# ASYMMETRIC GENE FLOW MAINTAINS AN ABUNDANT RANGE CENTER

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 an abundant range center


## ABSTRACT

The abundant center hypothesis postulates that population density is the highest in the center of the range and declines towards the periphery. While many evolutionary hypotheses have been offered to explain what generates this pattern, ecologically it's assumed the pattern emerges due to increasing distance from a species ideal niche with edge populations occupying a "marginal" niche of low suitability. However, the ubiquity of the abundant center pattern has recently been questioned. Furthermore, it has proven difficult to assess both the pattern of the abundant center and the potential generative scenarios (e.g., gene swamping, migration limitation) due to the inherent complexity of species ranges. In this study, we evaluate the suitability of the abundant center hypothesis to explain patterns of genetic diversity and inbreeding in shovel bugs – small, sand-burrowing coastal crustaceans. In addition, we evaluate potential biases in migration direction due to the dominant current regimes in the North Atlantic using a common method for querying the site frequency spectrum, the directionality index ($\psi$). While patterns of genetic diversity are consistent with an abundant center, we find a strong signature of recent central colonization from the range periphery. Using simulations, we evaluate the performance of $\psi$ under different migration regimes, finding that, in the presence of asymmetric migration, it strongly infers an edge-to-center colonization dynamic. We conclude that asymmetric migration towards the center can lead to signatures of an abundant center without necessarily implying a reduction in niche quality towards the range periphery.

**KEYWORDS:** abundant center hypothesis, range evolution, directionality index, migration, coastal, shovel bug, isolation-by-distance

## 1 INTRODUCTION

All species are restricted to finite ranges ultimately due to a failure to locally adapt along the periphery (Haldane 1956; Mayr 1963). Why populations fail to adapt and continue to expand their range, what is the long-term stability of range size, and how edges form and persist across a continuous environment are fundamental questions in the ecology and evolution of ranges (Kirkpatrick & Barton 1997; Sexton et al. 2009; Gaston 2003). Answering these questions has major implications for predicting species' responses to climate change (Doak & William 2010) as well as the propensity of invasive species to rapidly expand their range (Alexander & Edwards 2010).

A prime driver molding species' ranges is heterogeneity in population abundance, which has the potential to generate disparate evolutionary responses across populations. Grinnell (1922) envisioned a species range as an amoeboid organism with the center representing the "species metropolis," an area of high abundance and adaptiveness, with decreasing density towards the periphery. Along the edges of the range are "accidentals", occasional migrants that diffuse from the center like gas particles to the unsuitable fringe (Grinnell 1922). This concept of species ranges has since been formalized into the "abundant center" or "central marginal" hypothesis, which states that the region of the highest abundance is in the center of the

range and decreases monotonically towards the edge (Hengeveld & Haeck 1982; Brown 1984). While intuitive, the generality of the abundant center hypothesis has generated considerable debate (e.g., Dallas et al. 2020; Santini et al. 2019).

Ecologically, range limits are often conceptualized as the relationship between the species' fundamental and realized niche, where the former represents the full potential breadth of a species range whereas the latter reflects biotic and abiotic constraints preventing full occupation of the fundamental niche (Hutchinson 1957, 1965). While appeals to a species' niche are intuitive for present distributions, often we are curious about how they evolved in the first place and what forces are maintaining its current bounds. These forces are fundamentally *evolutionary* in nature; niches are the result of past adaptation; therefore, their limits should represent limits to a species adaptive capacity. Haldane (1956) was among the first to suggest that population replacement fails along the periphery due to the inability of edge populations to adapt on account of their reduced size relative to central populations; genetic drift overwhelms selection on the periphery. Mayr (1963) suggested this failure to adapt was due to the influx of migrants from the range center, which carried with them alleles adapted to the center but maladapted on the edge. This asymmetric migration was hypothesized to swamp local adaptation, preventing edge populations from establishing. Kirkpatrick & Barton (1997) formalized this "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. Thus, 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, in a large review of the literature, Kottler et al. (2021), find little empirical evidence for gene swamping. Alternatively, they suggest that migration limitation causes edge populations to suffer from a greater drift load and hence fail to adapt; as opposed to gene swamping, an increase in migration is expected to rescue edge populations, not doom them (Halbritter et al. 2015; Evans et al. 2014; Silva-Arias et al. 2017). In this model, the rate at which stochastic extinction occurs in each patch exceeds the rate at which migrants can colonize the patch. Migration limitation requires that rates of dispersal decrease toward the edge, which implies some external barriers exist along the edges but not, presumably, among central populations.

In the ocean, particularly for pelagic organisms, dispersal limitation may be a product of 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 the settling time and current velocity (Pringle & Wares 2007; Wares & Pringle 2008). Thus, marine ranges can be determined by complex abiotic factors that can cause mass movement of individuals in asymmetric ways that may not be due to biases in population abundance or reduction in niche quality. 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. Such an asymmetry in dispersal could theoretically generate an abundant center pattern in the absence of reduced niche quality along the periphery if the current was strong enough that successive colonization events became rarer and rarer at the range extremes due to most individuals migrating downstream.

Determining the limits of ranges (e.g., Dallas et al. 2017; Pironon et al. 2017; Soberón et al. 2018) and the underlying processes generating those limits remain challenging (Kottler et al. 2021; Fristoe et al. 2023), especially in organisms not amenable to transplant experiments. From a population genetic perspective, declines in abundance along the periphery of ranges are often investigated using patterns of genetic diversity (e.g., Ntuli et al. 2020; Dixon et al. 2013; Jin et al. 2020). Regions of high abundance are expected to harbor the greatest diversity, whereas edge populations are expected to have increased inbreeding (Haldane 1956; Arnaud-Haond et al. 2006; De Ryck et al. 2016) and differentiation (Wright 1943; Eckert et al. 2008). Furthermore, several techniques seek to detect asymmetries in the direction of migration (Peter & Slatkin 2013, Beerli et al. 2019), whereas others model regions of the range with high resistance to gene flow (Petkova et al. 2016; Marcus et al. 2021).

