Regional connectivity, local environmental conditions, and biotic interactions jointly structure a temperate 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:** keila.stark@zoology.ubc.ca

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

Abstract: 271 words
Main text: 6196 words
Number of references: 47
Number of figures: 5
Number of tables: 2

Author contributions: KAS conceived the study with guidance from MIO. LL provided samples from RA and HL. JY provided samples from IN, DK, and EB. KAS collected samples from GB, JB, CB, LH, SS, RB, SA, and DC, processed samples, analyzed data, and wrote the first draft of the manuscript. PLT provide substantial recommendations for the analysis. KAS, PLT, and MIO contributed to writing the manuscript. All authors gave final approval for publication.

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

Dispersal, environmental niche filtering, and biotic interactions all structure patterns of abundance and distribution in ecological communities. Metacommunity ecology provides a framework for understanding how these processes work together to determine community composition in sites within a region that are linked by dispersal of multiple potentially interacting populations. Previous metacommunity approaches have sometimes had limited success due to statistical limitations and incompatibilities between theory and methods. Here, we use the *Hierarchical Modelling of Species Communities* statistical framework to identify the contribution of the aforementioned assembly processes to biodiversity patterns in an important coastal habitat, seagrass meadows, distributed across approximately 1000 km on the coast of British Columbia, Canada. We used this method to assess the degree to which seagrass epifaunal invertebrate communities are structured by dispersal, environmental niche filtering, and biotic interactions. Observed species distributions and co-occurrence patterns suggest that seagrass epifauna form well-connected metacommunities across multiple meadows, but that dispersal rates are not so high that they homogenize the region through mass effects. We observed niche filtering along an abiotic gradient, with average phosphates, sea surface temperature and dissolved oxygen being the most important environmental variables influencing species distributions. However, environmental gradients only explained a subset of the variation in community composition. Strong structuring in pairwise species co-occurrences led us to hypothesize that local community composition may be determined by priority effects resulting from biotic interactions. This approach has provided new insight into the local and regional processes influencing seagrass epifaunal community assembly. This knowledge can inform targeted experiments testing for metacommunity processes, monitoring efforts, and management

71 decisions for preserving biodiversity and the future health of this important coastal ecosystem.

72

73 **Introduction**

74 Understanding how local environmental conditions, biotic interactions, and spatial
75 connectivity combine to structure the composition of communities is a central challenge in
76 ecology (Vellend 2016, Leibold and Chase 2017). Determining the relative importance of these
77 processes in particular ecosystems is a critical step in understanding the spatial and temporal
78 scales of stability and resilience in spatially structured communities (Leibold and Chase 2017).
79 Metacommunity theory offers an integrative framework for understanding the structure and
80 dynamics of local and regional species diversity but previous attempts to quantitatively apply
81 metacommunity concepts to real-world ecological systems have had limited success (Logue et al.
82 2011). Consequently, the potential benefits to conservation prioritization that could be gleaned
83 by applying metacommunity theory to biodiversity have not been fully realized (Economo 2011).
84 Many marine ecosystems that host high biodiversity occupy spatially structured habitats such
85 as rocky reefs, coral reefs, seagrass meadows, and marine species have various dispersal
86 strategies that connect populations in distinct habitat patches. Marine systems are characterized
87 by strong local species interactions (Sala and Graham 2002, Berlow 1999) and the importance of
88 'supply-side' or dispersal-driven population dynamics has long been recognized (Levin and
89 Paine 1974, Gaines and Roughgarden 1985). The combined roles of dispersal, niche filtering,
90 and species interactions suggest that marine biodiversity patterns at local scales may in fact
91 reflect local and regional processes characteristic of metacommunities (Guichard et al. 2003).
92 Despite the likely importance of metacommunity processes in marine systems, and the
93 historically strong focus on dispersal and species interactions in marine community ecology

(Bertness book), metacommunity dynamics have only been explicitly documented in a few marine systems (Valanko et al. 2015, Whippo et al. 2018, Yamada et al. 2014, Guichard et al 2003.

Seagrasses, like kelps, corals and salt marsh plants, are foundation species that support high productivity and animal biodiversity (Duffy et al 2015, Heck and Thoman 1984, Orth et al 1984). Seagrasses form meadows separated from each other by deeper water, un-vegetated seafloor or other terrain, and faunal species disperse among meadows and interact within meadows through predation, competition and facilitation (Bostrom et al 2006, Williams and Heck 2001). This scenario is consistent with the kind of landscape that supports communities with diversity patterns governed by metacommunity dynamics (Mouquet and Loreau 2003, Leibold et al. 2004). Yet, our understanding of seagrass-associated faunal diversity and its consequences for conservation and ecosystem function does not yet reflect the possibility that seagrass biodiversity is maintained by processes operating across numerous seagrass meadows and adjacent habitats in a seascape. Numerous seagrass community ecology studies have focused on the influence of local, site-level environmental characteristics such as habitat structure (seagrass shoot density and length) and primary productivity (algal and detritus biomass) on the structure of the faunal communities they host (Williams and Heck 2001). Metacommunity processes allow for co-existence at local and regional scales that would not be possible in unconnected local habitats (Loreau et al. 2003, Leibold and Chase 2017), so failure to recognize metacommunity processes could lead to underestimates of the amount of habitat required to sustain biodiversity and ecosystem functions (Boström et al 2006).

