Beyond a single patch: local and regional processes explain biodiversity patterns in a seagrass epifaunal metacommunity

Keila A Stark[1]*, Patrick L Thompson[1], Jennifer Yakimishyn[2], Lynn Lee[3], Mary I O'Connor[1]

1. Department of Zoology
Biodiversity Research Centre
University of British Columbia
V6T 1Z4
Vancouver, BC, Canada

2. Pacific Rim National Park Reserve
P.O. Box 280
Ucluelet, BC, Canada

3. Gwaii Haanas National Park Reserve, National Marine Conservation Area Reserve, and Haida Heritage Site
60 Second Beach Road
Skidegate, BC, Canada

***Correponding author:** keilastark@zoology.ubc.ca

Abstract: 260 words
Main text: 5167 words
Number of references: 57
Number of figures: 5
Number of tables: 1

Data accessibility: Data will be made available on Dryad should the manuscript be accepted. All code for analysis can be found at https://github.com/keilast/HMS-Seagrass.

Competing interests: We have no competing interests.

Abstract

Ecological communities are structured by dispersal, environmental niche filtering, and biotic interactions. Metacommunity ecology provides a framework for understanding how these processes combine to determine community composition in sites that are connected through the dispersal of multiple potentially interacting populations. Previous approaches for describing metacommunity dynamics from observational data have had limited success due to statistical and theoretical limitations. We used the Hierarchical Modelling of Species Communities (HMSC) statistical framework to assess the contributions of these metacommunity processes to faunal community structure in an important coastal habitat, seagrass meadows, across approximately 1000 km along the coast of British Columbia. Following the HMSC framework, we parametrized a hierarchical joint species distribution model describing the relationship between environmental variables, spatial distance between sites, species co-occurrence patterns, and random effects on epifaunal species distributions across thirteen eelgrass (*Zostera marina* (L.)) meadows. Species distribution patterns suggested that sites are well-connected by dispersal, and that dispersal rates are not so high that they homogenize regional diversity. We found evidence of niche filtering in several species along an abiotic gradient; however, environmental conditions only explained a small fraction of the overall variation in species distributions (22%), suggesting variable niche width among species. Strong structuring in pairwise species co-occurrences after controlling for environmental and spatial covariates led us to hypothesize the importance of priority effects and trophic interactions. This work has yielded refined hypotheses about the specific local and regional dynamics influencing seagrass epifaunal community structure; these hypotheses in turn can inform targeted experiments, monitoring efforts, and management decisions for preserving diversity in this important coastal ecosystem.

Key words: Hierarchical Modelling of Species Communities, priority effects, dispersal, niche filtering, British Columbia, *Zostera marina* (L.)

Introduction

Understanding how local environmental conditions, biotic interactions, and regional connectivity jointly structure the composition of communities is a central challenge in ecology (Ricklefs and Schluter 2003, Vellend 2010, 2016, Leibold and Chase 2017). Determining the relative importance of these assembly processes in particular ecosystems is a critical step in understanding the spatial and temporal scales of stability and resilience in spatially structured communities (Economo et al. 2011). Metacommunity theory offers a promising framework for understanding drivers of community assembly across spatial scales; however, previous attempts to quantitatively apply metacommunity concepts to real-world systems have had limited success (Logue et al. 2011, Brown et al. 2017). Over the past decade and a half, the primary focus of metacommunity studies has been seeking in natural systems signatures of one of four idealized paradigms: the neutral model, patch-dynamics, species sorting, and mass effects (Leibold et al. 2004, Cottenie 2005). An emerging perspective holds that natural metacommunities often do not fit neatly under these paradigms because they do not reflect all possible combinations of the various assembly processes (Logue et al. 2011, Brown et al. 2017, Leibold and Chase 2017). Thus, the revised goal is to determine how dispersal, environmental conditions, and biotic interactions combine to structure species composition and relative abundance within a metacommunity.

Many marine ecosystems that host high biodiversity occupy spatially structured habitats (e.g. rocky reefs, coral reefs, seagrass meadows), and marine species have dispersing life histories that

connect populations in distinct habitat patches. Marine systems are characterized by strong local species interactions (Berlow 1999, Sala and Graham 2002), and the importance of 'supply-side' or dispersal-driven population dynamics has long been recognized (Levin and Paine 1974, Gaines and Roughgarden 1985, Lee and Bruno 2009). These combined roles of dispersal, the local environment, and biotic interactions suggest that marine biodiversity patterns may reflect processes characteristic of metacommunities (Guichard et al. 2003). Despite this, metacommunity dynamics have only been explicitly investigated in a few marine systems (Guichard et al. 2003, Yamada et al. 2014, Valanko et al. 2015, Whippo et al. 2018).

Seagrasses, like kelps, corals, and salt marsh plants, are foundation species that support high productivity and faunal diversity (Orth et al. 1984, Duffy et al. 2015). Seagrasses form meadows separated from each other by deeper water, un-vegetated seafloor, or other vegetated habitats such as kelp forests, and associated fauna disperse among meadows (Boström et al. 2006). This scenario is consistent with the type of landscape that is governed by metacommunity dynamics (Mouquet and Loreau 2003, Leibold et al. 2004). Yet, our understanding of the causes and conservation implications of seagrass-associated faunal diversity patterns has not accounted for processes operating across numerous meadows and adjacent habitats in a seascape. Local characteristics such as habitat structure (Orth et al. 1984, Heck and Valentine 2006), primary biomass (Cébrian and Duarte 1998, Gullström et al. 2012), or trophic interactions (Amunudrud et al. 2015, Huang et al. 2015) affect faunal species diversity, but the importance of ecological processes operating on a regional or inter-site scale remains unclear. Our limited understanding of the importance of these regional processes could lead to underestimates of the amount of habitat required to sustain biodiversity and ecosystem function (Boström et al 2006).

To date, evidence for the importance of metacommunity dyamics and regional-scale

114 processes in seagrass associated communities is based on two different quantitative
115 metacommunity approaches (Yamada et al. 2014, Whippo et al. 2018). Inferences were based on
116 statistical methods available at the time that assessed the contribution of environment and space,
117 but were not able to identify possible biotic interactions and were therefore did not account for
118 this critical community assembly process. Using variation partitioning on constrained ordinations
119 (Borcard et al. 1992, Cottenie 2005), Yamada et al. (2014) found that the relative contribution of
120 space and environment varied among invertebrate functional groups. However, the variation
121 partitioning analysis was not able to distinguish between the possible metacommunity dynamics
122 at play owing to the ambiguous statistical significance of the results (Yamada et al. 2014).
123 Simulation studies have shown that this method can over- or under-estimate spatial and
124 environmental components of variation, and produce inflated R^2 values depending on the
125 researcher's statistical choices (Gilbert and Bennett 2010, Tuomisto et al. 2012). The other
126 approach applied to seagrass epifaunal systems, known as elements of metacommunity structure
127 (EMS), seeks idealized metacommunity structures in species occurrence patterns (Leibold and
128 Mikkelson 2002, Presley et al. 2010, Whippo et al. 2018). While useful for visualizing regional
129 diversity patterns, EMS does not robustly elucidate the processes contributing to these patterns
130 (Brown et al. 2017).
131 These limitations, along with the revised goal of individually assessing the contribution of
132 dispersal, niche filtering, and biotic interactions to diversity patterns, invite novel methodological
133 approaches. A new statistical framework, Hierarchical Modelling of Species Communities
134 (HMSC) (Ovaskainen et al. 2017), shows promise for achieving this goal. HMSC uses Bayesian
135 inference to parametrize a hierarchical joint species distribution model summarizing the
136 influence of space, environment, phylogeny, and traits on species distributions. The framework

allows the partitioning of variation in species distributions due to the environment, spatial distance, and random effects for all species distributions, allowing the observer to discern patterns at the whole-community level as well as in individual species. Additionally, it specifies a way to visualize patterns of positive or negative pairwise species co-occurences. For these reasons, HMSC achieves what researchers aimed to achieve with the previous methods, and much more.

Here we test the hypothesis that seagrass-associated epifaunal invertebrate communities are jointly structured by the spatial distances between meadows (a proxy for inferring dispersal dynamics), local environmental conditions, and biotic interactions. We follow the HMSC framework to answer the following questions: (1) Is the metacommunity characterized by dispersal limitation, intermediate dispersal, or high dispersal? (2) Do species distribution patterns suggest environmental niche filtering, and if so along which environmental axes? (3) Do species co-variance patterns suggest possible biotic interactions influencing community assembly? The order in which these questions are posed is significant. For example, if evidence for dispersal limitation is found, this knowledge affects how we may interpret observed species co-occurrence patterns. We address these three questions with a biodiversity survey of invertebrates in 13 *Zostera marina* (L.) meadows in British Columbia, Canada. We find evidence of intermediate dispersal rates among meadows across 1000 km, but no sign of complete regional homogenization resulting from high dispersal. We also find that local environmental conditions have some influence on species distributions across the 13 sites, but that individual species' responses to environmental gradients vary dramatically among taxa. Strong species' covariance patterns led us to hypothesize possible biotic interactions influencing local community composition.

Methods

Study sites and field sampling of eelgrass and associated epifauna

We sampled epifaunal biodiversity in 13 eelgrass *Zostera marina* (L.) meadows across a spatial scale of 1000 km from late June through early August in 2017. The sites were located within four coastal areas: Haida Gwaii within Gwaii Haanas National Park Reserve, National Marine Conservation Area Reserve and Haida Heritage Site; southern Clayoquot Sound (one site was within Pacific Rim National Park Reserve); Barkley Sound on the West coast of Vancouver Island; and the Southern Gulf Islands on the south-eastern tip of Vancouver Island within Gulf Islands National Park Reserve (Fig. 1).

We sampled three sites in Barkley Sound (SA, DC, RB) with SCUBA because they required boat access, and the deepest parts of the meadow reached a depth of 5 m. All other sites were sampled by wading or snorkeling at a low tide of 0.6 m or less. We used a standardized, nested sampling regime with 0.25 m x 0.25 m quadrats to collect aboveground eelgrass, epifaunal invertebrates, detritus, and macroalgae (Whippo et al. 2018). Six quadrats were set in a 15 m x 30 m array, such that every quadrat was situated 15 metres from its nearest neighbour (Appendix S1: Fig. 1). The array was placed in the middle of the eelgrass meadow to avoid edge effects where possible. We removed all above-ground biomass (eelgrass, detritus, macroalgae) within each quadrat by hand and immediately placed the contents into a mesh or ziploc bag.

