith, B. T. (2017). Isolation with asymmetric gene flow during the nonsynchronous divergence of dry forest birds. *Molecular Ecology*, *26*(5), 1386-1400.

Paz-Vinas, I., Quéméré, E., Chikhi, L., Loot, G., & Blanchet, S. (2013). The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Molecular Ecology*, *22*(12), 3279-3291.

Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data. *Evolution*, *67*(11), 3274-3289.

Peter, B. M., & Slatkin, M. (2015). The effective founder effect in a spatially expanding population. *Evolution*, *69*(3), 721-734.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94-100.

Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B., & Thompson, J. D. (2017). Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews*, *92*(4), 1877-1909.

Polechová, J., & Barton, N. H. (2015). Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences*, *112*(20), 6401-6406.

Pringle, J. M., & Wares, J. P. (2007). Going against the flow: maintenance of alongshore variation in allele frequency in a coastal ocean. *Marine Ecology Progress Series*, *335*, 69-84.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559-575.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573-589.

Ringbauer, H., Kolesnikov, A., Field, D. L., & Barton, N. H. (2018). Estimating barriers to gene flow from distorted isolation-by-distance patterns. *Genetics*, *208*(3), 1231-1245.

Rivadeneira, M. M., Hernáez, P., Antonio Baeza, J., Boltana, S., Cifuentes, M., Correa, C., Cuevas, A., del Valle, E., Hinojosa, I., Ulrich, N., Valdivia, N., Vásquez, N., Zander, A., & Thiel, M. (2010). Testing the abundant-centre hypothesis using intertidal porcelain crabs along the Chilean coast: linking abundance and life-history variation. *Journal of Biogeography*, *37*(3), 486-498.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular ecology*, *28*(21), 4737-4754.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219-1228.

Samis, K. E., & Eckert, C. G. (2007). Testing the abundant center model using range-wide demographic surveys of two coastal dune plants. *Ecology*, *88*(7), 1747-1758.

Santini, L., Pironon, S., Maiorano, L., & Thuiller, W. (2019). Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. *Ecography*, *42*(4), 696-705.

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, & Systematics*, *40*, 415-436.

Siegel, D. A., Kinlan, B. P., Gaylord, B., & Gaines, S. D. (2003). Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series*, *260*, 83-96.

Silva-Arias, G. A., Reck-Kortmann, M., Carstens, B. C., Hasenack, H., Bonatto, S. L., & Freitas, L. B. (2017). From inland to the coast: spatial and environmental signatures on the genetic diversity in the colonization of the South Atlantic Coastal Plain. *Perspectives in Plant Ecology, Evolution and Systematics*, *28*, 47-57.

Soberón, J., Peterson, T.A., Osorio-Olvera, L. (2018). A comment on “Species are not most abundant in the centre of their geographic range or climatic niche”. *Rethinking Ecology*, 3: 13-18.

Verheye, M. L., Backeljau, T., & d’Udekem d’Acoz, C. (2016). Looking beneath the tip of the iceberg: diversification of the genus Epimeria on the Antarctic shelf (Crustacea, Amphipoda). *Polar Biology*, *39*, 925-945.

Wares, J. P., & Pringle, J. M. (2008). Drift by drift: effective population size is limited by advection. *BMC Evolutionary Biology*, *8*(1), 1-7.

White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1688), 1685-1694.

Wiens, B. J., & Colella, J. P. (2024a). triangulaR: an R package for identifying AIMs and building triangle plots using SNP data from hybrid zones. *bioRxiv*, 2024-03.

Wiens, B. J., & Colella, J. P. (2024b). That's Not a Hybrid: How to Distinguish Patterns of Admixture and Isolation By Distance. *Molecular Ecology Resources*, e14039.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*(2), 97.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Zhan, S., Zhang, W., Niitepold, K., Hsu, J., Haeger, J. F., Zalucki, M. P., Altizer, S., de Roode, J.C., Reppert, S.M., & Kronforst, M. R. (2014). The genetics of monarch butterfly migration and warning colouration. *Nature*, *514*(7522), 317-321.