Two studies that have tested for metacommunity dynamics in seagrass epifauna found evidence of important regional-scale processes (Yamada et al. 2014, Whippo et al. 2018), but

117 their ability to describe which metacommunity dynamics were contributing to seagrass
118 metacommunity structure was limited by the statistical methods available at the time. Yamada et
119 al. (2014) used the variation partitioning method (Borcard et al. 1992, Cottenie 2005), and found
120 that the relative contribution of space and environment varied among invertebrate functional
121 groups, suggesting that the scale and relative contribution of metacommunity dynamics could
122 vary among functional groups. However, the variation partitioning analysis did not allow for
123 nference of the specific metacommunity dynamics contributing to the result (Yamada et al.
124 2014). Simulation studies have shown that variation partitioning can over- or under-estimate
125 spatial and environmental components of variation, and produce inflated R^2 values depending on
126 the researcher's statistical choices, sometimes yielding inconclusive results (Gilbert and Bennett
127 2010, Tuomisto et al. 2012). Whippo et al. (2018) used another method, the elements of
128 metacommunity structure (EMS) (Leibold and Mikkelson 2002, Presley et al. 2010), to
129 distinguish among possible metacommunity types as the cause of spatial patterns in diversity in
130 seagrass meadow landscapes. While useful for visualizing regional diversity patterns, EMS does
131 not explicitly elucidate the ecological mechanisms contributing to these patterns (Brown et al.
132 2017). Neither EMS nor variation partitioning assesses the contribution of interspecific
133 interactions, thereby entirely missing a critical aspect of community assembly.
134 Another limitation in previous metacommunity studies has been the way researchers have
135 conceptualized and investigated the processes underlying metacommunity structure. Over the
136 past decade and a half, the primary focus of metacommunity studies has been seeking signatures
137 of one of the four original metacommunity paradigms in natural systems: the neutral model,
138 patch-dynamics, species sorting and mass effects (Leibold et al. 2004, Cottenie 2005). An
139 emerging perspective holds that natural communities do not fit neatly under one of the four

140 paradigms, but rather that they are governed by multiple metacommunity processes operating in
141 concert (Logue et al. 2011, Brown et al. 2017, Leibold and Chase 2017). Thus, the goal is to
142 determine how biotic interactions, abiotic environmental gradients, and dispersal combine to
143 structure the composition and dynamics in a given metacommunity.

144 This revised goal invites novel methodological approaches to attribute community patterns to
145 the underlying ecological processes. A new statistical framework introduced by Ovaskainen et al.
146 (2017) addresses the multiple limitations in previous metacommunity approaches. Hierarchical
147 Modelling of Species Communities (HMSC) uses Bayesian inference to estimate parameters
148 summarizing influence of space, environment, patterns of co-occurrence, phylogeny and traits on
149 species distributions. The HMSC framework allows for variation partitioning of environmental
150 and spatial variables on all species distributions, allowing the observer to discern patterns at the
151 whole community level as well as individual variation in responses to the environment and
152 space. Additionally, the framework models the residual variance in species distributions
153 unexplained by environment and space to random effects as a variance-covariance matrix, which
154 can be used to describe the extent to which species positively or negatively co-occur,
155 highlighting potential interspecific interactions. For these reasons, HMSC achieves what
156 researchers aimed to achieve with the previous variation partitioning method, and much more.

157 Here we test the hypothesis that seagrass-associated epifaunal invertebrate communities are
158 jointly structured by the spatial distances (a proxy for inferring dispersal dynamics) between
159 meadows, local environmental conditions, and biotic interactions. Our approach moves away
160 from the four original metacommunity paradigms (Leibold et al. 2004), instead focusing on
161 individually characterizing the effects of dispersal, environmental niche filtering, and
162 interspecific interactions on the composition of seagrass associated communities across space.

We follow the Hierarchical Modelling of Species Communities (Ovaskainen et al. 2017) framework to model the effect of spatial distance, environmental conditions, and patterns in species covariance on abundance and distribution in epifaunal invertebrates in BC seagrass meadows to answer the following questions: (1) Is the metacommunity characterized by dispersal limitation, intermediate dispersal, or high dispersal? (2) Do patterns of species' abundance and distribution suggest environmental niche filtering, and if so along which environmental axes? (3) Do species co-variance patterns suggest possible biotic interactions influencing community assembly? In answering these questions with a biodiversity survey of invertebrates in 13 meadows in British Columbia, we also provide a comprehensive assessment of species' abundance and diversity from local to regional scales in a temperate seagrass habitat. We find evidence of intermediate dispersal rates among meadows across 1000 km, but no sign of complete biodiversity homogenization resulting from high dispersal. We find that local environmental conditions have a moderate effect on community composition, but that biotic interactions, likely resulting in priority effects, have a strong influence on the composition of local meadows. Together, these results indicate that local and regional processes interact to influence biodiversity in seagrass ecosystems.

Methods

Study sites

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 Marine Conservation Area Reserve, and Haida Heritage Site; southern Clayoquot Sound

186 (one site was within the boundaries of Pacific Rim National Park Reserve), Barkley Sound on the
187 West coast of Vancouver Island, and the Southern Gulf Islands on the south-eastern tip of
188 Vancouver Island within Gulf Islands National Park Reserve (Fig. 1).