Acquiring biotic and abiotic environmental variables

We acquired water quality and biotic data associated with hypothesized niche differences among epifauna. The water quality variables included in-situ measurements of pH, salinity, dissolved oxygen, nitrates and chlorophyll-*a* at each meadow, and averages acquired from Bio-

183 ORACLE (Assis et al. 2018), including surface chlorophyll-*a*, dissolved oxygen, surface nitrates,
184 surface silicates, surface phosphates, pH, salinity, and sea surface temperature. Bio-ORACLE is
185 a database of marine environmental variables gathered from several satellite and in-situ sources
186 at a spatial resolution of 0.08° (Assis et al. 2018). All sites (meadows) were situated in different
187 spatial cells, therefore all sites had distinct values for the aforementioned parameters. We also
188 quantified fetch, which is a proxy for site exposure.

189 For each quadrat sampled, we measured five biotic attributes summarizing habitat
190 structure and food availability: eelgrass leaf area index (LAI) and shoot density, eelgrass dry
191 mass, total algal dry mass, and detritus dry mass. To quantify LAI, we measured leaf length,
192 width, and number of blades in five haphazardly chosen shoots within each quadrat. Leaf area
193 index was calculated by multiplying the average blade area (m^2) per shoot by the shoot density
194 (number of shoots per m^2). We separated and dried eelgrass, detritus, and macroalgae in a
195 desiccator oven (60° C for 48 hours) to measure dry mass.

196

197 *Quantifying epifaunal diversity*

198 We rinsed the eelgrass blades from each quadrat to remove epifaunal invertebrates, and
199 passed samples through a 500 μm sieve. Invertebrates were transferred into a centrifuge tube and
200 preserved in 95% EtOH for identification with light microscopy (10 times magnification). We
201 identified invertebrates to the lowest taxonomic level possible using keys in the Light and Smith
202 Manual (Carlton 2007) and Kozloff (1996).

203

204 *Analysis*

205 We followed the methodology of Ovaskainen et al. (2017) to construct a hierarchical

206 joint species distribution model describing the relationship between measured environmental
207 variables, spatial distance, species co-occurrences and species distributions using the HMSC
208 package (Guillaume Blanchet et al. 2018) in R (version 3.4.2; R Development Core Team 2014).
209 The model is comprosed of two linear predictor terms: one describing fixed effects of
210 environment (nine water quality and five biotic variables) and space, and one describing random
211 effects that cannot be attributed to environment or space (Equations 2 & 3 in Ovaskainen et al.
212 2017). The random-effects linear predictor term describes random variation in individual species
213 distributions as well as covariation for every species pair in the analysis.

214 We used Markov chain Monte Carlo to estimate the model parameters. Species'
215 responses to environmental variables were assumed to follow a multivariate Gaussian
216 distribution. We assumed a Poisson distribution for species count data and used the same priors
217 as in the examples presented in Ovaskainen et al. (2017). The Markov Chain included 100 000
218 iterations, a burn-in length of 1000 iterations, and thinning of 100. We confirmed that parameters
219 were well mixed by visually inspecting trace plots. We tested the model's performance by
220 calculating explained deviance (D^2_{adj}) as described in Appendix S1: Eq. S1. D^2_{adj} is comparable
221 to R^2_{adj} in least square regression analysis (Guisan and Zimmermann 2000); in this context, it
222 describes the model fit where a D^2_{adj} value of 1 means the model perfectly predicts the observed
223 data. Our calculations yielded a D^2_{adj} value for each individual species distribution model, and
224 the mean D^2_{adj} across all species gave a single metacommunity-level value of the model's
225 predictive power.

226 To answer our three main questions about the roles of metacommunity processes, we
227 visualized the parametrized model in several ways (Table 1). We investigated three possible
228 dispersal conditions: dispersal limitation, intermediate dispersal, and dispersal surplus. To

229 investigate dispersal limitation at large spatial scales, we plotted distance-decay of community
230 similarity. To do so, we randomly drew 200 pairwise distances between sampling plots, and
231 plotted predicted community similarity against distance. Distance-decay of community similarity
232 could reflect dispersal limitation due to isolation by distance, or it could reflect differences in
233 environmental conditions with increasing spatial scale. To distinguish these possibilities, we
234 compared predicted species compositional similarity as a function of both distance and
235 environmental covariates (the full HSMC model), and predicted compositional similarity using
236 an alternative model in which the effect of the environment was removed (all environmental
237 covariates were set to their mean value across all sites; Equation 5 in Ovaskainen et al. 2017). To
238 distinguish between dispersal sufficiency (intermediate) and dispersal surplus (high), we visually
239 inspected our site-by-species matrix for signatures of regional homogenization by high dispersal.
240 We used the degree of environmental niche filtering (as estimated by D^2_{adj}) as another line of
241 evidence to distinguish intermediate and high dispersal: theory predicts that intermediate
242 dispersal allows species to colonize sites where local environmental conditions are optimal for
243 growth and reproduction, whereas high dispersal overrides this (Mouquet and Loreau 2003).
244 To estimate the strength of niche filtering, we conducted variation partitioning analysis of
245 the fixed and random effects as defined in Ovaskainen et al. (2017). This summarized the
246 relative contribution of all environmental variables, spatial distance, and random effects to the
247 distribution of each species in the model; we visualized this in a stacked bar plot. We then
248 estimated the pure and shared fractions of variation explained by environment and space, as well
249 as the unexplained fraction using type III sum of squares (Borcard et al. 1992), but using the
250 HMSC model as opposed to constrained ordination. To do so, we followed the steps and
251 equations in Appendix S1: Eq. 2.

To identify species co-occurrence patterns resulting from possible biotic interactions, we represented modelled pairwise associations in a correlation matrix describing the extent to which species pairs co-occurred more negatively or positively than by random chance, controlling for the effect of the environment. We then conducted hierarchical cluster analysis on sites according to predicted species abundances from the HMSC model using Ward's criterion (Ward 1963). We mapped the cluster dendrogram onto an abundance heatmap of the 20 species in our study with the highest mean abundance, and colour coded the cells according to the species co-occurrence groupings (groups of species that tended to positively co-occur amongst themselves). Full versions of these figures with all species are provided in Appendix S1: Figs. S4 and S5.

Results

Species abundance and distribution patterns

We counted and identified 39 644 individuals representing at least 58 species across the region (Appendix 1: Table S1). The represented taxa span multiple phyla, diet types (herbivores, detritivores, suspension feeders, and carnivores), and life history strategies (brooding, broadcast spawning). The mean plot or quadrat-level species richness was approximately 13 species, and mean meadow or site-level richness was approximately 16 species. A full summary of traditional diversity metrics can be found in Appendix S1: Table S2.

Several taxa were present at all sites (Appendix S1: Fig. S2). These include harpacticoid copepod *Porcellidium* sp., bivalves *Saxidomus gigantea* and *Clinocardium nutallii*, snails *Alvania compacta*, *Lacuna* spp., and the tanaid *Leptochelia* sp. Genera *Mytilus* (a mussel) and *Nereis* (a polychaete worm) were also represented at all sites, however we were unable to confidently identify these organisms to species. Ten species were present in at least one site in

every region: pycnogonid *Anoplodactylus viridintestinalis,* isopod *Idotea resecata,* gammarid amphipods *Monocorophium insidiosum, Ampithoe valida, Pontogeneia rostrata, Ampithoe dalli, Aoroides* spp*.,* snails *Amphissa columbiana* and *Alia carinata,* and limpet *Lottia pelta*.

Dispersal: Distance-decay of community similarity due to spatial distance and environment

In the full HMSC model summarizing species' responses to environment and distance between sites, predicted community similarity decreased with increasing distance between meadows (Fig. 2, red points and line). However, community similarity showed a slight increase with distance when the effect of environmental variables was removed from the model (all environmental variables set to their mean value), suggesting that spatial distance alone did not correspond with differences in invertebrate compositional similarity (Fig. 2, blue points and line).

Environmental niche filtering: variation partitioning

Patterns in abundance and distribution were consistent with weak environmental niche filtering along an abiotic axis; however, the importance of environmental conditions varied across species. The relative contributions of all 14 environmental variables, spatial distance and random effects in the HMSC model can be seen in Fig. 3. Out of all the measured environmental variables, the water quality variables (all blue bars in Fig. 3) explained the largest proportion of variation across nearly all species distributions. Across all taxa, herbivore food availability (seagrass, algae and detritus biomass) and habitat structure (eelgrass shoot density and LAI) had smaller influences on species distributions than abiotic predictors (Fig. 3). The relative contribution of abiotic parameters to species abundance and distribution patterns also varied

298 among species and major taxonomic groups. While the mean deviance explained (D^2_{adj}) across
299 all species distribution models was 0.29, individual values varied dramatically (0 to 0.82,
300 Appendix 1: Table S1), implying high variation in the match between our model and the
301 observed species distributions. Accounting for this, the true fraction of variation explained by
302 environment was 22.4%, space explained -3%, the shared fraction between environment and
303 space was 4%, and the residual fraction of variation was 0.76 (Appendix S1: Fig. S3).
304
305 *Species interactions: Patterns in pairwise species co-occurrences*
306 We identified, post hoc, three main site-level species co-occurrence groupings after
307 controlling for effects of environmental covariates. The first group of species indicated by purple
308 cells (*Nereis* sp. through *Caprella laeviuscula* along the vertical axis in Fig. 4a), henceforth
309 known as the *Nereis* group, tended to positively co-occur with each other, and negatively co-
310 occur with the second group. This second group of species indicated by blue cells, henceforth
311 known as the *C. californica* group (*Caprella californica* through *Ampithoe valida,* and
312 Harpacticoid 2 through *Lacuna* sp. along the vertical axis of Fig. 4a), positively co-occurred with
313 members of the same group and negatively co-occurred with members of the *Nereis* group.
314 *Alvania compacta* and *Lacuna* sp. negatively co-occurred with members of the *C. californica*
315 group and positively co-occurred with members of the *Nereis* group, however their co-
316 occurrence values tended to be weaker. The third co-occurrence grouping was comprised of
317 species that did not show strong positive or negative co-occurrence with other species. *Spirorbis*
318 sp. was the only species in this category among the 20 species with highest mean abundances.
319 Species representing each of the three major co-occurrence groupings were present at all
320 sites (Fig. 4b), indicating that members from these groups can be simultaneously present.