An expectation of the abundant center hypothesis is that edge populations may be experiencing stochastic extinction and recolonization (in Grinnell's words, the "accidentals" are expanding and contracting along the periphery; Grinnell 1922). Theoretically, this might lead to a signature in the site frequency spectrum (SFS) of

recurrent range expansion in edge populations. Peter & Slatkin (2013; 2015) introduced a method they call the “directionality index”, $\psi$, which detects shifts in the frequency of alleles across populations due to founder events in the expanding front. The logic behind the index is that in large, stable populations there should be an abundance of low frequency alleles, but subsequent founding events can cause those alleles to reach high frequency via drift. Hence, alleles that are still segregating in the entire population should have increasingly higher frequency in the direction of a range expansion. In the case of repeated expansion along the edge in an abundant center range, the directionality index should therefore detect a signature of higher frequency alleles on the periphery relative to the center.

Small amphipod crustaceans called shovel bugs (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*), shovel bugs are expected to be relatively poor dispersers, and this has been largely confirmed in past studies (Jones 2005; Hancock et al. 2019). In the western North Atlantic, one species of shovel bug (*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). A curious feature of their southernmost distribution is the abundance of suitable sandy beaches farther south than they are found, with no 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 — ease of identifying range boundaries, and the one-dimensional nature of their range, shovel bugs are ideal organisms for evaluating hypotheses concerning range evolution and for testing the abundant center hypothesis.

In this study, we evaluate whether the shovel bug *H. canadensis* in the western North Atlantic constitutes a single, continuous population or contains discrete structure. Next, we investigate whether patterns of genetic diversity and inbreeding are consistent with expectations of the abundant center. 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 ranges 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 Study system*

Deemed “the most interesting group of amphipods” by J. Laurens Barnard, shovel bugs are haustoriid amphipods that are highly modified from their hypothesized gammarid ancestor which include 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 (hence their common name; **Fig. 1A**) (Barnard 1969; Hancock & Wicksten 2018). Shovel bugs are entirely fossorial, living in the first 1–10 cm of sand and occupying almost exclusively the narrow swash zone (Croker & Hatfield 1980).

179 They lack a pelagic larva; instead, the females brood the juveniles through the first few
180 molts in their marsupium. Adult shovel bugs range in body size from 2–18 mm,
181 sometimes within the same species, and past work has found strong correlations
182 between body size and latitude (Hancock et al. 2021), though there may also be an
183 effect of sand grain size (LeCroy et al. 2002; Hancock & Wicksten 2018).
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**). Collections began in West Palm Beach,
188 Florida, USA to identify the southernmost established population; no specimens were
189 located until Cocoa Beach, Florida. Similarly, sampling continued beyond Crescent
190 Beach, Maine, but no samples were located further north (last attempted location was
191 Popham Beach, ~93 km north). Specimens were sifted from sand using a 435 $\mu$m sieve
192 plate, plucked from the plate with forceps, and preserved in 95% ethanol. Individuals
193 were identified to species in the lab using a dissecting microscope and key from
194 Hancock & Wicksten (2018), and then stored at -20°C prior to DNA extraction.
195
196 *2.3 Molecular techniques*
197 Whole genomic DNA was extracted from either the whole individual or from pereopods
198 4–7 using Qiagen DNeasy Blood & Tissue Kit following manufacturer's protocols.
199 Library prep was performed using Restriction site Associated DNA sequencing (2RAD)
200 following Bayona-Vásquez et al. (2019), with slight modifications. All 96 samples were
201 quantified using a Qubit™ 1X dsDNA High Sensitivity (HS) kit on a Qubit 2.0

202 fluorometer. Samples were then normalized to have 100ng per sample (5 $\mu$L) on a 96
203 well plate, and then digested in a 15 $\mu$L solution with 1.5 $\mu$L 10X CutSmart Buffer (New
204 England Biolabs, USA), with 10 units of EcoRI (0.5 $\mu$L), 10 units MseI (0.5 $\mu$L), 3.5 $\mu$L
205 Invitrogen™ UltraPure™ DNase/RNase-Free Distilled Water, 2.5 $\mu$L of EcoRI
206 compatible adapter (2 $\mu$L), and 0.25 $\mu$M MseI compatible adapter ($\mu$L) on a Eppendorf
207 Mastercycler PRO S 6325 Thermal Cycler thermocycler at 37°C for one hour with the
208 heated lid turned off. Immediately after digestion we added 5 $\mu$L of ligation mastermix
209 composed of 2.75 $\mu$L of nuclease free water, 0.5 $\mu$L 10X ligation buffer (NEB), 1.5 $\mu$L 10
210 mM rATP (Promega), and 0.25 $\mu$L T4 DNA Ligase 400,000 U/ml to each sample. The
211 ligation reaction was implemented on the same thermocycler as above for two cycles of
212 22°C for 20 min, 37°C for 10 min, followed by 80°C for 20 min to stop the enzyme
213 activity. We then pulled 5 $\mu$L of each sample and split these into two microcenterfuge
214 tubes and performed a 1.25X SpeedBead cleanup (Rohland and Reich, 2012).
215 To reduce PCR duplicates and add the iTru5–8N barcode (which has a segment
216 of eight random base pairs) we implemented a single PCR cycle with six tubes each
217 with a 50 $\mu$L reaction. Each reaction contained 10 $\mu$L nuclease free water, 25 $\mu$L KAPA
218 HiFi hot start ready mix (Roche, USA), 5 $\mu$L iTru5–8N barcode (5 $\mu$M), and 10 $\mu$L of
219 library and were added using a single cycle of 98˚C for 1 min, 60˚C for 30s, and 72˚C
220 for 6 min. All six PCR reactions were then pooled and a 1.5X SpeedBead cleanup was
221 performed and resuspended in 30 $\mu$L of nuclease free water. Next, we split the library
222 into three 55 $\mu$L reactions with three unique iTru7 adapters. Each reaction had 10 $\mu$L of
223 nuclease free water, 25 $\mu$L KAPA HiFi hot start ready mix (Roche, USA), 5 $\mu$L P5 primer
224 (5 $\mu$M) and 5 $\mu$L of each iTru7 (5 $\mu$M) added separately. The PCR was implemented on