189 *Zostera marina* (L.) is the most common subtidal seagrass species in the Northeast Pacific
190 (Alaska to southern California). In British Columbia, Canada, *Z. marina* meadows are found in
191 discrete patches along the coast. They vary in several biotic and abiotic conditions, such as
192 seagrass shoot surface area and density, macroalgal species composition, fish community
193 structure, temperature, salinity, and water column nutrient levels.

194 *Field sampling of eelgrass and associated epifauna*

195 We sampled three sites in Barkley Sound (Sarita, Dodger Channel, Robbers Passage)
196 with SCUBA because they required boat access, and the deepest parts of the meadow reached a
197 depth of 5m during sampling. All other sites were sampled by wading or snorkeling at low tide.
198 We used a standardized, nested quadrat sampling regime to collect aboveground eelgrass,
199 epifaunal invertebrates, detritus, and macroalgae (Whippo et al. 2018). Six 0.25 m x 0.25 m
200 quadrats were set in a 15 m x 30 m array, such that every quadrat was situated 15 metres from its
201 nearest neighbour (S1, Supplementary Material). The array was placed in the middle of the
202 eelgrass meadow to avoid edge effects where possible. We removed all above-ground biomass
203 (eelgrass, detritus, macroalgae) within each 0.25 m x 0.25 m quadrat by hand and immediately
204 placed the contents into a 250 µm mesh bag for the Barkley Sound and Southern Gulf Islands
205 sites, and into a Ziploc bag for the Clayoquot Sound and Haida Gwaii sites.

206

207 *Acquiring biotic and abiotic environmental variables*

208 To test our hypothesis about environmental niche filtering, we quantified 14

environmental attributes of each meadow and plot sampled. We used in-situ field measurements and satellite-derived averages from the Bio-ORACLE dataset (Assis et al. 2018) as estimates of water quality conditions associated with niche differences among epifauna. We gathered in-situ measurements of water quality parameters (pH, salinity, dissolved oxygen, nitrates and chlorophyll-*a*) at each meadow. We also acquired data on mean surface chlorophyll-*a* ($\mu g\ L^{-1}$), dissolved oxygen, surface nitrates, surface silicates, surface phosphates, pH, salinity, and mean sea surface temperature (°C) from Bio-ORACLE: a global database of marine ecological variables gathered from several satellite and in-situ sources at a spatial resolution of 0.08° (Assis et al. 2018). All sites (meadows) were situated in different spatial cells, therefore all sites had distinct values for the aforementions parameters. However all quadrats within a given meadow (spaced 15m from its adjacent neighbor) were situated in the same spatial cell and thus were given the same site-level value for all the Bio-ORACLE variables; we accounted for this spatially nested random effect while constructing our model. We also quantified fetch, which is a proxy for site exposure. We chose these abiotic variables because we expected them to be the most important variables influencing species distributions, either directly because species differ in their physiological tolerances, or indirectly by influencing primary productivity and food quality.

For each sample, we measured five biotic attributes: eelgrass leaf area index (LAI), shoot density, and dry mass, total algal dry mass, and total detritus dry mass. For our leaf area index estimates, we measured leaf length, width, and number of blades in five representative shoots within each quadrat. Leaf Area Index (m^2 eelgrass blade area per m^2 of ground) was calculated by multiplying the average blade area (m^2) per shoot by the shoot density (number of shoots per m^2). We standardized blade length measurements from the end of the blade to just above the first

232 rhizome node because some shoots had intact rhizomes while others were missing after
233 uprooting in the field. We separated and dried eelgrass, detritus, macroalgae in a desiccator oven
234 (60° C for 48 hours) to measure dry mass.
235
236 *Response variable: Quantifying epifaunal diversity*
237 We rinsed the eelgrass blades from each quadrat to remove epifaunal invertebrates, and
238 passed samples through a 500 μm sieve. Invertebrates were transferred into a centrifuge tube and
239 preserved in 95% EtOH for identification with light microscopy (10 times magnification). We
240 identified invertebrates to the lowest taxonomic level possible using keys in the Light and Smith
241 Manual (Carlton, 2007) and Kozloff (1996). For each quadrat, we identified every individual
242 collected and retained.
243
244 *Analysis*
245 We followed Ovaskainen et al. (2017) to construct a hierarchical joint species distribution
246 model describing the relationship between species environmental variables, spatial distance, and
247 species co-occurrences on species distributions using the "HMSC" package (Guillaume Blanchet
248 et al. 2018) in R (version 3.4.2). The joint species distribution model comprises two linear
249 predictor terms: one describing fixed effects of environment and space, and one describing
250 random effects that cannot be attributed to environment or space (Equations 2 & 3 in Ovaskainen
251 et al. 2017). The random-effects linear predictor term describes random variation in individual
252 species distributions as well as covariation for every species pair in the analysis. The full
253 parametrized model summarizes species' responses (presence and abundance) to the 9 water
254 quality and 5 biotic variables, spatial distance between sites, and co-occurences with other

species.