321 However, the abundances of members from the different co-occurrence groups tended to
322 negatively co-vary; when species from one group were abundant, species from the others were
323 less abundant (Fig. 4b). At sites JB, HL, CB, RA, EB, DC, and DK, members of the *Nereis*
324 group were predicted to be most abundant, whereas *C. californica* group members were
325 predicted to be most abundant at SS, SA, LH, GB and RB. The site IN was predicted to be
326 strongly dominated by *Spirorbis* sp. in. The emergent species assemblages identified in Fig. 4a
327 from the HMSC modeled distributions do not clearly correspond with the spatial distance
328 because the cluster analysis did not group sites by spatial proximity or regional membership
329 (black symbols in Fig. 4b).
330
331 **Discussion**
332 By applying a new statistical framework, we have gained new insight into the possible
333 metacommunity processes structuring seagrass-associated biodiversity in coastal British
334 Columbia. We report evidence that both biotic interactions and environmental conditions
335 influence seagrass meadow invertebrate composition. Spatial patterns of species abundance and
336 distribution suggest regional-scale dispersal patterns are possible, such that even distant
337 meadows may host the same species. However, these dispersal rates are clearly not so high as to
338 overwhelm the local meadow scale effects of biotic interactions and environmental conditions.
339 Below, we review these conclusions and the evidence to support them.
340
341 *Epifaunal communities may be well-connected and experience intermediate dispersal rates*
342 Seagrass epifaunal communities appear to be well-connected by dispersal at the present
343 spatial scale and distribution of seagrass habitat along the British Columbia coast. This is true for

344 the seven species that were present at all sites, and the additional ten species that were not
345 present at every site but present in all regions (Appendix S1: Fig. S2). It is noteworthy that these
346 widely distributed groups had representation from across phyla and life history traits, suggesting
347 that no single taxonomic group or dispersal life history strategy had a consistently greater spatial
348 distribution among the sites sampled. Many of the meadows sampled were small (< 1 ha), and
349 epifaunal abundances can vary substantially within meadows from year to year such that local
350 extinctions followed by rescue from nearby populations are likely important for diversity and
351 composition (Thom et al. 1995). However, the specific pathway the organisms travelled to arrive
352 at these sites remains unknown. We used Euclidean distances in our analysis, which preserved
353 the rank order of distances across sites, but actual distances are likely greater due to
354 oceanographic circulation patterns (Kinlan and Gaines 2003, Mitarai et al. 2008, Treml et al.
355 2015).

356 It is unclear whether distant sites share the same species because they are (1) linked by
357 direct dispersal; (2) indirectly linked by dispersal via unsampled "stepping-stone" sites; or (3)
358 were colonized by populations of the same species in a historical dispersal event, but have not
359 seen the exchange of individuals since. Seagrass-associated epifauna generally disperse passively
360 either by larval transport in currents or by "rafting" on floating pieces of seagrass or macroalgae.
361 Kelp rafts have been observed hosting benthic and epifaunal invertebrates across phyla
362 (echinoderms; peracarids; molluscs; annelids), and are important in maintaining connectivity
363 between coastal ecosystems across 100s of kilometres (Wichmann et al. 2012). Thus, while
364 incidences of direct dispersal between the farthest sites may be rare, they are not impossible.

365 Species in our analysis that were only found at one site or region (e.g. *Jassa marmorata,*
366 *Neoamphitrite robusta*, Appendix S1: Fig. S2) may be dispersal limited for reasons other than

367 prohibitively long travel distances. Those with low abundances (fewer than five individuals per
368 meadow) may not disperse in appreciable numbers, and therefore cannot establish populations at
369 all sites as readily as more abundant species. This is supported by our observation of a generally
370 positive relationship between range size (number of sites a species was represented at) and
371 average abundance (Appendix 1: Fig. S6).

372 We can also conclude that dispersal rates are high enough for several species to be
373 present at most sites, but not so high as to completely overwhelm the effects of the environment
374 and biotic interactions. Thus, dispersal rates are likely intermediate on average. A subset of
375 species was present at all sites, but often at varying abundances (Fig 4b); this may be indicative
376 of weak mass effects, wherein immigration allows populations to persist even if local conditions
377 are not optimal (Schmida and Wilson 1985, Mouquet and Loreau 2003). Metacommunity theory
378 predicts that such weak mass effects are likely to be found whenever dispersal rates are not
379 limiting (Thompson et al. 2017), particularly in organisms that cannot control their own dispersal
380 (Leibold and Chase 2017) such as seagrass-associated epifauna. This may be supported by our
381 result on the contribution of environmental conditions, particularly in ubiquitous species for
382 which the model fit was poor (e.g. *Aoroides* spp., *Spirorbis* sp., $D^2_{adj} = 0$); the poor model fit
383 may indicate that these persist at sites even if environmental conditions are not optimal, due to
384 sufficient immigration. An alternative explanation for the model's poor predictive power in these
385 species is that these taxa have broader environmental niches, and thus the regional environmental
386 gradient is not sufficient to drive clear presence-absence patterns.

387

388 *Seagrass epifauna show varying degrees of environmental niche filtering*

389 The most important environmental axis informing species distributions is the abiotic one,

with phosphates, salinity, and dissolved oxygen explaining the largest fractions of variation (Fig. 3). However, after accounting for the model fit, the total variation explained by the environment was small (22%). In some species for which the model had high predictive power (e.g. *Ampithoe dalli* and *Caprella natalensis,* Appendix 1: Table S1*)*, niche filtering may be important, but overall there remains a large amount of variation unexplained by the environment.

Environmental heterogeneity may influence species distributions in one of three ways. First, environmental conditions may influence food availability through primary productivity. This is possible despite the fact that primary biomass explained less variance than the abiotic parameters in our HMSC model (Figure 3); the primary biomass measurements were taken at a single time point. Conversely, the abiotic parameters represent a four-year average (Assis et al. 2018) and may therefore better represent the growing conditions these organisms experience over their lifetimes. Second, sea surface temperature, dissolved oxygen, and salinity may affect species' distributions by via their environmental tolerance ranges. For example, estuarine isopod *Idotea baltica* (a relative of *I. resecata* and *I. montereyensis* in our study) experiences significantly lower survival in prolonged exposure to higher water temperatures and lower salinities than those experienced in their home habitat (Rugiu et al. 2018). Third, temperature may influence the metabolic demands of the epifauna (many of which are herbivores) thereby influencing grazing rates and altering the available primary biomass (O'Connor 2009). Decreased food resources may increase competition, thereby influencing the number and identity of species a seagrass patch can host. Determining whether individual species' responses to competitive stress influence whole-community structure may provide key information on how environment and species interactions interact to influence community structure.

Environmental conditions in temperate eelgrass meadows fluctuate seasonally, driving

413 changes in the relative abundance of epifaunal taxa throughout the year (Wlodarska-Kowalczuk
414 et al. 2014, Markel et al. 2017). As a result, the epifaunal metacommunity may not show these
415 distributions year-round; sampling several times in one year to capture these temporal changes
416 would give a clearer picture of species' responses to all metacommunity processes.
417
418 *Species co-variance patterns may suggest priority effects and trophic interactions*
419 Three main co-occurrence groupings emerged showing antagonistic abundances across
420 sites (Fig. 4b). Stochasticity alone is unlikely to have produced such a strong pattern, especially
421 given deterministic processes tend to be more important in systems that are not dispersal-limited
422 (Lee and Bruno 2009, Ron et al. 2018). While there are multiple possible explanations for non-
423 random species occurrence patterns (Connor and Simberloff 1979), environmental heterogeneity
424 was ruled out as the co-occurrence values represented pairwise covariation after controlling for
425 spatial and environmental covariates (Ovaskainen et al. 2017). This, along with the high
426 unexplained fraction in our variation partitioning result (Appendix S1: Fig. 3), leads us to
427 suggest the potential importance of biotic processes. Several types of interactions may be
428 occurring in the metacommunity; many species in the study are herbivores or detritivores, and
429 thus may compete for primary biomass. Laboratory experiments have demonstrated that
430 behaviours such as grazing rate and habitat selection in amphipods are altered in the presence of
431 morphologically and functionally similar intrerspecific competitors (Beermann et al. 2018).
432 Other species in our study are filter feeders, and thus may not be competing for food resources
433 but possibly predator-free space or substrate on the eelgrass blades.
434 The community structure of a given site with certain local environmental conditions and
435 access to a certain regional pool may have multiple possible compositional states depending on

436 the arrival order of species; this phenomenon is known as priority effects (Fukami et al. 2016, Ke
437 and Letten 2018) and has been documented in marine fouling communities (Vieira et al. 2018).
438 The antagonistic co-occurrence groupings in our analysis may suggest priority effects. We
439 hypothesize that following the establishment of a new meadow or a disturbance event, the first
440 few species to colonize the meadow determine the success of subsequent colonizers, regardless
441 of local environmental conditions. As an illustrative example, *Lacuna* spp. might have been
442 among the first to colonize sites JB, HL, GB, and CB (Fig. 4b), thereby gaining a numerical
443 advantage over species in the alternative cohort. Distibution patterns in the sessile calcifying
444 polychaete *Spirorbis* sp. may also indicate priority effects. While it did not strongly negatively or
445 positively co-occur with other species, it dominated the HL site; 16 300 individuals were found
446 across the six quadrats, several orders of magnitude higher than its abundance at other sites (Data
447 S1). Site HL also had the lowest species richness of all 13 sites (9 species). Eelgrass shoots there
448 were completely covered by these *Spirorbis* sp., possibly preventing grazers in the community
449 from accessing epiphytic algae or the seagrass itself. The age, colonization history, and
450 disturbance regime of the meadows likely matter if priority effects are important.
451 The negatively correlated abundance patterns could also reflect the presence of a predator
452 such as *Nereis* at some sites and not others (Fig. 4b). Field experiments have demonstrated that
453 changes in predation pressure by fish, shorebirds, and predaceous invertebrates can shift
454 seagrass-associated epifaunal assemblages in a matter of weeks (Amundrud et al. 2015; Huang et
455 al. 2015; Lewis & Anderson 2016). Differential predation pressure could also lead to the pattern
456 of diverging community composition that we have witnessed (Guzman et al. 2018). A more
457 comprehensive sampling effort of fish and bird predators would be required to confirm the
458 degree to which trophic interactions structure these communities.

459 The positive co-occurrence patterns between certain species may additionally be a result
460 of positive biotic interactions. An example is seen in *Orchomenella recondita* - a gammarid
461 amphipod that lives in the gastrovascular cavity of the anemone *Anthopleura elegans* (Carlton
462 2007). This species was only recorded at the SS site, and specifically only found in quadrats
463 where *A. elegans* was collected with the seagrass shoots.
464

465 **Conclusion**

466 Using a new statistical approach, we have arrived at refined hypotheses about the local
467 and regional metacommunity processes driving temperate seagrass epifaunal diversity. Our study
468 also gives a baseline spatial scale (1000 km) for comparing the contribution of these processes
469 across spatial scales and other ecological systems (Viana and Chase 2018). Manipulative
470 experiments and monitoring are now needed to test our refined hypotheses. One could test the
471 hypothesis that species sorting is relatively weak by measuring fitness differences along the most
472 important environmental axes. Given our hypothesis on the importance of priority effects, one
473 might experimentally manipulate the colonization order of species from the different co-
474 occurrence groupings to determine if resulting compositional states differ (Fukami et al. 2016).
475 Testing our hypothesis that populations are not dispersal limited is more challenging, but could
476 be done using particle tracking models similar to those used to estimate connectivity between
477 coral reefs via larval dispersal across thousands of kilometres (Treml et al. 2008). A population
478 genetic analysis could also ground-truth the extent and time scale of dispersal between allopatric
479 populations. All together, we suggest that a modern metacommunity perspective combined with
480 new advances in our statistical toolbox provide a powerful and relevant framework for studying
481 natural ecosystems.