887 **FIGURE LEGENDS**

888 **Figure 1**. *Sample locations and characterization of population genetic structure*. A)
889 Points are sampled localities, colored by latitude. Northbound arrows are the Gulf
890 Stream, southbound are the Labrador Current, size differences represent different
891 current intensities. Inset is a photograph of a shovel bug collected from New Jersey. B)
892 PCA results; points are individuals, colored by the latitude of their sample site. C)
893 Isolation-by-distance; points are sample localities plotted by their pairwise linearized F_{ST}
894 against geographic distance in miles. Colors are comparisons within groups (north or
895 south of the Chesapeake) or between them.

896

897 **Figure 2.** *Admixture and effective migration surface*. A) Clustering (*conStruct*) analysis
898 results using k = 2. B) Effective migration surface from *FEEMS*. Cooler colors represent
899 regions of higher migration, warmer indicate potential barriers. Pies represent the
900 admixture proportions inferred from *conStruct* averaged across all individuals sampled
901 at a locality. Stars represent inferred origins of the population expansion from the ψ
902 analysis. Population abbreviations are defined in *Table S1*.

903

904 **Figure 3**. *Patterns of genetic diversity*. A) Individual genetic diversity (π) as a function of
905 latitude. B) Inbreeding coefficient as a function of the distance from the center of the
906 range, taken as 36° latitude, in miles. Points are colored by the latitude of their sample
907 location as in **Fig. 1A**.

908

Figure 4. *Directionality index (ψ) results*. A) Empirical estimates of ψ, points are the summed values of ψ for all comparisons of population *i* to populations *j*. B) Simulation results for each stepping-stone model: “center_asym” is abundant center with asymmetric migration; “center_sym” is abundant center with symmetric migration; “equal_asym” is equal deme size with asymmetric migration; and “equal_sym” is equal deme size with symmetric migration. Each point per population is a different replicate. Points are colored by estimated genetic diversity (π) of the sampled locality. Dashed line is when ψ = 0, which is expected under equilibrium isolation-by-distance.

FIGURES

Figure 1. *Sample locations and visualization of population genetic variation*. A) Points are sampled localities, colored by latitude. Northbound arrows are the Gulf Stream, southbound are the Labrador Current, size differences represent different current intensities. Inset is a photograph of a shovel bug collected from New Jersey. B) PCA results; points are individuals, colored by the latitude of their sample site. C) Isolation-by-distance; points are sample localities plotted by their pairwise linearized F_{ST} against geographic distance in miles. Colors are comparisons within groups (north or south of the Chesapeake) or between them.