We used Markov chain Monte Carlo to estimate the model parameters. Species' responses to environmental variables were assumed to follow a multivariate Gaussian distribution. We assumed a Poisson likelihood distribution for species count data and used the same priors as in the examples presented in Ovaskainen et al. (2017). The Markov Chain included 100 000 iterations, a burn-in length of 1000 iterations, and thinning of 100. We confirmed that parameters were well mixed by visually inspecting trace plots. We tested the model's performance by calculating explained deviance (D^2_{adj}) as described in S2 (Supplementary Material). D^2_{adj} is comparable to R^2_{adj} in least square regression analysis (Guisan and Zimmermann 2000); in this context, it describes the model fit where a D^2_{adj} value of 1 means the model perfectly predicts the observed data. This yielded a D^2_{adj} value for each individual species distribution model, and the mean proportion of deviance across species gave a single metacommunity-level value of our model's predictive power.

To answer our three main questions about the relative roles of dispersal, environment, and species interactions as causes of biodiversity patterns, we visualized the parametrized model in several ways. We investigated three possible dispersal conditions: dispersal limitation, intermediate dispersal, and dispersal surplus. To investigate dispersal limitation at large spatial scales, we plotted distance-decay of community similarity. To do so we randomly drew 200 pairwise distances between sampling plots, and plotted predicted community similarity against distance. Distance decay of community similarity could reflect dispersal limitation, or it could reflect differences in environmental conditions with increasing spatial scale. To distinguish these two possibilities, we compared predicted species compositional similarity as a function of distance with site-level biotic and abiotic predictors, and an alternate model in which all sites and

plots shared the same environmental attributes set to the average values across sites (Equation 5 in Ovaskainen et al. 2017). To distinguish between dispersal sufficiency (intermediate) and dispersal surplus (high), we looked for evidence of biodiversity homogenization across sites- a signature of swamping by high dispersal- by visual inspection of our site-by-species matrix. The strength of environmental niche filtering (as measured by the proportion of deviance explained) also provides some insight on dispersal levels: intermediate dispersal allows species to settle at sites where local environmental conditions are optimal for growth and reproduction, whereas high dispersal overrides this.

To assess the degree of environmental niche filtering, we conducted variation partitioning analysis of the fixed and random effects as defined in Ovaskainen et al. using the variPart() function in the HMSC package. This yielded the relative contribution of every abiotic and biotic environmental variable, spatial distance, and random effects to the distribution of each species in the model; we visualized this in a stacked bar plot. We then used type III sum of squares (Borcard et al. 1992) to estimate the fractions of variation explained by environment only, space only, the shared variation between space and environment, and the residual or unexplained variation. To do so, we estimated three versions of the HMSC model: the original global model which included environmental and spatial variables, site- and region-level random effects, environmental variables only with the random effects, and spatial distances only with random effects. From each model, we obtained estimates of the total variance explained by environment and space, the variance explained by environment, and the variance explained by space, but not accounting for any shared variation between space and environment. To determine this, we then 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 in S5.a

301 (Supplementary Material) to calculate the pure fractions of variation explained by environment
302 and space (correcting for the shared variation between the two).
303 To identify possible within-site biotic interactions that may result in species' co-
304 occurrence patterns, we represented modelled species-to-species associations with a correlation
305 matrix describing the extent to which species pairs co-occurred more negatively or positively
306 than random chance, controlling for the effect of environmental covariates. For clarity, we
307 included only the twenty most abundant species in the main figure, but we provide a full species
308 co-occurrence matrix in the supplemental materials (S5, Supplementary Material). We then
309 conducted hierarchical cluster analysis on sites according to predicted species proportional
310 abundances from the HMSC model using Ward's criterion (Ward 1963), or "ward.D2" in R. We
311 mapped this dendrogram onto an abundance heatmap of the 20 most abundant species, and
312 colour coded the cells according to the species co-occurrence groupings (groups of species that
313 tended to positively co-occur amongst themselves).
315
316 **Results**
317 *Species abundance and distribution patterns*
318 The aim of this section is to describe basic biodiversity patterns across the seagrass
319 meadows. Specific results aimed at answering our three questions about metacommunity
320 processes follow in the subsequent sections. Seagrass-associated biodiversity varied spatially,
321 both within meadows and among meadows. We counted and identified 39 644 individuals
322 representing at least 58 species across all sites (Table 2). The metacommunity represented
323 multiple phyla, diet types (herbivores, detritivores, suspension feeders, and carnivores), and life
324 history strategies (brooding, broadcast spawning). A full summary of traditional diversity metrics

can be found in S3 of the Supplementary Material.

Several taxa were present at all sites (S4). These include the harpacticoid copepod *Porcellidium* sp., clams *Saxidomus gigantea* and *Clinocardium nutallii*, snails *Alvania compacta*, *Lacuna variegata* and *Lacuna vincta*, and the tanaid *Leptochelia* sp. Genera *Mytilus* and *Nereis* were also present at all sites, however the number of species represented remains unknown because we were unable to confidently identify these organisms to species. Ten species were present in at least one site in every region: the pycnogonid *Anoplodactylus viridintestinalis,* isopod *Idotea resecata,* gammarid amphipods *Monocorophium insidiosum, Ampithoe valida, Pontogeneia rostrata, Ampithoe dalli, Aoroides columbiae,* snails *Amphissa columbiana* and *Alia carinata,* and the limpet *Lottia pelta*.

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

In the full HMSC model with latent variables 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, this decrease in community similarity with spatial distance disappeared, and 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). The range of predicted similarities was greater at small distances than larger distances, indicating that spatial proximity between sites does not necessarily confer compositional similarity resulting from the exchange of individuals.