Our findings are particularly relevant to conservation initiatives that aim to maintain diversity in seagrass meadows; identifying the contribution of local and regional processes allows the conservation practitioner to better predict how habitat fragmentation and climate change might influence diversity. Seagrass-associated epifaunal communities may be quite resilient, given our findings that environmental conditions make a minimal contribution to species distributions, and several species across phyla and functional groups are consistently present across a fairly large spatial scale.

Acknowledgements

KAS conceived of the study with guidance from MIO. JY provided samples from Clayoquot Sound. LL provided samples from Haida Gwaii. KAS collected and processed samples from the Gulf Islands and Barkley Sound, analyzed data, and wrote the first manuscript draft. PLT provided recommendations for the analysis. KAS, PLT, and MIO contributed to writing the manuscript. This research was financially supported by a UBC Zoology SURE Grant, and NSERC Discovery Grants, Canadian Foundation for Innovation (CFI), and is sponsored by the NSERC Canadian Healthy Oceans Network and its Partners: Department of Fisheries and Oceans Canada and INREST (representing the Port of Sept-Îles and City of Sept-Îles). PLT is supported by NSERC and Killam postdoctoral fellowships. We are grateful to Duarte Viana for providing critical recommendations and code for the D^2_{adj} and variation partitioning calculations. Matt Whalen pulled the Bio-ORACLE dataset, and John Cristiani made the map in Fig. 1. Special thanks to Michelle Paleczny and Gulf Islands National Park Reserve for boat and laboratory space during their eelgrass field season. We acknowledge staff, students, and volunteers from Pacific Rim National Park Reserve for their assistance collecting eelgrass

samples in Clayoquot Sound. Thanks goes to Rob Underhill, Abbie Sherwood, Emily Adamczyk, Minako Ito, Ariane Comeau, Giordano Bua for assistance with sample collection, processing, and ID. We thank Coreen Forbes, Matt Whalen and Jeff Groh for their comments on the manuscript drafts.

Literature Cited

Amundrud, S.L., D. S. Srivastava, and M. I. O'Connor. 2015. Indirect effects of predators control herbivore richness and abundance in a benthic eelgrass (*Zostera marina*) mesograzer community. Journal of Animal Ecology 84:1092–1102.

Assis J., L. Tyberghein L. S. Bosch, H. Verbruggen, E. A. Serrão, and O. De Clerck. 2018. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. Global Ecology and Biogeography 27:277–284.

Beermann, J., K. Boos, L. Gutow, M. Boersma, and A. C. Peralta. 2018. Combined effects of predator cues and competition define habitat choice and food consumption of amphipod mesograzers. Oecologia 186:645–654.

Berlow, E. L. 1999. Strong effects of weak interactions in ecological communities. Nature 398: 330–334.

Boström, C., E. L. Jackson, and C. A. Simenstad. 2006. Seagrass landscapes and their effects on associated fauna: A review. Estuarine, Coastal and Shelf Science 68:383-403.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the Spatial Component of Ecological Variation. Ecology. 73:1045–1055.

Brown, B.L., E. R. Sokol, J. Skelton, and B. Tornwall. 2017. Making sense of metacommunities: dispelling the mythology of a metacommunity typology. Oecologia 183:643–652.

Carlton, J. 2007. The Light and Smith Manual: Intertidal invertebrates from Southern California to Oregon. Fourth edition. University of California Press, Oakland, California, USA.

Cébrian, J. and C. M. Duarte. 1998. Pattems in leaf herbivory on seagrasses. Aquatic Botany 60:67–82.

Connor, E. F., and D. Simberloff. 1979. The assembly of species communities: Chance or competition? Ecology 60:1132–1140.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8:1175–1182.

Duffy J.E, et al. 2015. Biodiversity mediates top-down control in eelgrass ecosystems: A global comparative-experimental approach. Ecology Letters 18:696–705.

Economo, E. P. 2011. Biodiversity Conservation in Metacommunity Networks: Linking Pattern and Persistence. The American Naturalist 177:167-180.

Fukami, T., E. A. Mordecai, and A. Ostling. 2016. A framework for priority effects. Journal of Vegetation Science 27:655–657.

Gaines, S., and J. Roughgarden. 1985. Larval settlement rate: A leading determinant of structure in an ecological community of the marine intertidal zone. Proceedings of the National Academy of Sciences. 82:3707–3711.

Gilbert, B, and J. R. Bennett. 2010. Partitioning variation in ecological communities: do the numbers add up? Journal of Applied Ecology 47: 1071–1082.

Guichard, F. 2003. Mussel disturbance dynamics: Signatures of oceanographic forcing from local interactions. American Naturalist 161:889–904.

Guillaume Blanchet, F., with contributions from G.Tikhonov and A. Norberg. 2018. HMSC: Hierarchical Modelling of Species Community. R package version. 2.1-2.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 145:147–186.

Gullström, M., S. Baden, and M. Lindegarth. 2012. Spatial patterns and environmental correlates in leaf-associated epifaunal assemblages of temperate seagrass (*Zostera marina*) meadows. Marine Biology 159:413–425.

Guzman, L. M., R. M. Germain, C. Forbes, S. Straus, M. I. O'Connor, D. Gravel, D. S. Srivastava, and P. L. Thompson. 2018. Towards a multi-trophic extension of metacommunity ecology. In press. Ecology Letters doi: 10.1111/ele.13162.

Heck K.L., and J. F. Valentine. 2006. Plant-herbivore interactions in seagrass meadows. Journal of Experimental Marine Biology 330:420–436.

Huang A.C., M. Essak, and M. I. O'Connor. 2015. Top-down control by great blue herons *Ardea herodias* regulates seagrass-associated epifauna. Oikos 124:1492–1501.

Ke, P. J., and A. D. Letten. 2018. Coexistence theory and the frequency-dependence of priority effects. Nature Ecology and Evolution 2:1691–1695.

Kinlan, B.P., and S. D. Gaines. 2003. Propagule Dispersal in Marine and Terrestrial Environments: A Community Perspective. Ecology 84:2007–2020.

Kozloff, E. 1996. Marine Invertebrates of the Pacific Northwest. University of Washington Press, Seattle, Washington, USA.

Leibold, M.A., et al. 2004. The metacommunity concept: A framework for multi-scale community ecology. Ecology Letters 7:601–613.

Leibold, M. A. and J. M. Chase. 2017. Metacommunity Ecology, Volume 59. Princeton University Press, Princeton, New Jersey, USA.

Leibold, M. A., and G. M. Mikkelson. 2002. Species Turnover , and Boundary Clumping:

Elements of Meta-Community Structure. Oikos 97:237–250.

Lee, S. C, and J. F. Bruno. 2009. Propagule supply controls grazer community structure and primary production in a benthic marine ecosystem. Proceedings of the National Academy of Sciences 106:7052–7057.

Levin, S. A., and R. T. Paine. 1974. Disturbance, patch formation, and community structure. Proceedings of the National Academy of Sciences. 71:2744–2747.

Lewis, L. S., and T. W. Anderson. 2016. Top-down control of epifauna by fishes enhances seagrass production. Ecology 93:2746–2757.

Markel, R. W., K. E. Lotterhos, and C. K. Robinson. 2017. Temporal variability in the environmental and geographic predictors of spatial-recruitment in nearshore rockfishes. Marine Ecology Progress Series 574:97–111.

Mitarai, S., D. A. Siegel, and K. B. Winters. 2008. A numerical study of stochastic larval settlement in the California Current system. Journal of Marine Systems 69:295–309.

Mouquet, N., and M. Loreau. 2003. Community patterns in source-sink metacommunities. American Naturalist 162:554–557.

O'Connor, M.I. 2009. Warming strengthens an herbivore-plant interaction. Ecology 90:388–398.

Orth, R. J., K. L. Heck, and J. van Montfrans. 1984. Faunal Communities in Seagrass Beds: A Review of the Influence of Plant Structure and Prey Characteristics on Predator-Prey Relationships. Estuaries 7:339-350.

Ovaskainen, O., G. Tikhonov, A. Norberg, F. Guillaume Blanchet, L. Duan, D. Dunson, T. Roslin, and N. Abrego. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecology Letters 20:561–576.

Presley, S.J., C. L. Higgins, and M. R. Willig. 2010. A comprehensive framework for the

evaluation of metacommunity structure. Global Ecology and Evolution 19:908–917.

R Development Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org.

Ricklefs, R. E., and D. Schluter. 1993. Species Diversity in Ecological Communities: Historical and Geographical Perspectives. University of Chicago Press, Chicago, Illinois, USA.

Ron, R., O. Fragman-Sapir, and R. Kadmon. 2018. Dispersal increases ecological selection by increasing effective community size. Proceedings of the National Academy of Sciences 115:11280–11285.

Rugiu, L., I. Manninen, J. Sjöroos, and V. Jormalainen. 2018. Variations in tolerance to climate change in a key littoral herbivore. Marine Biology 165:1–11.

Sala, E., and M. H. Graham. 2002. Community-wide distribution of predator-prey interaction strength in kelp forests. Proceedings of the National Academy of Sciences 99:3678–3683.

Schmida, A., and M.V. Wilson. 1985. Biological determinants of species diversity. Journal of Biogeography 12:1–20.

Thom, R., B. Miller, and M. Kennedy. 1995. Temporal patterns of grazers and vegetation in a temperate seagrass system. Aquatic Botany 50:201–205.

Thompson, P. L., B. Rayfield, and A. Gonzalez. 2017. Loss of habitat and connectivity erodes species diversity, ecosystem functioning, and stability in metacommunity networks. Ecography 40:98–108.

Treml, E.A., P. N. Halpin, D. L. Urban, and L.F. Pratson. 2008. Modeling population connectivity by ocean currents: a graph-theoretic approach for marine conservation. Landscape Ecology 23:19–36.

Treml, E. A., J. R. Ford, K. P. Black, and S. E. Swearer. 2015. Identifying the key biophysical

drivers, connectivity outcomes, and metapopulation consequences of larval dispersal in the sea. Movement Ecology 3:17.

Tuomisto, H., L. Ruokolainen, and K. Ruokolainen. 2012. Modelling niche and neutral dynamics: on the ecological interpretation of variation partitioning. Ecography 35:961-971.

Valanko, S., J. Heino, J. M. Westerbom, M. Viitasalo, and A. Norkko. 2015. Complex metacommunity structure for benthic invertebrates in a low-diversity coastal system. Ecology and Evolution 5:5203–5215.