225 the thermocycler mentioned above under the following conditions: 95˚C for 3 mins, 8
226 cycles of 98˚C 20s, 60˚ 15s, 72˚C 30s, and a final 72˚C for 5 mins. All PCR reactions
227 were pooled and were clean using a 1.5X SpeedBead cleanup and resuspended in 30
228 $\mu$L of nuclease free water.

229 We performed size selection on the pooled library using a Pippin Prep (Sage
230 Science, USA) with internal standards on a 1.5% dye free cassette, target 350–480 bp
231 range (mean= 415 bp). After size selection to enrich the library, we performed a PCR
232 under the same conditions for the iTru7 reaction above, except with only four 50 $\mu$L
233 reactions and 6 cycles, also both the P5 and P7 primers (5 $\mu$M) were used. Finally, we
234 performed a 1X SpeedBead clean up and quantified the library on the Qubit 2.0 to
235 ensure that our library had a high enough concentration for sequencing. The library was
236 then sequenced on 12.5% percent of a NovaSeq X lane at the University of Michigan
237 Advanced Genomics Core.

238 Raw sequence reads were assembled and filtered using the de novo pipeline in
239 STACKS v.2.67 (Catchen et al. 2013; Rochette et al. 2019). A custom *process_radtags*
240 function was provided to us by J. Catchen to handle the unique FASTQ files of 2RAD,
241 which include an additional adapter sequence. Reads were assembled using the de
242 novo pipeline (Catchen et al. 2011); briefly, stacks were built in *ustacks* with a maximum
243 nucleotide distance allowed of 2 and minimum number of reads needed to seed a stack
244 of 3. Stacks were merged allowing a maximum of 2 gaps and 6 mismatches. Next, we
245 used the *populations* function, filtering out alleles that did not occur in at least 10
246 populations and keeping those that were present in 75% of individuals per population.
247 Sequences were output as two VCF files, one that included only variant sites and a

248 second that contained both variant and monomorphic sites. VCF files were additionally
249 filtered using *vcftools* v.1.16 (Danecek et al. 2011); first, individuals with >80% missing
250 data were removed. Next, we filtered out loci with more than two alleles (--max-alleles
251 2) and with a minor allele frequency of less than 0.05 (--maf 0.05). We then filtered out
252 loci with a minimum mean depth of less than 20 (--min-meanDP 20) and removed all
253 indels (--remove-indels). After filtering, we evaluated the relationship between mean
254 depth, estimated heterozygosity, and distance from the range center via linear
255 regression to determine if mean depth might be a confounding factor impacting the
256 detection of an abundant center.
257
258 *2.4 Population genetic analysis*
259 To evaluate potential patterns in genetic diversity, we estimated individual-level
260 heterozygosity using *pixy* (Korunes & Samuk 2021) on the VCF that included invariant
261 sites. The inbreeding coefficient (*F*; Wright, 1931), estimated as the difference between
262 the expected and observed homozygosity per individual, was estimated using *vcftools* (-
263 -het).
264 To test for geographic structure, we first used *pixy* to estimate population-level
265 pairwise $F_{ST}$ (Weir & Cockerham 1984), which we then linearized and regressed against
266 geographic distance to test for isolation-by-distance (Wright 1943; Rousset 1997). Next,
267 we performed a principal component analysis (PCA) to identify signs of genetic
268 clustering using *VCF2PCACluster* (He et al. 2024). Lastly, we performed a cluster
269 analysis using *fastSTRUCTURE* (Raj et al. 2014) with $k$ = 2–5. To determine whether
270 identified clusters and admixed individuals were merely a product of gapped sampling

271 or represented discrete populations, we performed a triangle plot analysis using
272 triangulaR (Wiens & Colella 2024a). In the presence of discrete structure, individuals
273 with a hybrid index approaching 0.5 should have concomitantly high interclass
274 heterozygosity. However, a continuous population characterized solely by isolation-by-
275 distance will display a high hybrid index but no increase in interclass heterozygosity
276 (Wiens & Colella 2024b).