348 *Environmental niche filtering: variation partitioning*

349 Patterns of species' abundance and distribution are consistent with weak environmental niche
350 filtering along an abiotic axis, however the importance of environmental conditions varies across
351 species. The contribution of all 14 environmental variables, spatial distance and random effects
352 can be seen in Fig. 3. Out of all the measured environmental variables, the water quality
353 variables (all blue bars in Fig. 3) explained the largest proportion of variation across nearly all
354 species distributions Across all taxa, herbivore food availability (algae and detritus biomass) and
355 habitat structure (eelgrass shoot density and Leaf Area Index) had smaller influences on species
356 distributions than abiotic predictors (Fig. 3). The relative contribution of abiotic parameters to
357 species' abundance and distribution patterns also varied among species and major taxonomic
358 groups. For instance, site and region membership, food availability and habitat structure
359 generally explained less variation in gastropod distributions than gammarid amphipod
360 distributions. While the mean deviance explained (D^2_{adj}) across all species distribution models
361 was 0.29, individual values varied dramatically (0 to 0.82, Table 2), implying high variation in
362 the match between our model and the observed data. Accounting for this, the true fraction of
363 variation explained by environment was 22.4%, space explained 0%, the shared fraction between
364 environment and space was 0.04%, and the residual fraction of variation was 0.76 (S5,
365 Supplementary Material).

367

368 *Species interactions: Patterns in pairwise species co-occurrences*

369 We identified, post hoc, three main site-level species co-occurrence groupings across all
370 78 samples from 13 meadows and 4 regions after controlling for effects of environmental
371 covariates. The six species in the first group (*Nereis* sp. through *Caprella laeviuscula* in Fig. 4a)

372 henceforth known as the purple group, tended to positively co-occur with each other, and
373 negatively co-occur with non-members of this group. The second group- henceforth known as
374 the turquoise group- included *Caprella californica* through *Ampithoe valida,* and Harpacticoid 2
375 through *Lacuna* sp. *Spirorbis* sp. – the only truly sessile organism in the study, and having an
376 unusually high abundance at the HL site- did not fall into either category. *Alvania compacta* and
377 *Lacuna* sp. negatively co-occurred with members of the purple group and positively co-occurred
378 with members of the turquoise group, however their co-occurrence values tended to be weaker.
379
380 *Species interactions: Mapping co-occurrence groupings onto a site by species matrix*
381 Species associated with all three major co-occurrence groupings were present at all sites
382 (Figure 4b), indicating that these three sets of co-occurring species groups are not mutually
383 exclusive. It is their abundances that tend to negatively co-vary; when species of one assemblage
384 are abundant, species of the other assemblage are rare (Figure 4b). Fig. 4b shows clustering of
385 sites according to the identity of the most abundant species by co-occurrence colour grouping. At
386 JB, HL, CB, RA, EB, DC, and DK, members of the blue group tended to be most abundant,
387 whereas purple group members were most abundant at SS, SA, LH, GB and RB. IN was an
388 outlier site, as *Spirorbis* sp. was predicted to vastly outnumber other species. The emergent
389 species assemblages identified in Fig. 4a from the HMSC modeled distributions do not clearly
390 correspond with the four regions, or even with sites within regions, because the cluster analysis
391 did not group the sites by spatial proximity or regional membership (black symbols in Fig. 4b).
392
393 **Discussion**
394 By applying a new statistical framework, we have gained new insight into the

395 metacommunity processes structuring seagrass-associated biodiversity in coastal British
396 Columbia. We report evidence suggesting that both biotic interactions and environmental
397 conditions have an influence on seagrass meadow invertebrate composition. Spatial patterns of
398 species abundance and distribution suggest regional-scale dispersal patterns such that even
399 distant meadows may host the same species. However, these dispersal rates are clearly not so
400 high as to overwhelm the local meadow scale effects of biotic interactions and environmental
401 conditions. Below, we review these conclusions and the evidence to support them.
402
403 *Seagrass animal communities may experience moderate dispersal rates*
404 Dispersal limitation is likely not an important dynamic at the present spatial scale (an
405 extent of 1000 km) for the seven species that were present at all sites, and the additional ten
406 species that were not present at every site but present in all regions (S3, Supplemental Material).
407 It is noteworthy that these groups had representation from across phyla and life history traits
408 (polychaete worms, gammarid amphipods, snails, bivalves, tanaids), suggesting that no single
409 taxonomic group or dispersal life history strategy had a consistently longer dispersal range. Two
410 further pieces of evidence suggest that dispersal limitation resulting from isolation by distance
411 between sites is unlikely: (1) modelled community similarity did not decay with increasing
412 spatial distance (Fig. 2), and (2) hierarchical cluster analysis on species counts did not group
413 sites according to spatial proximity (Fig 4b). Species in our analysis that were only found at one
414 site or region may be dispersal limited for reasons other than isolation by distance. Those with
415 low site-level abundances (fewer than five individuals detected across all plots from a site) may
416 not disperse in appreciable numbers, and thus are not able to establish populations at all sites as
417 easily as more abundant species. This group included *Jassa marmorata*, *Orchomenella*

418 *recondita, Glycinde armigera, Neoamphitrite robusta, Lepidonotus squamatus,* and others (S3,
419 Supplementary Material).