Viana, D. S., and J. M. Chase. 2018. Spatial scale modulates the inference of metacommunity assembly processes. In press. Ecology doi: 10.1002/ecy.2576

Vieira, E. A., A. A. V. Flores, and G. M. Dias. 2018. Persistence and space preemption explain species-specifc founder effects on the organization of marine sessile communities. Ecology and Evolution 8:3430-3442.

Vellend, M. 2010. Conceptual Synthesis in Community Ecology. The Quarterly Review of Biology, 85:183–206.

Vellend, M. 2016. The Theory of Ecological Communities: Princeton University Press.

Ward, J. H. 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58:236-244.

Wichmann, C. S., I. A. Hinojosa, and M. Thiel. 2012. Floating kelps in Patagonian Fjords: An important vehicle for rafting invertebrates and its relevance for biogeography. Marine Biology 159:2035–2049.

Whippo, R., N. S. Knight, C. Prentice, J. Cristiani, M. R. Siegle, and M. I. O'Connor. 2018. Epifaunal diversity patterns within and among seagrass meadows suggest landscape-scale biodiversity processes. Ecosphere 9:02490.10.1002/ecs2.2490.

Wlodarska-Kowalczuk, M., E. Jankowska, L. Kotwicki, and P. Balazy. 2014. Evidence of Season-Dependency in Vegetation Effects on Macrofauna in Temperate Seagrass Meadows (Baltic Sea). PLoS One 9:e100788.

Yamada, K., Y. Tanaka, T. Era, and M. Nakaoka. 2014. Environmental and spatial controls of macroinvertebrate functional assemblages in seagrass ecosystems along the Pacific coast of northern Japan. Global Ecology and Conservation 2:47–61.

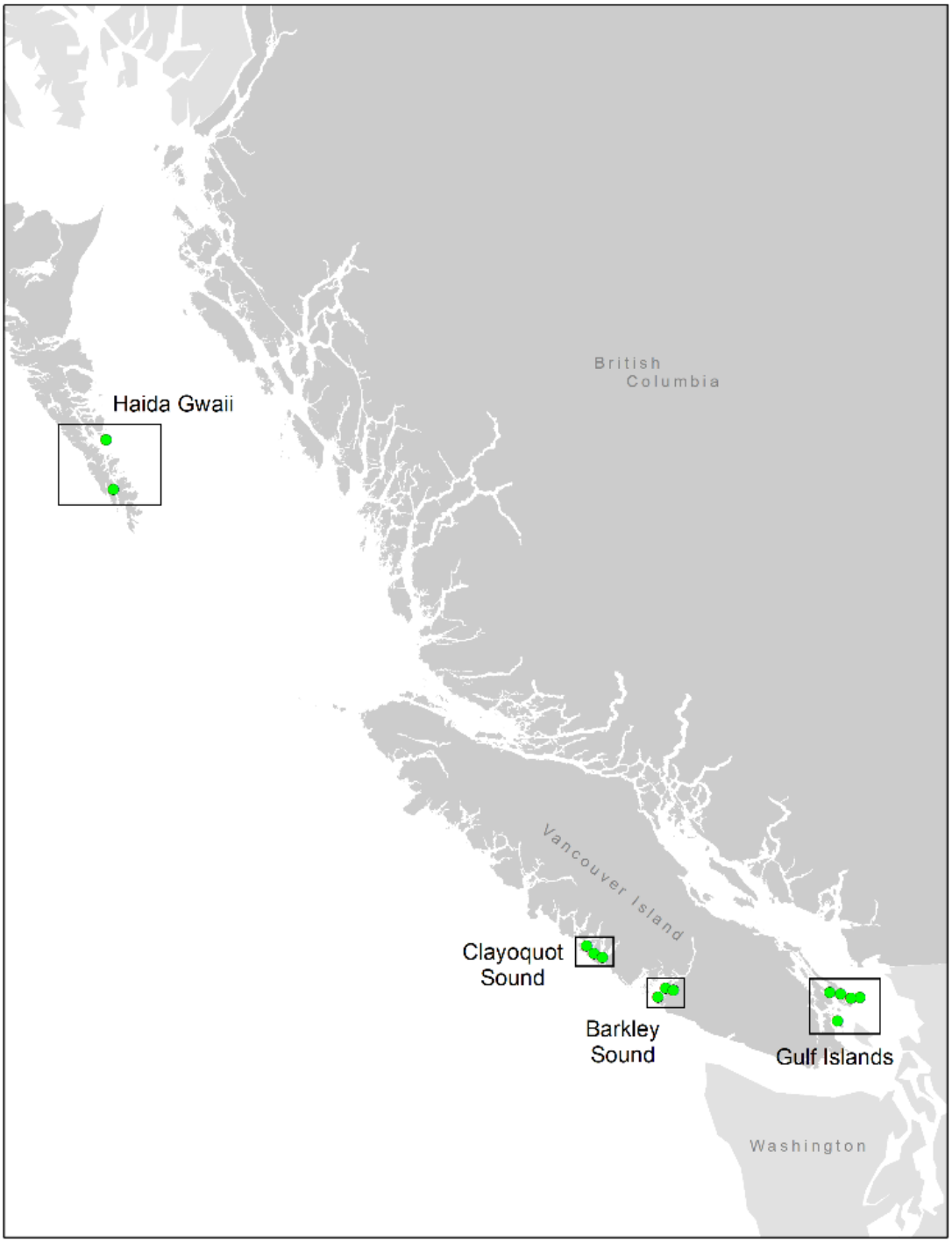
666 **Table 1.** Summary of evidence used to answer questions about metacommunity processes

Question	Analyses	Figure number
(1) Is the metacommunity characterized by dispersal limitation, intermediate dispersal, or high dispersal?	Dispersal limitation: Distance-decay of community similarity from HMSC joint species distribution models with and without accounting for environmental similarity	Fig. 2
	Intermediate versus high dispersal: Biodiversity patterns across sites for signatures of high dispersal (compositional homogenization); HMSC model results indicating structured co-occurrence patterns and variation in composition explained by environmental	Fig. 4a, 4b
(2) Do species distribution patterns suggest environmental niche filtering, and if so along which environmental axis?	Variation partitioning of modelled site and region-level fixed and random effects of environmental covariates and calculating predictive power (proportion of deviance) of the fixed effects model	Fig. 3
(3) Do species co-variance patterns suggest possible biotic interactions influencing community assembly?	Correlation matrix of pairwise species co-occurences, after controlling for the effect of space and environment	Fig. 4a & 4b

667

668

669



670 **Figure 1.** Map of coastal British Columbia showing eelgrass field sites in Haida Gwaii (HL,
671 RA), Clayoquot Sound (DK, EB, IN), Barkley Sound (SA, DC, RB), and the Southern Gulf
672 Islands (GB, JB, CB, LH, SS).

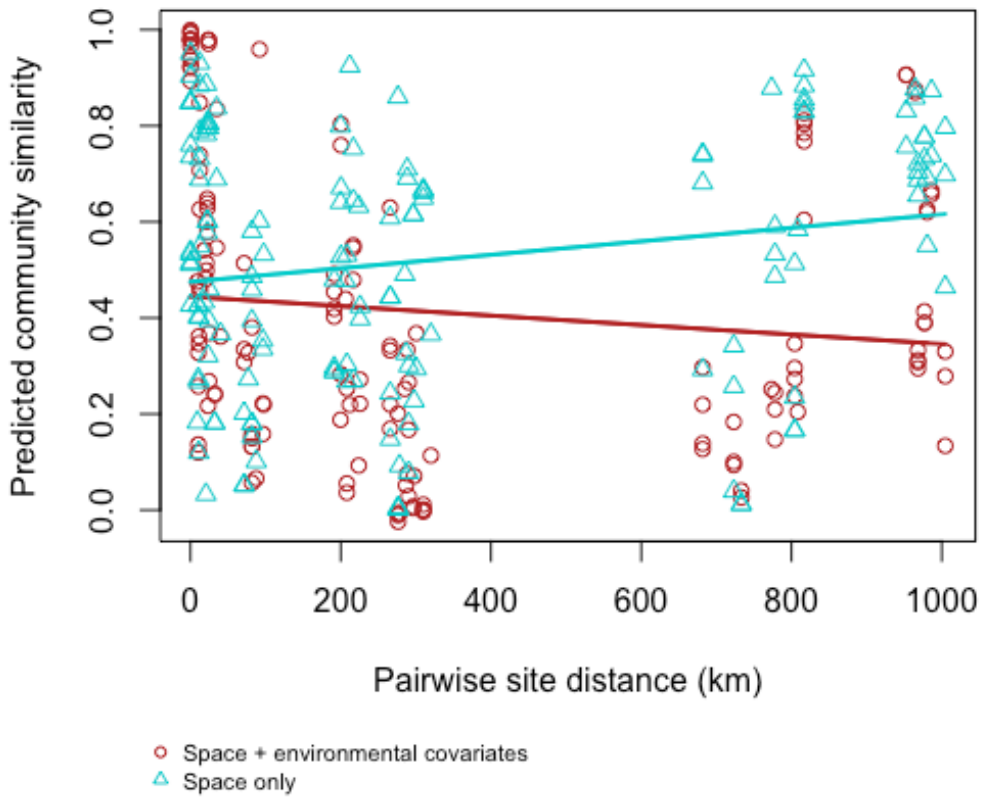


Figure 2. HMSC model predictions of community similarity relative to the spatial distance (Euclidean) between communities. The red circles show similarity predictions based on the full model. The blue triangles show predictions wherein the effect of all environmental parameters (9 water quality and 5 biotic) has been removed (standardized to their mean values), thus showing predicted community similarity as a result of spatial distance only.