277 We evaluated potential barriers to migration using the program *FEEMS* (Marcus
278 et al. 2021), which estimates an effective migration surface using edge-specific
279 parameter estimation in a pseudo-likelihood framework. Lastly, we tested for a signature
280 of range expansion along the periphery of the population using Peter & Slatkin's (2013)
281 directionality index, $\psi$, as implemented in *rangeExpansion*
282 (https://github.com/BenjaminPeter/rangeexpansion). The directionality index is
283 calculated by comparing the difference in allele frequencies between pairwise
284 populations, conditioned on the allele being present in all populations. If the allele is not
285 lost by drift, then it is expected to be at higher frequency in founder populations than in
286 the source, because if founders carry the allele at all it will necessarily be at greater
287 frequency. Thus, successively more recently founded populations in an expanding
288 range will have increasingly positive values of $\psi$, permitting the identification of both the
289 origin and the direction of the range expansion (Peter & Slatkin 2013; 2015). In
290 populations at an isolation-by-distance equilibrium, simulations show that $\psi$ should be
291 zero. Therefore, we predict that if $\psi$ is zero, then there is no marked departure from
292 migration-drift equilibrium and hence no signature of reduced fitness or recurrent
293 colonization in edge populations. However, if dispersal limitation is driving edge

294 dynamics, we expect $\psi$ should be positive in the direction of the edge. We estimated the
295 hypothesized origin of the expansion both on the entire dataset and by subsetting it into
296 populations north and south of Chesapeake Bay.
297
298 *2.5 Simulations*
299 The direction of population expansion, as inferred from $\psi$, may be influenced by spatial
300 clines in population density even under scenarios with isotropic dispersal. To determine
301 whether the results of our $\psi$ analyses are robust to spatial variation in population
302 density (see Results/Discussion), we conducted forward time simulations using SLiM
303 v.4.0.1 (Haller & Messer 2023). We established a 10-deme stepping-stone model with
304 four different range scenarios: 1) equal population size per deme with symmetrical
305 migration; 2) equal population size per deme with asymmetrical migration from the edge
306 populations towards the center; 3) an abundant center model in which the central
307 demes had five times more individuals than the edge with symmetric migration; and 4)
308 abundant center with asymmetric migration from the edges. We performed 10 replicates
309 for each scenario. In addition, we performed simulations for the abundant center with
310 equal migration to evaluate how differences in population size between the central and
311 edge demes might drive patterns in $\psi$. We evaluated differences in population size of
312 5:1, 25:1, and 50:1, with 10 replicates each. Simulations were run for 100,000
313 generations (>10*N*) to ensure demographic equilibrium, and tree sequences were
314 recorded for subsequent processing using *pyslim* v.1.0.1 (Kelleher et al. 2018) and *tskit*
315 v.0.5.3 (Kelleher et al. 2018). For tree sequences with more than one root, we used
316 recapitation to simulate a neutral coalescent history until only a single root remained,

317 and then mutations were added at a rate of $1x10^{-7}$ to the trees using *msprime* v.1.2.0
318 (Kelleher et al. 2016). Next, we randomly sampled five individuals per deme and output
319 their sequences in VCF format. From here, we trimmed the VCFs to only include
320 biallelic SNPs, and converted them to BED files using *plink* v.1.9 (Purcell et al. 2007).
321 Next, to use the *rangeExpansion* package, we converted the individual coordinates from
322 SLiM, which were in discrete demes, into Earth coordinates as required by
323 *rangeExpansion*; thus, we chose 10 coordinates 5° apart in latitude. Because the units
324 of distance are arbitrary in SLiM, we needed only to ensure that deme 1 is closest to
325 deme 2, etc., to adequately capture the SLiM environment in an Earth coordinate
326 system. We then calculated $\psi$ for each simulation to see how the different range
327 scenarios might cause bias inferences made from $\psi$.
328
329 **3 RESULTS**
330
331 *3.1 ddRAD assembly*
332 A total of 2,605,149 loci passed filtering in STACKS with a mean effective per-sample
333 coverage of 23.0, a standard deviation of 13.6, with a minimum of 4.0 and maximum of
334 67.4. The mean number of sites per locus was 220.1. After filtering for biallelic SNPs
335 and missingness in *vcftools*, the final dataset included 7,405 SNPs. We find a significant
336 positive relationship between mean depth per individual and our estimates of
337 heterozygosity from *pixy* ($P < 5.3e–7$, $R^2 = 0.27$; *Table S1*); in addition, we find a
338 significant negative relationship between mean depth and the distance from the center
339 of the range (taken as 36° latitude; $P < 0.002$, $R^2 = 0.10$) (see *Table S1*). To account for

340 the potential confounding between measures of heterozygosity, mean depth, and their
341 relationship with the distance from the range center, all subsequent multivariate linear
342 regressions are performed with mean depth as a confounding variable. In addition, we
343 tested for a significant correlation after dropping all individuals with a mean depth
344 greater than 30 to determine if a few individuals with high depth were driving the
345 patterns. While a significant correlation between heterozygosity and mean depth
346 remained ($P < 2.7e–5$; $R^2 = 0.38$), it was no longer significantly correlated with distance
347 from the center ($P = 0.57$; $R^2 = –0.02$) (*Fig. S1*).
348
349 *3.2 Genetic structure*
350 Results of the *fastSTRUCTURE* analysis supported three population clusters ($k = 3$) as
351 the model that maximized the marginal likelihood – one from Florida to South Carolina,
352 a second from North Carolina to Virginia, and a third that included the remaining
353 populations from Maryland to Maine (*Fig. S2*). However, the PCA reveals relatively
354 continuous structure, with a possible break between populations north and south of the
355 Chesapeake (**Fig. 1B**). A discrete pattern of ancestry was also not supported by the
356 triangle plot analysis, which found that, despite increasing hybrid ancestry towards
357 Chesapeake Bay, interclass heterozygosity did not increase dramatically (*Fig. S3*).
358 Furthermore, we found no discrete breaks in patterns of isolation-by-distance when
359 regressing linearized $F_{ST}$ against geographic distance between comparisons within
360 southern and northern populations and between them (**Fig. 1C**). Subsequent analyses
361 therefore treat the shovel bug population as if it is one continuous population without
362 discrete breaks.