420 Seagrass-associated epifauna are generally passive dispersers, with individuals colonizing
421 neighbouring sites either by larval transport in currents or by "rafting" on floating pieces of
422 seagrass, detritus, or macroalgae. Kelp rafts have been observed hosting benthic and epifaunal
423 invertebrates across phyla (echinoderms; peracarids; molluscs; annelids), and are important in
424 maintaining connectivity between coastal ecosystems (Wichmann et al. 2012). Thus, our results
425 suggest that this passive dispersal is sufficient to facilitate connectivity of these sites across the
426 region so that species distributions are not limited by dispersal. However, the specific pathway
427 the organisms travelled to arrive at these sites remains unknown. We used Euclidan distances in
428 our analysis, which preserved the rank order of distances across sites, but actual distances are
429 likely greater due to oceanographic circulation patterns (Kinlan and Gaines 2003, Mitarai et al.
430 2008, Treml et al. 2015). It is unclear whether distant sites share the same species because they
431 are (1) linked by direct dispersal; (2) indirectly linked by dispersal via unsampled "stepping-
432 stone" sites; or (3) were colonized by populations of the same species in a historical dispersal
433 event, but have not seen the exchange of individuals since.

434 We can also conclude that dispersal rates are not so high as to overwhelm the effects of
435 biotic interactions and environmental differences at the local scale; thus, dispersal rates are likely
436 intermediate. Our patterns of co-occurrence indicate that a subset of species was present in most
437 patches, but often at varying abundances. This may be reflective of weak mass effects, which
438 allow these populations to persist even if local conditions are not sufficient for them to persist
439 without immigration (Mouquet and Loreau 2003). Metacommunity theory predicts that such
440 weak mass effects are likely to be found whenever dispersal rates are not limiting (Thompson et

441 al. 2017), particularly in organisms that cannot control their own dispersal (Leibold and Chase
442 2017), such as seagrass-associated epifauna. While this is plausible for species found at several
443 sites, mass effects do not explicitly explain how the rare and singleton species in our dataset
444 persist (S4, Supplementary Material). Another possibility is that these ubiquitous species have
445 broader environmental niches, and species occupying only one or a few sites must have narrower
446 niches. While this scenario cannot be ruled out without an empirical test of species niches, it is
447 unlikely given our results on the contribution of the environment; the weak fit of the
448 environmental covariate model (Fig. 3), particularly for ubiquitous species for whom the model
449 fit was poor (e.g.'s *Aoroides* spp., *Spirorbis* sp., $D^2_{adj} = 0$), suggest that niche filtering as in
450 species sorting is not a large contributor to species distributions and abundances.
451
452 *Seagrass species show varying degrees of environmental niche filtering*
453 The fact that environmental covariates across space- and not space alone- are associated
454 with differences in community composition suggest some influence of the environmental niche
455 filtering (Fig. 2). For most species, environmental variables explained more variance in species
456 distributions than random effects. The most important environmental axis informing species
457 distributions is the abiotic one, with phosphates, salinity, and dissolved oxygen being the top
458 three parameters by percent variation explained (Fig. 4a). Accounting for the model fit (D^2_{adj} =
459 0.29), the fraction of variation explained by environment only was 0.22, leaving a great deal of
460 unexplained variation in the observed data at the whole-community level. However, in some
461 species for whom our model had high predictive power (including *Ampithoe dalli, Pontogeneia*
462 *rostrata, Caprella natalensis)*, environmental niche filtering may be important.
463 Species may respond to environmental heterogeneity across the metacommunity in one of

464 four ways. First, the most important abiotic environmental parameters (nitrates, phosphates, sea
465 surface temperature, and dissolved oxygen) may indirectly influence epifaunal presence and
466 abundance via primary productivity levels. This is possible despite the fact that primary biomass
467 explained less variance than the abiotic parameters in our HMSC model (Figure 3). Our measure
468 of primary biomass was based on a measurement at a single time point. But, our abiotic
469 parameters represent a four-year average (Assis et al. 2018) and so may provide a better
470 representation of what the organisms experience over their lifetimes. Second, sea surface
471 temperature, dissolved oxygen, and salinity may also directly influence optimal tolerance ranges.
472 For example, estuarine isopod *Idotea baltica* (a relative of *I. resecata* and *I. montereyensis* in our
473 study) shows significantly lower survival in prolonged exposure to higher water temperatures
474 and lower salinities than those experienced in their home habitat (Rugiu et al. 2018). Third,
475 temperature may influence the metabolic demands of the epifauna (most of which are grazers
476 such as some gammarid amiphopids, *Idotea* spp., *Lacuna* spp.), thereby influencing grazing rates
477 and altering the available primary biomass (O'Connor 2009). Decreased food resources may
478 result in competition, thereby influencing the number of individuals, number of species, or
479 identity of species a seagrass patch can host. It is not easy to determine whether individual
480 species responses to competitive stress influence whole-community structure, however doing so
481 may provide key information on how environmental conditions and interspecific interactions
482 interact to influence biodiversity.