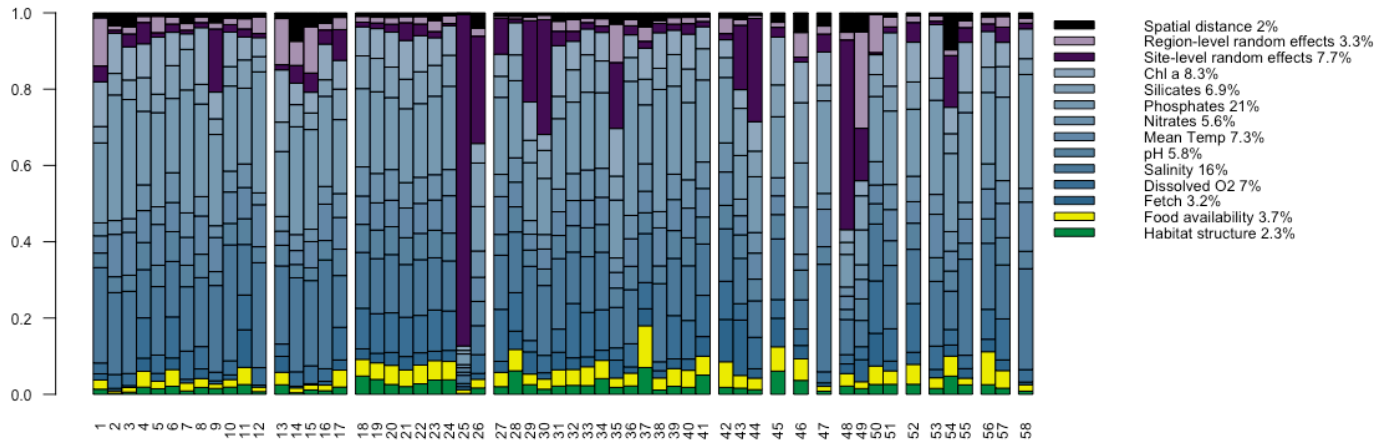


Figure 3. Variance partitioning of fixed (environmental variables, spatial distance) and random effects across all species (numbers correspond to species names in Appendix 1: Table S1) in the HMSC model. Purple cells represent site and region-level random effects. Blue cells represent abiotic water quality parameters. Yellow cells represent food availability (sum of variances explained by eelgrass, algae and detritus biomass). Green cells represent habitat structure, which is the sum of variances explained by eelgrass leaf area index and shoot density. Bars are organized according to broad taxonomic groupings as indicated in Appendix 1: Table S1. The percentages shown next to the legend labels indicate mean variance explained by that environmental variable across all species. Adjusted deviance explained (D^2_{adj}) for the individual species distribution models varied (Appendix 1: Table S1). The mean D^2_{adj} across all species was 0.29.

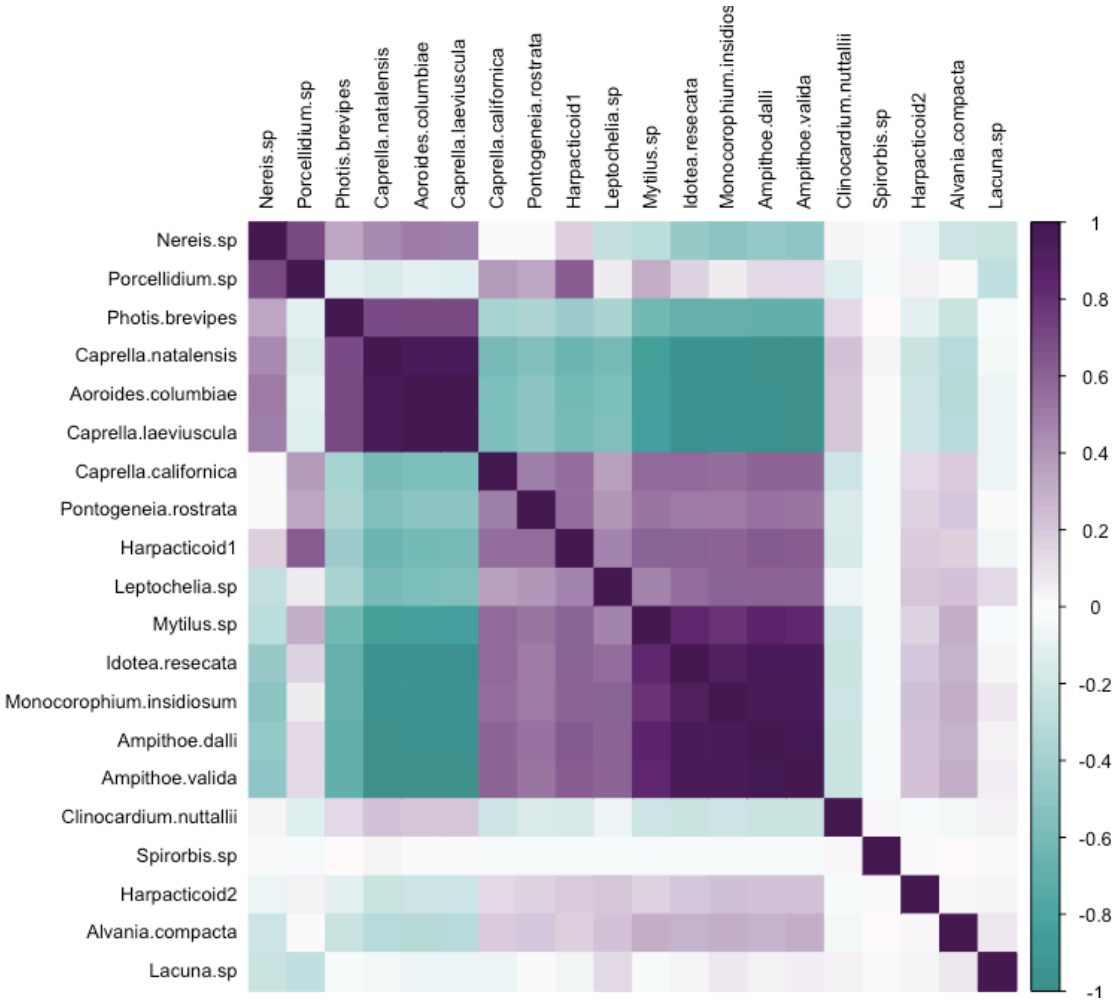


Figure 4a. A correlation matrix showing modelled site-level co-occurrence of species pairs in the 20 most abundant species. Purple cells represent positively co-occurring species pairs, and turquoise cells represent negatively co-occurring species pairs. Species names along both axes are ordered according to the output of hierarchical clustering with Ward's criterion on pairwise co-occurrence values. A version of this figure with all 58 species is shown in Appendix S1: Fig. S4.

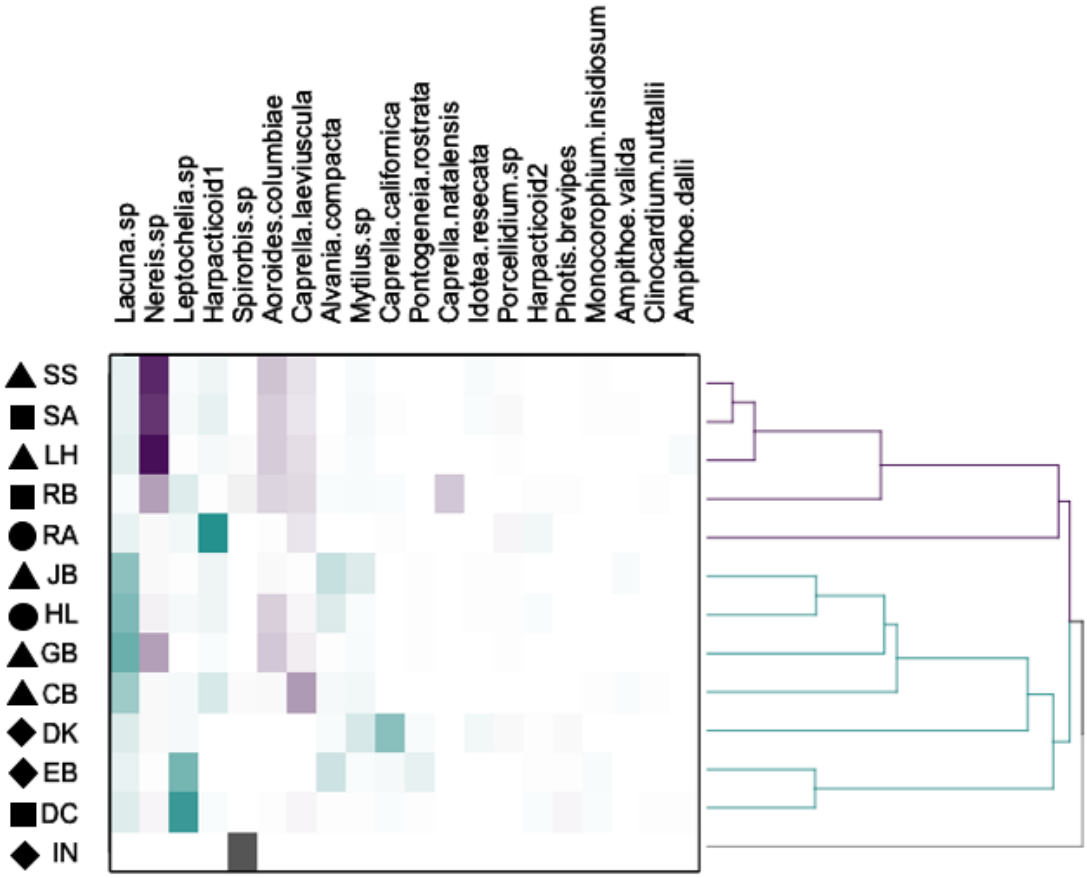


Figure 4b. Heatmap and cluster dendrogram depicting modelled (HMSC) species abundance and compositional similarity across sites. The heat map shows the twenty species with the highest predicted mean proportional abundance at the site level, and is ordered from highest to lowest predicted mean proportional abundance. Cell colours correspond to the HMSC co-occurrence groups shown in Fig. 4a: *Nereis* (turquoise group), *C. californica* (purple) group, and *Spirorbis* (grey) group. Cell shade strength represents proportional abundance at a given site (darker means higher abundance). The symbols to the left of the site abbreviations indicate region membership; circles show Haida Gwaii sites, squares show Barkley Sound sites, diamonds show Clayoquot Sound sites, and triangles show Southern Gulf Islands sites.

727 **Appendix S1:** Supporting Information

728

729 **Table S1.** Species names, taxonomic groupings, and proportion of deviance (predictive power).

730 Numbers correspond to the axes in Figure 3. The top twenty species with highest predicted

731 abundances according to the HMSC model are bolded.