363

364 *3.3 Patterns of diversity*

365 The patterns of genetic diversity and inbreeding we inferred are strongly indicative of an
366 abundant center. After incorporating sequencing depth as a possible confounding effect,
367 measures of individual heterozygosity are significantly predicted by the distance from
368 the center of the range ($P < 3.79e–16$; mean depth $P = 5.1e–5$). This relationship
369 remained significant after dropping individuals with mean depth greater than 30 ($P <$
370 $0.0009$; $R^2 = 0.25$) (*Fig. S4*). Similarly, measures of inbreeding are the highest at the
371 range edges, declining as a function of distance from the center (**Fig. 2B**). These
372 patterns are consistent with field observations, in which the ease of collecting and
373 observed highest abundance was in the central range locations, with shovel bugs at the
374 edge sample locations being rarer and thus more difficult to collect. However, we note
375 that while distance is a much stronger predictor of patterns of genetic diversity and
376 inbreeding, mean sequencing depth was still a significant predictor in each model
377 (*Table S1*).

378 The effective migration surface estimated from *FEEMS* suggests the center of
379 the range has the highest ease of movement, with reduced effective migration towards
380 the range edges (*Fig. S5*). However, in accordance with results from *fastSTRUCTURE*
381 and PCA, the migration surface indicates a strong depression in effective migration
382 across the Chesapeake, which might suggest it acts as a partial barrier to gene flow.

383

384 *3.4 Directionality index & simulations*

The inferred origin of the range expansion of the entire sample of individuals, as indicated by the lowest value of $\psi$, was 28.3° latitude in northern Florida (**Fig. 3A**). When the data were subset into two regions, south and north (based on the cluster analysis), the inferred origins were at 28.3° and 43.5° latitude, respectively. The sum of $\psi$ for each population resulted in a pattern of increasing $\psi$ towards the range center, mirroring the increase in heterozygosity (**Fig. 3A**). This pattern of $\psi$ superficially suggests a range expansion toward the center of the range from the edges following the dominant current regimes of the Gulf Stream running south to north and the Labrador Current running north to south, each converging in the center of the range (**Fig. 1A, 3A**). This pattern of expansion is in opposition to expectations from the gene swamping hypothesis and suggests that asymmetric migration due to the currents might be maintaining range edges, which would be consistent with the migration limitation hypothesis.

Simulation results from SLiM support this interpretation. Across replicates, the strongest divergences from $\psi = 0$, which is expected at migration-drift equilibrium, occur in simulations with asymmetric migration from the edges toward the center, with the most striking inflation in the scenario in which there is both an abundant center and asymmetric migration (**Fig. 3B**). While there is some deviation from zero along the edges in scenarios with symmetric migration in both the equal and center-abundant scenarios, they do not systematically produce a positive $\psi$ in the center as in the case of asymmetric migration (**Fig. 3B**). Furthermore, increasing the population size discrepancy between the center and the edge does not consistently lead to an increase

407 in central estimates of $\psi$ (*Fig. S6*), indicating that asymmetric migration, not increased
408 drift along the edges, is the dominant force shaping the observed dynamics.
409

410 **4 DISCUSSION**

411 The persistence of range edges despite the potential for local adaptation is an enduring
412 evolutionary mystery. Several hypotheses have been proposed to explain the finite
413 nature of species ranges, principle among these being gene swamping (Mayr 1963;
414 Kirkpatrick & Barton 1996) and dispersal limitation (Sexton et al. 2009; Hargraves et al.
415 2014). Each of these are expected to generate a pattern of an abundant range center in
416 which fitness, population density, migration, or some combination of the three, decline
417 as a function of distance from the center.

418 In this study, we investigated the range evolution of shovel bugs (*Haustorius*
419 *canadensis*; Bousfield 1962) distributed along the Eastern Seaboard of North America.
420 Past work has suggested that shovel bugs have the potential for large amounts of
421 cryptic diversity and are sensitive to environmental barriers (Jones 2005; Hancock et al.
422 2019), so our first aim was to establish that shovel bugs in the North Atlantic represent a
423 single, continuous population. Given that patterns of linearized $F_{ST}$ were consistent with
424 a continuous population and the triangle plot analyses did not reveal a strong
425 relationship with interclass heterozygosity and the hybrid index, we tentatively conclude
426 that shovel bugs in the North Atlantic represent a single, continuous population. We
427 identify reduced effective migration on the range periphery consistent with the dispersal
428 limitation hypothesis (*Fig. S5*); however, we also find evidence for resistance to
429 migration across the Chesapeake Bay. Shovel bugs of the genus *Haustorius* specialize

430 on open, wave-exposed beaches in which the sand is highly oxygenated (Bousfield
431 1970), and rarely occur in low salinity, brackish bays. While the Chesapeake Bay was
432 formed by an impact crater during the Eocene (Ivany et al. 2000), and this event has
433 been postulated to have driven haustoriid diversification in North America (Hancock et
434 al. 2022), its recent form as a large bay is young, filling in after the end of the Last
435 Glacial Maximum around 10,000 years ago (Colman & Mixon 1988). The bay has the
436 potential to be acting as a partial barrier to gene flow, and there is some support for this
437 from the PCA and *fastSTRUCTURE* analyses. However, the triangle plot and isolation-
438 by-distance regressions did not support discrete breaks (**Fig. 1C**; *Fig. S3*). This might
439 indicate that the barrier is weak or has occurred recently enough that it has not yet
440 influenced genome-wide patterns of allele frequency divergence, which can lag behind
441 other population summary statistics (Barton et al. 2013; Ringbauer et al. 2018).
442 Once the genetic structure of shovel bugs was characterized, our next aim was
443 to identify patterns of genetic diversity and inbreeding to determine if they were
444 consistent with the abundant center hypothesis. The abundant center hypothesis has
445 mostly been tested in terrestrial organisms; in the few instances in which it has been
446 tested in marine and coastal taxa, it has had mixed support. Rivadeneira et al. (2010)
447 found two of the five porcelain crab (Porcellanidae) species analyzed along the Chilean
448 coastline showed patterns of an abundant center, while the other three either showed
449 no pattern or greater diversity along an edge. Samis & Eckhert (2007), evaluating the
450 abundance and production of two species of dune flower, found no evidence for an
451 abundant center in one (pink sand verbena) while the other (beach evening primrose)
452 had the highest seed production in the center of the range. In this study, patterns of