483 Environmental characteristics of temperate eelgrass meadows fluctuate with the seasons,
484 likely driving changes in the relative abundance of epifaunal taxa throughout the year
485 (Wlodarska-Kowalczuk et al., Markel et al. 2017). As a result, the epifaunal metacommunity
486 may not evince these abundances or distributions year-round; sampling several times in one year

487 to capture temporal changes in community structure would give a clearer picture of species
488 responses to environmental conditions, as well as biotic interactions and dispersal events.
489
490 *Do species co-variance patterns reveal possible biotic processes influencing community*
491 *assembly?*
492 Our analysis revealed three distinct species co-occurrence groupings; the first two
493 (turquoise and purple groups in Fig. 4a) tended to positively co-occur with members of the same
494 group, and negatively co-occur with members of the other group. The third group, which
495 included only *Spirorbis* sp. in our top 20 most abundant species (Fig. 4a), did not show strong
496 negative or positive co-occurrence with any other species. These distinct groupings suggest that
497 species co-variance patterns are not random, but rather are shaped by interactions between
498 species and the order in which species colonize a site (i.e. priority effects, Fukami et al. 2016).
499 This strong signal of biotic interactions, along with the insignificant contribution of spatial
500 distance, structure and the variable signal of environmental filtering leads us to suggest the
501 importance of biotic processes in structuring the epifaunal metacommunity. There are several
502 possible biotic interactions taking place within the epifaunal metacommunity/communities. The
503 majority of species in the present study are known to be herbivores or detritivores, and thus may
504 be competing for primary biomass, however it is difficult to empirically test whether these
505 resources are limiting such that they result in competition. Some species are filter feeders, and
506 thus may not be competing for food resources but possibly space.
507 The original patch dynamic metacommunity paradigm describes a scenario where
508 competition-colonization tradeoffs facilitate regional coexistence. This dynamic is impossible to
509 exclude based on observational data alone, however one might expect to see mutually exclusive

510 occurrences among species pairs, or a checkerboard pattern (Diamond 1975). Patch-dynamics in
511 this system are unlikely given the observed co-occurrence patterns; species pairs that had
512 negative mean co-occurrence values often still co-existed, with one of the two showing higher
513 abundance at a given site. Still, experimental evidence is required to completely rule out the
514 possibility of competition-colonization tradeoffs contributing to diversity patterns in this system.
515 The community structure of a given site with a certain regional pool and local
516 environmental conditions may have multiple stable states depending on the arrival order of
517 species and the structure of competitive interactions; this phenomenon is known as priority
518 effects (Diamond 1975, Waters et al. 2013, Fukami et al. 2016, Ke and Letten 2018). The fact
519 that we found strong groupings of species with consistent patterns of negative and positive co-
520 occurrence (Fig. 4a) as well as the three different community groupings (Fig. 4b) is indicative of
521 priority effects. We hypothesize that following the establishment of a new meadow or a
522 disturbance event, the first few species to colonize a seagrass meadow greatly determine the
523 success of subsequent colonizers, regardless of the environmental conditions in the meadow. The
524 high abundance, and negative co-occurrence between *Nereis* sp. and *Lacuna* sp. suggests that
525 these species may be key determinants of community structure (Fig. 4b). As an illustrative
526 example, *Nereis* sp. might have been among the first to colonize sites SS, SA, LH, GB, and RB
527 (Fig. 4b). At HL in particular, hundreds of individuals were found covering a single blade of
528 eelgrass, limiting the amount of available epiphytic microalgae for grazers to feed on. At sites
529 RA, LH, and SS, *Nereis* sp. may have been the first to arrive, thereby gaining a numerical
530 advantage over *Lacuna* sp. and others. The disturbance regime, colonization history and age of
531 the eelgrass meadows themselves likely matter greatly if priority effects are an important
532 dynamic influencing community assembly. A key piece of information that would confirm the

533 nature of biotic processes is whether species from different co-occurrence groupings compete
534 more strongly with each other than with themselves (Ke and Letten 2018), but this would require
535 an experimental test.

536 The positive co-occurrence patterns between certain species may be a result of positive
537 interspecific interactions. An example is *Orchomenella recondita*: a gammarid amphipod found
538 living in the gastrovascular cavity of the anemone *Anthopleura elegans* (Carlton 2007). This
539 species was only recorded at the SS site, and specifically only found in quadrats where *A.*
540 *elegans* was collected with the seagrass shoots.

541 A potentially important biotic interaction missed in our analysis is the trophic effect of
542 predators. Field experiments have demonstrated that changes in predation pressure by fish,
543 shorebirds, and predaceous invertebrates can shift seagrass-associated epifaunal assemblages in a
544 matter of weeks (Amundrud et al. 2015; Huang et al. 2015; Lewis & Anderson 2016).
545 Differential predation pressure, or composition of predators could also lead to the pattern of
546 diverging community composition that we have witnessed (Guzman et al. 2018). Further study is
547 required to determine the degree to which these seagrasses associated communities are structured
548 by trophic interactions.