Number	Species name	Broad taxonomic group	Adjusted deviance explained for the HMSC model (D^2_{adj})
1	*Alia carinata*	Gastropod (snail)	0.15
2	***Alvania compacta***	**Gastropod (snail)**	**0.4**
3	*Amphissa columbiana*	Gastropod (snail)	0
4	*Crepidula* sp.	Gastropod (snail)	0.64
5	*Euspira lewisii*	Gastropod (snail)	0.46
6	*Harmothoe imbricata*	Gastropod (snail)	0
7	***Lacuna* spp. *(L. variegata and L. vincta)***	**Gastropod (snail)**	**0.45**
8	*Lirularia parcipincta*	Gastropod (snail)	0.46
9	*Lirularia* sp.	Gastropod (snail)	0
10	*Littorina* sp.	Gastropod (snail)	0.5
11	*Margarites pupillus*	Gastropod (snail)	0.45
12	*Lottia pelta*	Gastropod (limpet)	0.53
13	***Clinocardium nuttallii***	**Bivalve**	**0**
14	***Mytilus* sp.**	**Bivalve**	**0.28**
15	*Saxidomus gigantea*	Bivalve	0
16	*Unknown clam 1*	Bivalve	0.68
17	*Unknown clam 2*	Bivalve	0.35
18	*Dorvillea longicornis*	Polychaete	0.52
19	*Exogone sp.*	Polychaete	0.25
20	*Glycinde armigera*	Polychaete	0.05
21	*Lepidonotus squamatus*	Polychaete	0.55
22	*Neoamphitrite robusta*	Polychaete	0.22
23	*Unknown polychaete (Family Oenoidae)*	Polychaete	0.15
24	*Unknown polychaete (Family Opheliidae)*	Polychaete	0.36
25	***Spirorbis* sp.**	**Polychaete**	**0**

26	***Nereis* sp.**	**Polychaete**	**0.55**
27	***Ampithoe dalli***	**Gammarid amphipod**	**0.47**
28	*Ampithoe lacertosa*	Gammarid amphipod	0.34
29	***Ampithoe valida***	**Gammarid amphipod**	**0**
30	***Aoroides* spp.**	**Gammarid amphipod**	**0**
31	*Ceradocus spinicauda*	Gammarid amphipod	0
32	*Grandidierella japonica*	Gammarid amphipod	0
33	*Ischyocerus anguipes*	Gammarid amphipod	0.33
34	*Jassa marmorata*	Gammarid amphipod	0
35	***Monocorophium insidiosum***	**Gammarid amphipod**	**0.23**
36	*Orchomenella recondita*	Gammarid amphipod	0.73
37	***Photis brevipes***	**Gammarid amphipod**	**0**
38	***Pontogeneia rostrata***	**Gammarid amphipod**	**0.65**
39	*Unknown gammarid (Family Hyalidae)*	Gammarid amphipod	0.12
40	*Unknown gammarid (Family Isaeidae)*	Gammarid amphipod	0.10
41	*Unknown gammarid (Family Ischyoceridae)*	Gammarid amphipod	0.24
42	***Caprella californica***	**Caprellid amphipod**	**0.36**
43	***Caprella laeviuscula***	**Caprellid amphipod**	**0.06**
44	***Caprella natalensis***	**Caprellid amphipod**	**0.82**
45	*Cumella vulgaris*	Cumacean	0.36
46	***Leptochelia sp.***	**Tanaid**	**0.49**
47	*Nebalia gerkinae*	Leptostracan	0.39
48	***Porcellidium sp.***	**Copepod**	**0.33**
49	***Harpacticoid 1***	**Copepod**	**0**
50	***Harpacticoid 2***	**Copepod**	**0**
51	*Harpacticoid 3*	Copepod	0.50
52	*Anoplodactylus viridintestinalis*	Pycnogonid	0.01
53	*Idotea montereyensis*	Isopod	0.73
54	***Idotea resecata***	**Isopod**	**0.36**
55	*Munna sp.*	Isopod	0.20
56	*Pugettia producta*	Brachyuran crab	0.36
57	*Pugettia richii*	Brachyuran crab	0.34
58	*Hippolyte californiensis*	Caridean shrimp	0.45

735 **Table S2. Summary of traditional biodiversity metrics for each site**

736

Site abbrev	Region	Species richness	Raw abundance	Effective diversity (Hill number)
DC	Barkley Sound	21	1168	11.8
RP	Barkley Sound	17	1590	8.8
SA	Barkley Sound	17	924	6.8
DK	Clayoquot Sound	21	1168	11.8
EB	Clayoquot Sound	21	4812	8.5
IN	Clayoquot Sound	16	1803	7.5
CB	Gulf Islands	18	2112	5.0
GB	Gulf Islands	11	2239	4.3
JB	Gulf Islands	22	3522	7.1
LH	Gulf Islands	15	2259	6.6
SS	Gulf Islands	10	740	6.6
HL	Haida Gwaii	9	16743	1.3
RA	Haida Gwaii	14	934	8.6

737

738

739

740

741

Equation S1. Equations for D^2_{adj} calculations

Explained deviance or D^2 (Guisan and Zimmerman 2000) was calculated by the equation:

$$D^2 = 1 - \text{Residual deviance} / \text{Null deviance}$$

Where the residual deviance refers to the deviance in predicted species occurrences, and the null deviance refers to deviance in observed species occurrences.

Adjusted D^2 (Guisan and Zimmerman 2000) was calculated by the equation:

$$D^2_{adj} = 1-(((n-1)/(n-p-1)*(D^2-1))$$

Where n is the number of sites (13) in our study, and p is the number of parameters the model estimated (4, Ovaskainen et al. 2017, Supporting Information). Negative D^2_{adj} values (those which had greater residual deviance than null deviance) were set to 0.

Equation S2. Calculations for type III variation partitioning on HMSC model (Borcard et al. 1992)

We estimated three versions of the HMSC model: 1) environmental variables only with the random effects (*m2*), and 2) spatial distances only with random effects, and 3) the original global model which included environmental and spatial variables, site- and region-level random effects (*m1*). From each model, we extracted the total variation explained by environment and space, the variation explained by environment, and the variation explained by space. To determine the shared fraction of variation between environment and space, we multiplied each whole fraction by the calculated community-level D^2_{adj} to reflect the model deviance in our variance partitioning estimates, and followed the equations

m1 = the fraction of variation explained by all environmental variables in the HMSC model estimated with only environmental variables;

m2 = the fraction of variation by spatial distances in the version of the HMSC model estimated with only spatial distances;

m3 = the global model including both;

Fraction *ab* (pure environment + shared fraction) = $D^2_{adj}\ m1$

Fraction *bc* (pure space + shared fraction) = $D^2_{adj}\ m2$

Fraction *abc* (pure environment + shared fraction + pure space) = $D^2_{adj}\ m3$

Fraction *a* (pure environment) = *abc* - *bc*

Fraction *c* (pure space) = *abc* - *ab*

Fraction *b* (shared fraction) = *ab* + *bc* - *abc*

Fraction *d* (residuals) = 1 - *abc*

Fractions *a* ("Environmental conditions"), *c* ("Spatial distance"), *b* (shared fraction), and *d* ("Residuals") are shown in the Venn diagram in Fig. S3.

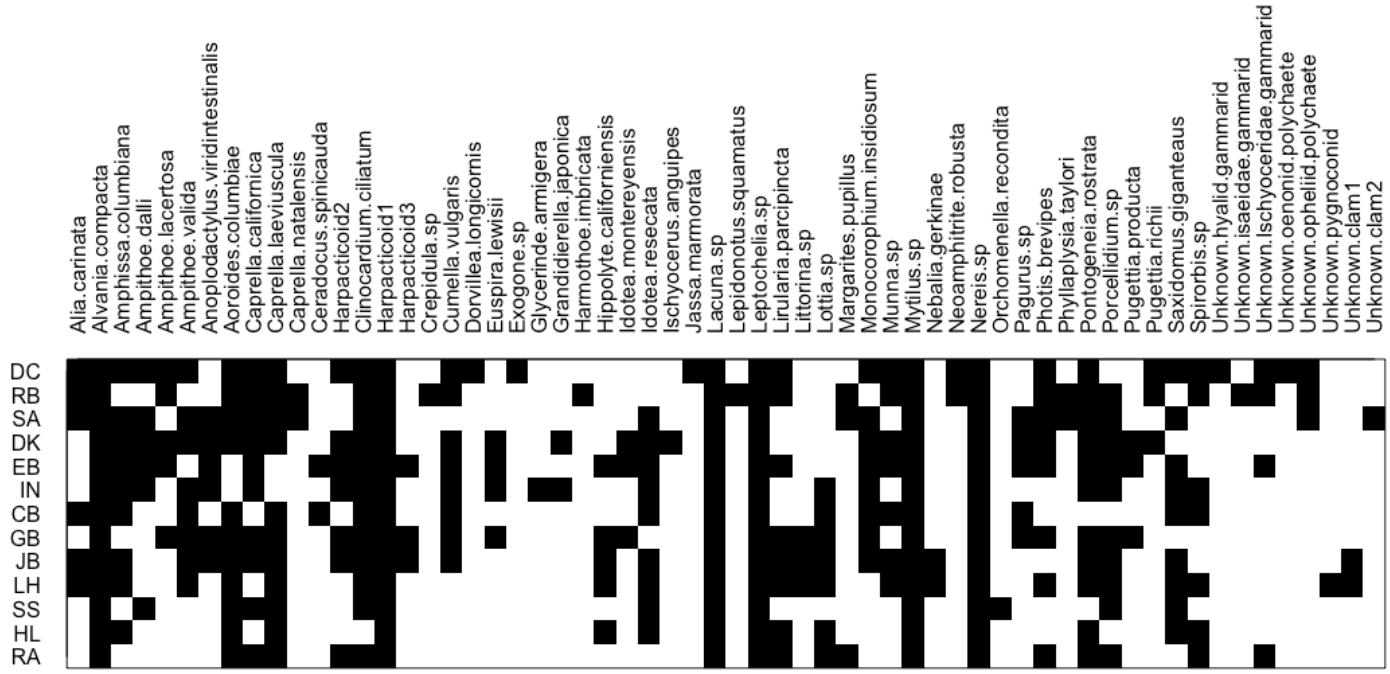


Figure S2. Site-by-species presence-absence matrix of all 58 invertebrate species in the study listed in alphabetical order. Black cells indicate species presences.

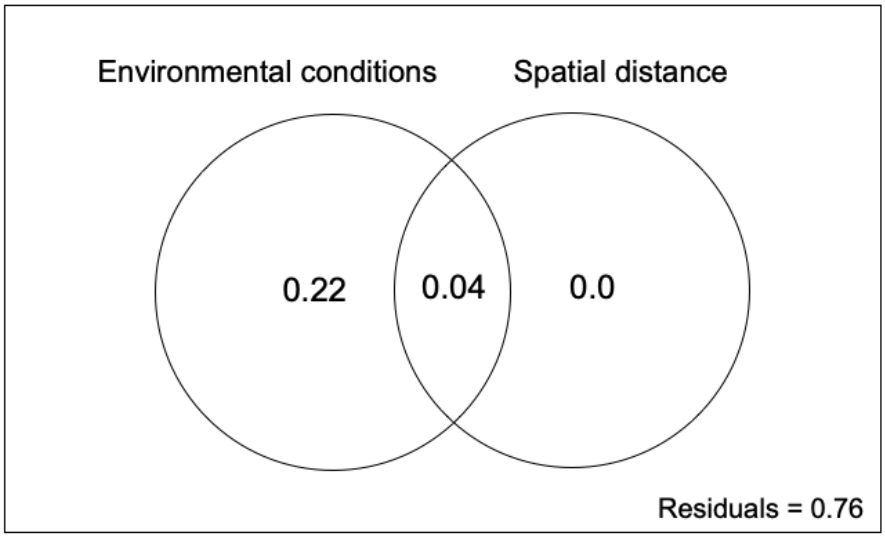


Figure S3. Venn diagram summarizing the fractions of variation explained by environmental covariates only (nine water quality and five biotic variables), spatial distance only, and the shared fraction explained by environment and space.