genetic diversity and inbreeding strongly support the hypothesis of an abundant range center (**Fig. 2**). We also observed higher densities of shovel bugs in the Carolinas and southern New Jersey, and ease of collection decreased towards the range periphery. Future work would benefit from quantitative estimates of census abundance across this shovel bug's range. In the northern part of the range, the decline in density corresponded with increasingly rocky beaches and, eventually, the replacement of beaches entirely with sheer cliffs sculpted by the retreating Laurentide Ice Sheet. In their southernmost range edge, however, suitable sandy beach habitat exists far beyond their occurrence (LeCroy 2002; Hancock & Wicksten 2018). Hargraves et al. (2013) propose that when the range limit is less than the niche limit, it is due either to the recency of niche expansion relative to dispersal potential, the difference in the rate of local extinction to migration, or fluctuations in suitability that exceed the rate of recolonization. Each of these explanations rely on reductions in dispersal on the range edge relative to the center, which our *FEEMS* analysis supports (*Fig. S5*).

If rates of recolonization of peripheral populations are less than that of stochastic extinction due to environmental or demographic fluctuations, edge populations might harbor signatures of repeated range expansions. Given that we found evidence of both an abundant center and reduced migration along the periphery, we sought to test for skews in the SFS consistent with an expansion (Peter & Slatkin 2013; 2015). However, skews in the SFS were in opposite directions of what we expected; we detected a linear increase in $\psi$ towards the range center from each edge, which implies a range expansion from the periphery towards the center. This is contradictory to patterns of inbreeding and genetic diversity, which clearly indicate that abundance is the highest in

the range center (**Fig. 2**). A possible explanation to these conflicting patterns is that the $\psi$ index is biased in the presence of asymmetric dispersal.

As discussed in the Introduction, the dominant current regimes in the North Atlantic flow in opposite directions, converging at a central latitude in the Chesapeake Bay region (**Fig. 1A**). If dispersal is mostly passive in shovel bugs, we might expect a net direction of movement towards the range center from the edges. Thus, asymmetric dispersal may be responsible for generating an abundant center as it has the potential to be acting as a population sink. In addition, because migrants are being pulled towards the center, this could affect the shape of the SFS such that the signal is consistent with recurrent colonization in the direction of the Chesapeake.

To evaluate whether our intuition about the impact of asymmetric migration on the $\psi$ index was correct, we conducted simulations intended to tease apart the impacts of differences in population size and migration (**Fig. 3B**; *Fig. S6*). Consistent with our explanation above, we find that $\psi$ is highly sensitive to asymmetric migration and will infer a population expansion in the direction of the range center even in equilibrium scenarios. In addition, this bias is strongest when under the combined influence of asymmetric migration and abundant center dynamics (**Fig. 3B**), but, importantly, does not occur in the presence of differences in population size alone (*Fig. S6*). This suggests that asymmetric migration may be playing the dominant role in shaping patterns in the SFS that are being identified by the $\psi$ index.

The $\psi$ index is a popular statistic for testing for range expansions and has been applied to a diverse array of different organisms (Manuel et al. 2016; Zhan et al. 2014; Fifer et al. 2022). In this study, we uncover a potential confounding influence on $\psi$ that

is likely to impact any organism with high asymmetric migration. While asymmetric gene flow has been identified in terrestrial organisms (Fedorka et al. 2012; Oswald et al. 2017), the taxa most likely to display directional dispersal are aquatic organisms with passive dispersal via river flow or oceanic currents (Siegal et al. 2003; Pringle & Wares 2007; Morrissey & de Kerckhove 2009). Past work has shown that directional dispersal in the ocean can mislead many population genetic analyses, including the inference of random dispersal via patterns of isolation-by-distance and the probability of coalescence (Nagylaki 1977; Lenormand 2002; Paz-Vinas et al. 2013). Using a simulation approach, we uncover that asymmetric migration can also bias the detection of a range expansion, potentially generating a false positive signature when, in fact, the population is in demographic equilibrium.

In addition to the impact asymmetric migration has on the inference of a range expansion, certain biased dispersal patterns may themselves be capable of generating an abundant center. Specifically, asymmetric migration from the edges towards the range center can maintain high abundance in the center, causing it to act like a population sink, while colonization towards increasingly distant patches becomes less likely as the number of available migrants dispersing away from the center becomes increasingly small. This is a kind of migration limitation (Sexton et al. 2009; Hargraves et al. 2013), in which the rate of extinction in peripheral patches exceeds that of colonization, and the cause of this decreased migration is asymmetries in the probability of migration. This hypothesis is in opposition to gene swamping, in which increased rates of migration to the edge is detrimental (Kirkpatrick & Barton 1996).