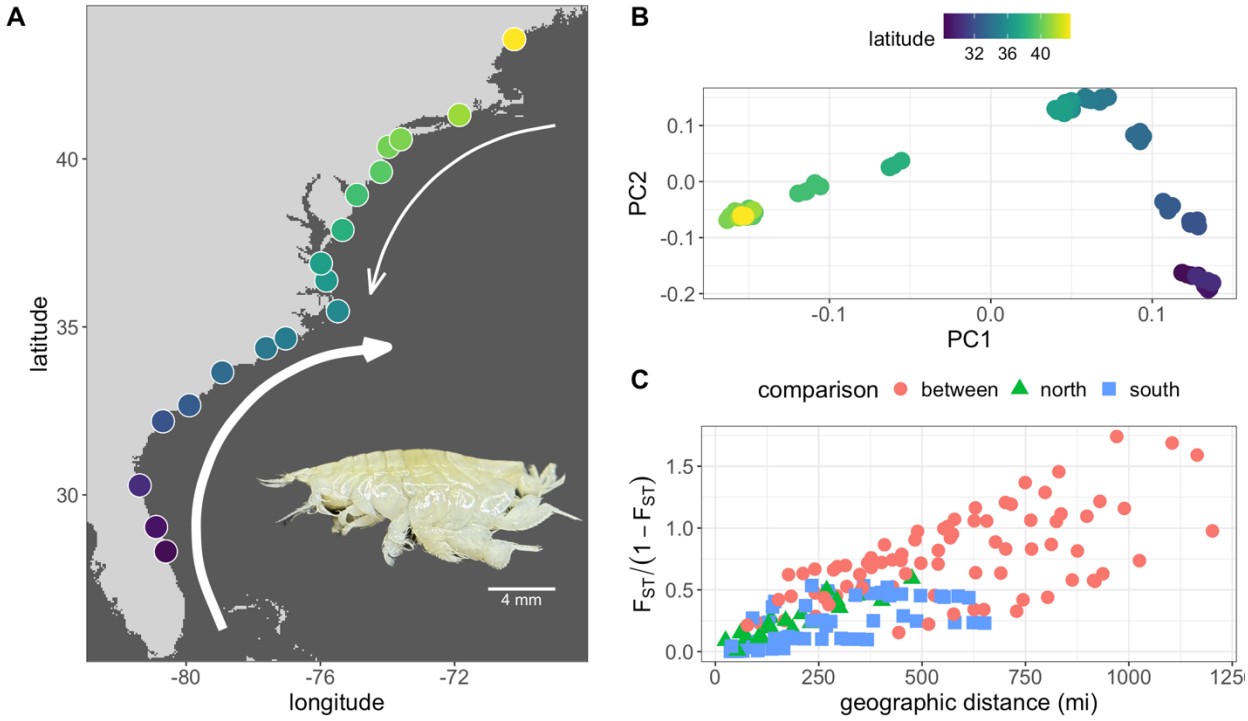
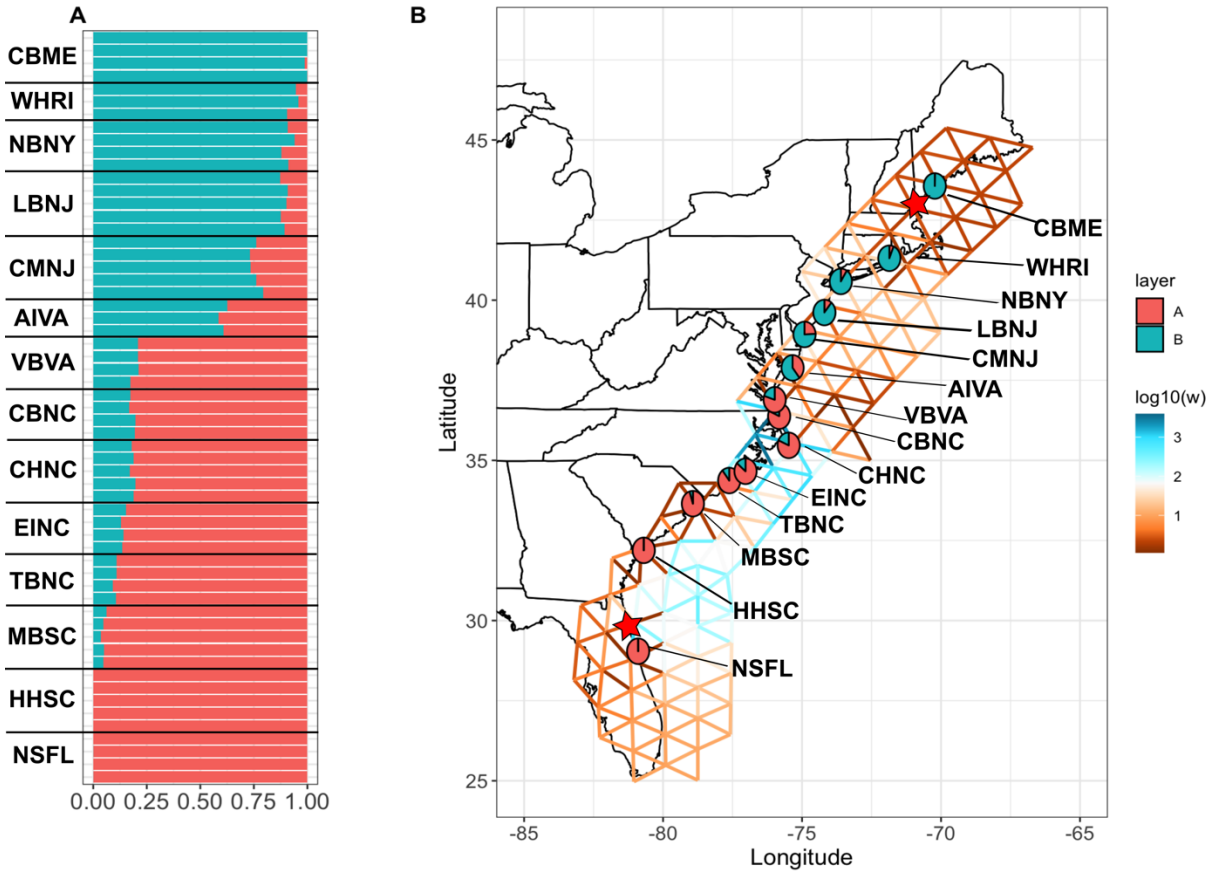