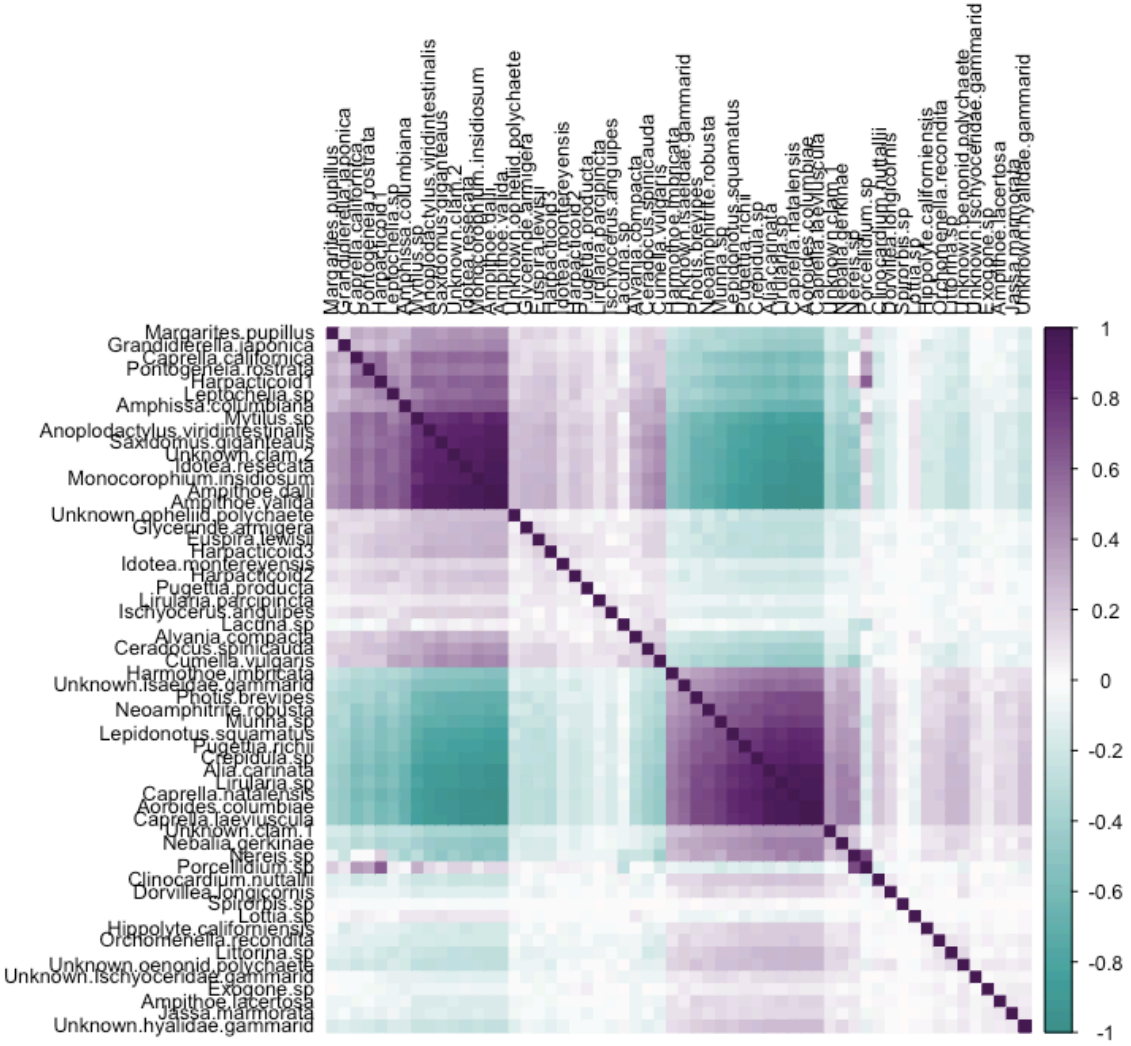


Figure S4. A correlation plot showing modelled site-level co-occurrence of species pairs across all species. Purple cells represent positively co-occurring species pairs, and turquoise cells represent negatively co-occurring species pairs. Species names along both axes are ordered according to the output of hierarchical clustering with Ward's criterion (Ward 1963) on pairwise co-occurrence values. This figure is a supplement to Fig. 4a in the main text.

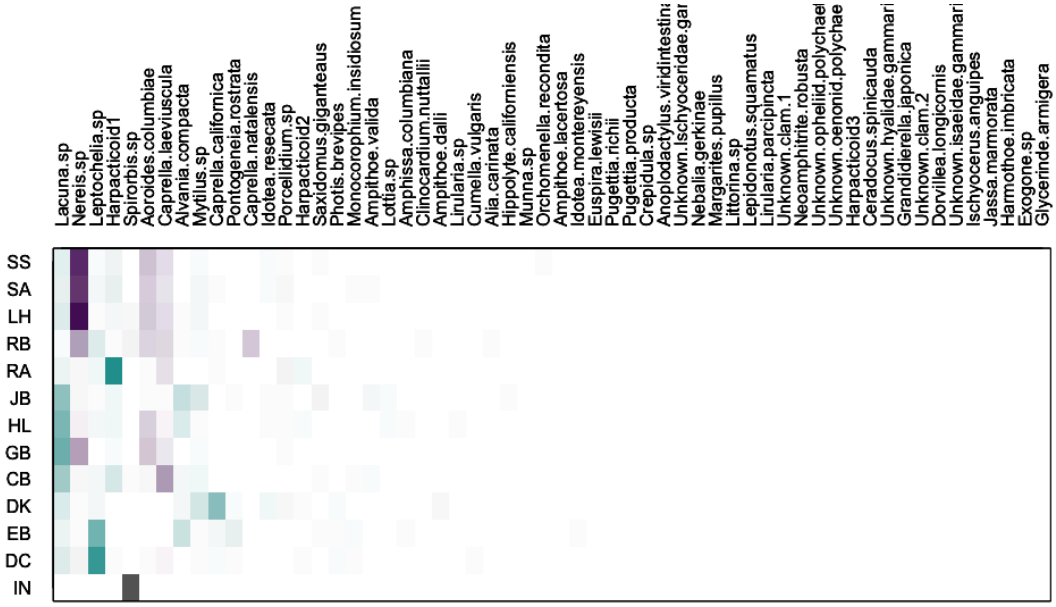


Figure S5. Heat map depicting abundance patterns in all 58 species in the study. Species names are ordered from highest to lowest predicted mean proportional abundance. Cell colours correspond to the co-occurrence groups in Fig. 4b, with turquoise cells representing members of the *Nereis* group, purple cells representing members of the *C. californica* assemblage, and grey cells indicating species that did not significantly co-occur negatively or positively with other species. Cell shade strength represents proportional abundance at a given site (darker means higher abundance). Most species outside the top twenty most abundant had extremely low predicted proportional abundances owing to their low raw abundances. This figure is a supplement to Fig. 4b in the main text.

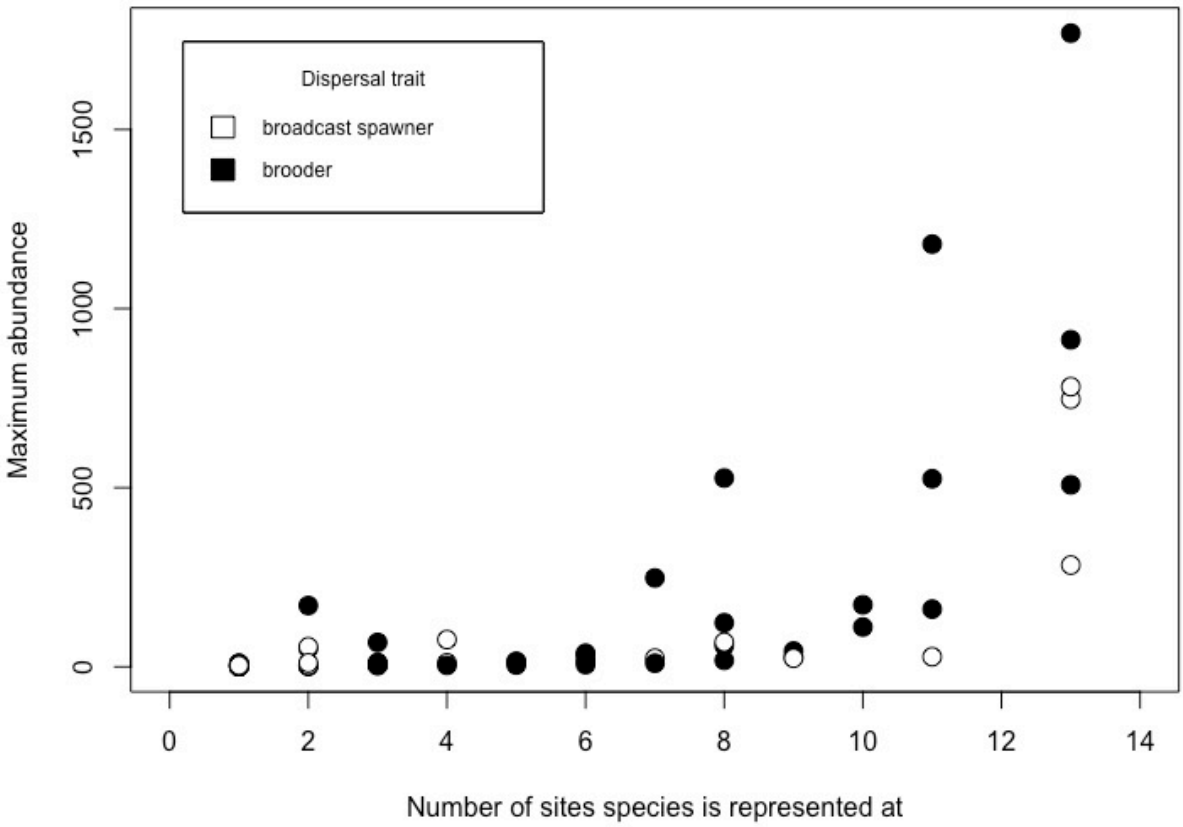


Figure S6. Plot showing relationship between maximum site-level abundance and site occupancy for all species. Black points represent broadcast spawners and white points represent brooders. Spirorbid polychaetes were excluded from this figure due to their unusually high maximum abundance (approximately 15000 at one site). A two-tailed t-test showed that there was no significant difference in the number of sites colonized between brooders and broadcast spawners ($t = 0.052$, $df = 42.5$, $p = 0.96$).