521 The generality of the abundant center hypothesis remains hotly contested (Dallas
522 et al. 2020; Santini et al. 2019; Sagarin et al. 2006), as do the evolutionary processes
523 that give rise to them (Kottler et al. 2023). We uncover strong evidence for an abundant
524 center in shovel bugs in the western North Atlantic and strong support for the
525 hypothesis that this pattern is maintained by a combination of passive dispersal and
526 directional currents promoting gene flow towards the center of the range. We suggest
527 this mechanism for generating an abundant center might be widespread in coastal
528 marine taxa, especially those with passive dispersal, as oceanic currents regularly
529 converge within species ranges. Furthermore, focusing on asymmetric migration via
530 directional water or wind currents may help clarify what maintains range edges in cases
531 where the fundamental niche appears to exceed the range limit.
532

533 **DATA AVAILABILITY**

534 Bioinformatic pipelines, including filtered VCFs, are available at
535 https://github.com/zachbhancock/asym_gene_flow_proj. Raw sequences were
536 deposited to GENBANK.
537

538 **AUTHOR CONTRIBUTIONS**

539 ZBH and GSB conceptualized and designed the study. ZBH carried out the field
540 collections. ZBH and PLW performed the laboratory work. ZBH and NEA carried out the
541 bioinformatics and analyzed the data. All authors wrote and corrected the manuscript.
542

543 **ACKNOWLEDGEMENTS**

Thanks to members of the Bradburd lab who provided helpful comments on this manuscript. Thanks also to Kristen Wacker and Teresa Pegan for thoughtful discussion and feedback. We are all indebted to Leonard Jones who coined the common name “shovel bug” for haustoriid amphipods, and proudly use it in print for the 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.

## REFERENCES

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

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.

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.

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.

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.

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.

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.

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.

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.

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

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.

823 **FIGURE LEGENDS**

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

832

833 **Figure 2**. *Abundant center results*. A) Individual genetic diversity (π) as a function of
834 latitude. B) Inbreeding coefficient as a function of the distance from the center of the
835 range, taken as 36° latitude, in miles. Points are colored by the latitude of their sample
836 location as in **Fig. 1A**.

837

838 **Figure 3**. *Directionality index ($\psi$) results*. A) Empirical estimates of $\psi$, points are the
839 summed values of $\psi$ for all comparisons of population *i* to populations *j*. B) Simulation
840 results for each stepping-stone model: "center_asym" is abundant center with
841 asymmetric migration; "center_sym" is abundant center with symmetric migration;
842 "equal_asym" is equal deme size with asymmetric migration; and "equal_sym" is equal
843 deme size with symmetric migration. Each point per population is a different replicate.
844 Points are colored by estimated genetic diversity (π) of the sampled locality. Dashed
845 line is when $\psi = 0$, which is expected under equilibrium isolation-by-distance.

## FIGURES

**Figure 1.** *Sample locations and characterization of population genetic structure*. 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.

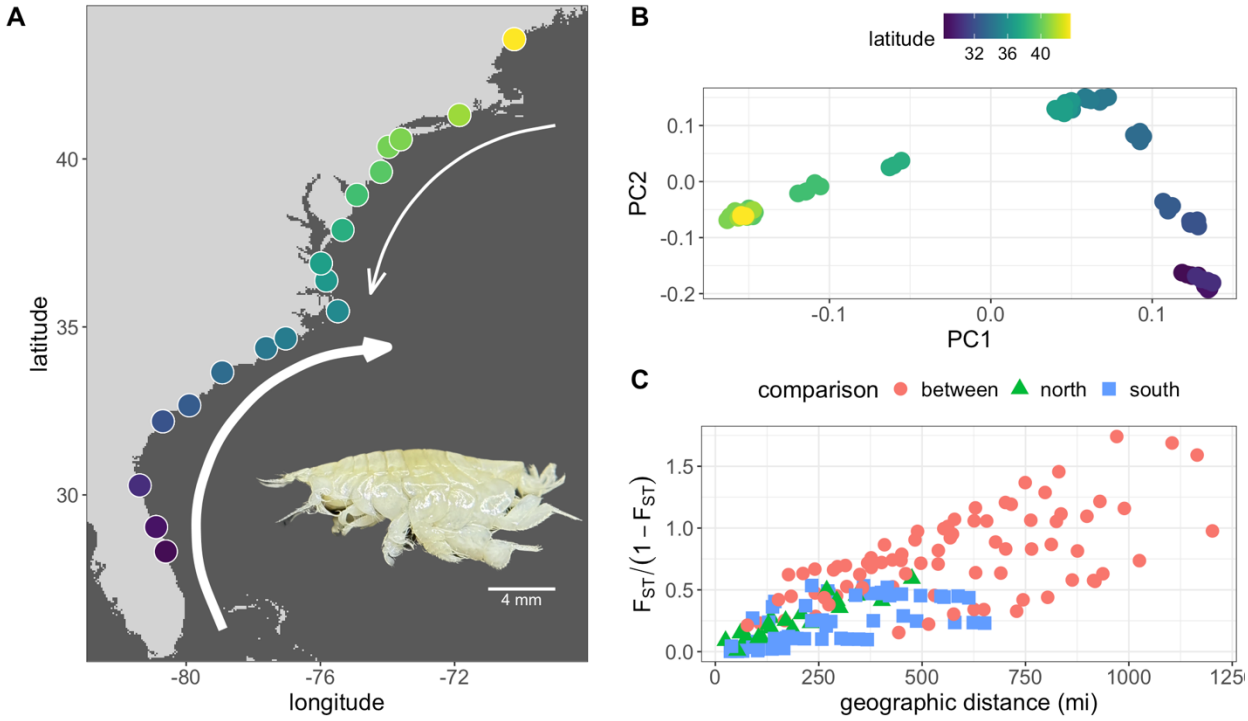

**Figure 2**. *Abundant center results*. A) Individual genetic diversity (π) as a function of latitude. B) Inbreeding coefficient as a function of the distance from the center of the range, taken as 36° latitude, in miles. Points are colored by the latitude of their sample location as in **Fig. 1A**.