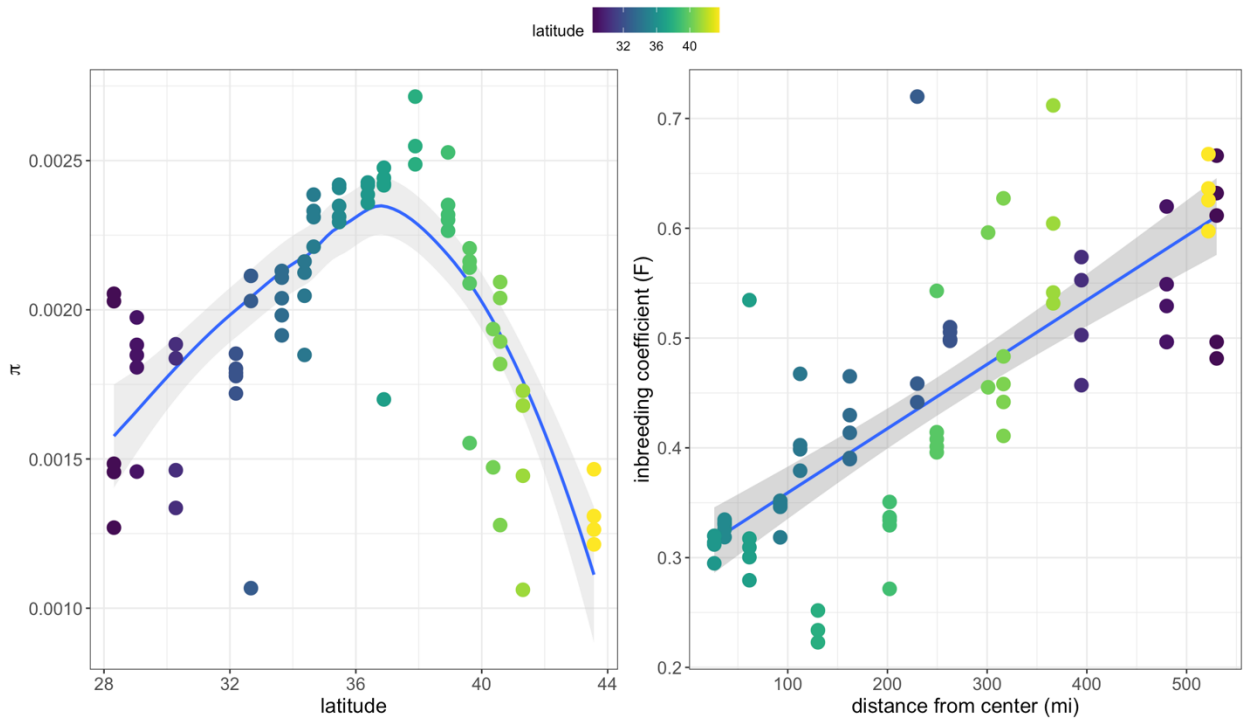


Figure 2. *Admixture and effective migration surface*. A) Clustering (*conStruct*) analysis results using k = 2. B) Effective migration surface from *FEEMS*. Cooler colors represent regions of higher migration, warmer indicate potential barriers. Pies represent the admixture proportions inferred from *conStruct* averaged across all individuals sampled at a locality. Stars represent inferred origins of the population expansion from the ψ analysis. Population abbreviations are defined in *Table S1*.



935 **Figure 3**. *Patterns of genetic diversity*. A) Individual genetic diversity (π) as a function of
936 latitude. B) Inbreeding coefficient as a function of the distance from the center of the
937 range, taken as 36° latitude, in miles. Points are colored by the latitude of their sample
938 location as in **Fig. 1A**.



939

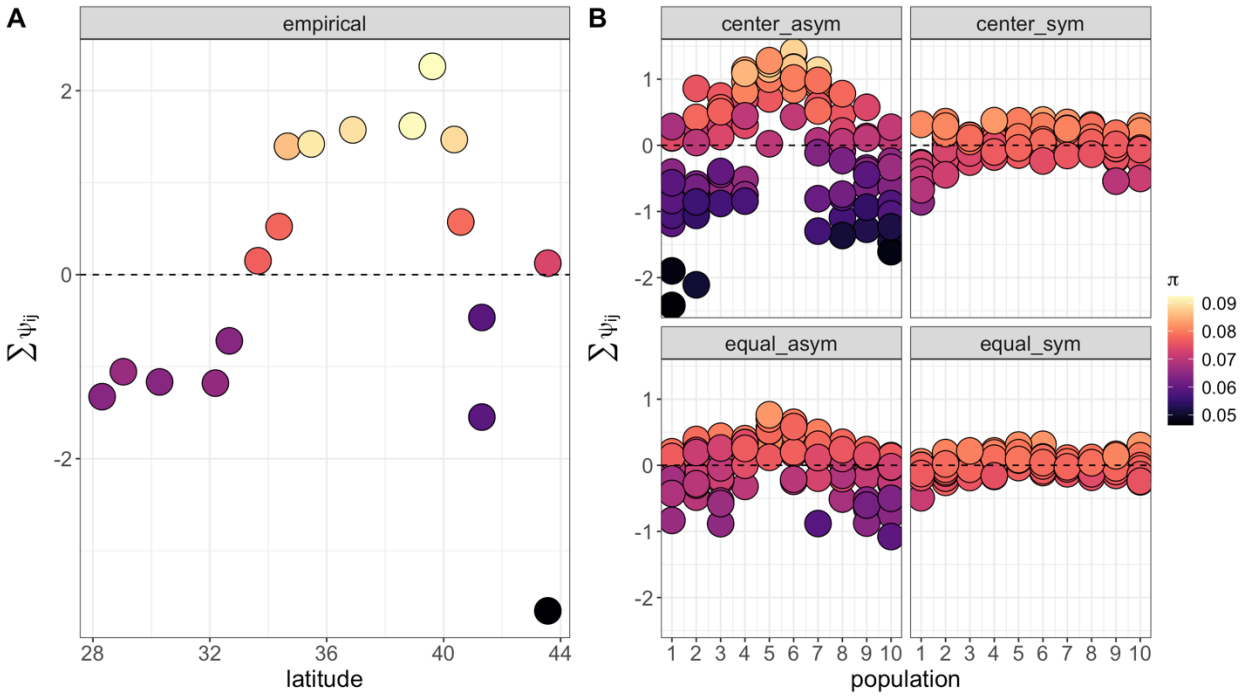



Figure 4. *Directionality index (ψ) results*. A) Empirical estimates of ψ, points are the summed values of ψ for all comparisons of population *i* to populations *j*. B) Simulation results for each stepping-stone model: “center_asym” is abundant center with asymmetric migration; “center_sym” is abundant center with symmetric migration; “equal_asym” is equal deme size with asymmetric migration; and “equal_sym” is equal deme size with symmetric migration. Each point per population is a different replicate. Points are colored by estimated genetic diversity (π) of the sampled locality. Dashed line is when ψ = 0, which is expected under equilibrium isolation-by-distance.