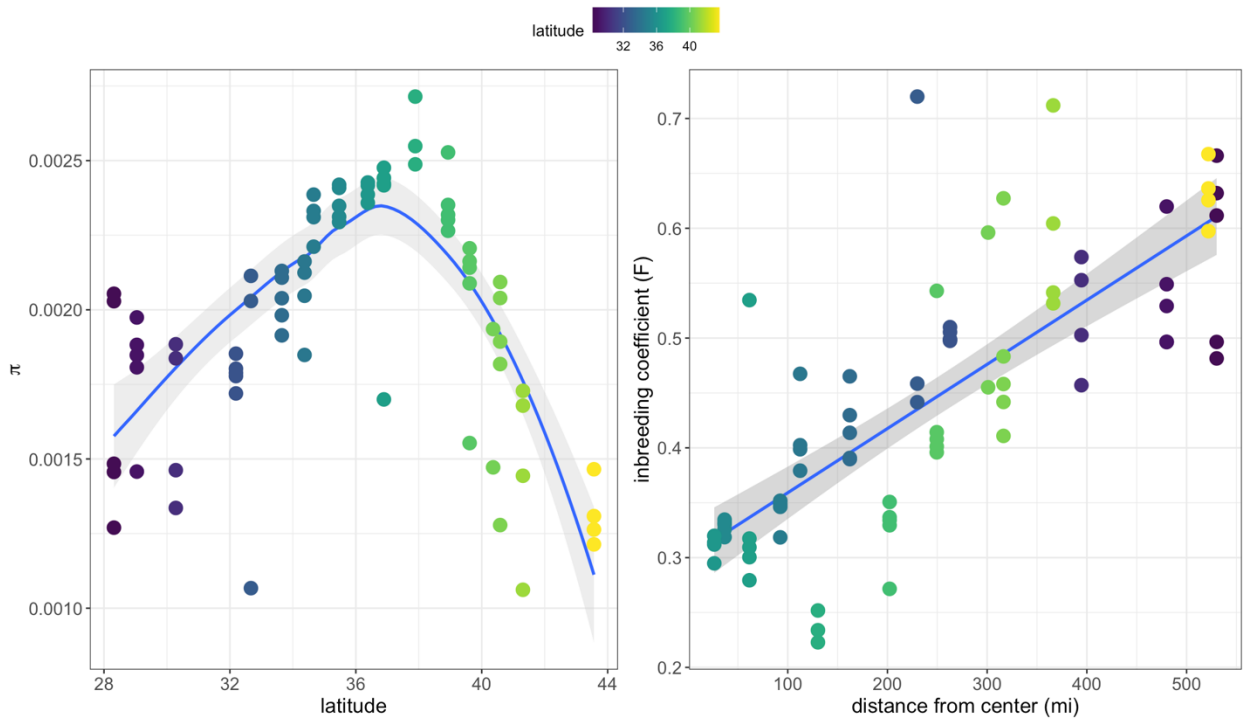
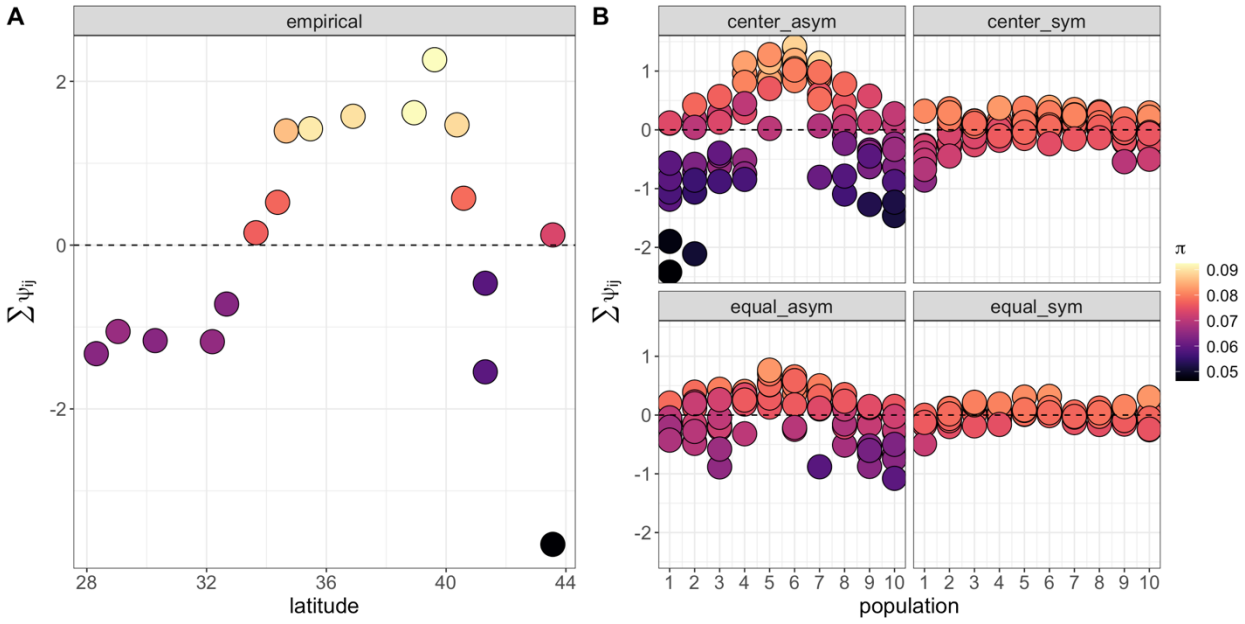
**Figure 3**. *Directionality index ($\psi$) results*. A) Empirical estimates of $\psi$, points are the summed values of $\psi$ for all comparisons of population *i* to populations *j*. B) Simulation results for each stepping-stone model: “center_asym” is abundant center with asymmetric migration; “center_sym” 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 $\psi$ = 0, which is expected under equilibrium isolation-by-distance.