549

550 **Conclusion**

551 Our study demonstrates how focusing on characterizing the processes of dispersal,
552 environmental heterogeneity, and biotic interactions can allow us greater insight into natural
553 metacommunities than previous approaches. Furthermore, we have shown how this can be
554 accomplished by combining the new HMSC framework with traditional multivariate statistics.
555 This approach provided insight into seagrass epifauna community assembly that would not have

556 been possible using traditional methods such as variation partitioning or EMS. We have
557 generated new hypotheses about the relative importance of dispersal, environmental
558 heterogeneity, and biotic interactions in shaping metacommunities. Manipulative experiments
559 and ongoing monitoring are now needed test and refine these hypotheses. Given our hypothesis
560 that priority effects play a strong role in structuring these communities, experiments are now
561 needed to test whether the composition of seagrass meadows depends on the order of species
562 colonization (Fukami et al. 2016). We could also test our hypothesis that species sorting is
563 relatively weak in this system by measuring fitness differences along the most important
564 environmental axes. Testing our hypothesis that these communities are not dispersal limited is
565 more challenging, but could be studied using particle tracking models (Treml et al. 2008) which
566 simulate the path travelled by epifauna among sites based on ocean currents and estimates of
567 pelagic larval duration. A population genetic analysis could also ground-truth the extent and time
568 scale of dispersal by determining whether populations share alleles, or have been isolated for
569 some time. All together, we suggest that a modern metacommunity perspective, combined with
570 new advances in our statistical toolbox provide a powerful and relevant framework for studying
571 natural ecosystems.

572 Our findings are particularly relevant to conservation initiatives that aim to maintain
573 diversity or prevent species loss in seagrass meadows. Specifically, identifying the contribution
574 of dispersal, environmental conditions, and interspecific interactions allows the conservation
575 practitioner to better predict how habitat fragmentation and climate change might influence
576 diversity. Epifaunal invertebrate communities in British Columbia seagrass meadows may be
577 quite resilient, given our findings that several species across phyla and functional traits are
578 consistently present and abundant across a fairly large spatial scale and environmental conditions

make a minimal contribution to species distributions.

Acknowledgements

This research was financially supported by a UBC Zoology SURE Grant to KAS, and funding from NSERC Discovery Grants, Canadian Foundation for Innovation (CFI). PLT is supported by NSERC and Killam postdoctoral fellowships. This research 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). Special thanks to Michelle Paleczny and Gulf Islands National Park Reserve of Canada for boat and laboratory space during their annual eelgrass sampling campaign. We acknowledge staff, students, and volunteers from Pacific Rim National Park Reserve for their assistance collecting eelgrass samples in Clayoquot Sound. We also sincerely thank Rob Underhill and Abbie Sherwood for support and assistance with the GB site sampling, Emily Adamczyk and Minako Ito for assistance with sampling and processing eelgrass from the RB, SA, and DC sites, to Giordano Bua for assistance with processing the Clayoquot Sound samples, and to Ariane Comeau for assistance with epifaunal identification of the Barkley Sound samples. John Cristiani created the map in Figure 1. Matt Whalen pulled the Bio-ORACLE dataset. We are grateful to Coreen Forbes and Matt Whalen for their constructive comments on the manuscript drafts.

Literature Cited

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

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

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

Brown, B.L., Sokol, E.R., Skelton, J., Tornwall, B. 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 (4th ed). University of California Press.

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

Diamond, J.M. 1975. Assembly of species communities - In: Cody, ML and Diamond, HM. (eds) Ecology and evolution of communities. Belknap Press: 342-444.

Duffy, J.E. 2006. Biodiversity and the functioning of seagrass ecosystems. Marine Ecology Progress Series 311: 233–250.

Duffy J.E, Reynolds, P.L., Boström, C., Coyer, J.A., Cusson, M., Donadi, S., Douglass, J.G., Eklöf, J.S., Engelen, A.H., Eriksson, B.K., 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., Mordecai, E. A., Ostling, A. 2016. A framework for priority effects. Journal of Vegetation Science 27: 655–657.

Gaines, S., Roughgarden, J. 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.

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

Guillaume Blanchet, F., with contributions from Gleb Tikhonov and Anna Norberg (2018). HMSC: Hierarchical Modelling of Species Community. - R package ver. 2.1-2.

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

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

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

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

Jaschinski, S., Sommer, U. 2008. Functional diversity of mesograzers in an eelgrass-epiphyte system. Marine Biology 154: 475–482.

Ke, P. J., Letten, A. D. 2018. Coexistence theory and the frequency- dependence of priority effects. Nature Ecology and Evolution, 1–7.

Kinlan, B.P., Gaines, S.D.. 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.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., et al. 2004. The metacommunity concept: A framework for multi-scale community ecology. Ecology Letters 7: 601–613.

Leibold, M.A., Mikkelson, G.M. 2002. Species Turnover , and Boundary Clumping: Elements of Meta-Community Structure. Oikos 97: 237–250.

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

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

Loreau, M., Mouquet, N., Gonzalez, A. 2003. Biodiversity as spatial insurance in heterogeneous landscapes. Proceedings of the National Academy of Sciences 28: 12765-12770.

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

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

Mouquet, N., and Loreau, M. 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., Heck, K. L., van Montfrans, J. 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., Tikhonov G., Norberg A., Guillaume Blanchet F., Duan L., Dunson D., Roslin T., Abrego N. 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., Higgins, C.L., Willig, M.R. 2010. A comprehensive framework for the evaluation of metacommunity structure. Global Ecology and Evolution 19: 908–917.

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

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

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

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

Treml, E. A., Ford, J. R., Black, K. P., Swearer, S. E. 2015. Identifying the key biophysical drivers, connectivity outcomes, and metapopulation consequences of larval dispersal in the sea. Movement Ecology 3:17

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

5: 5203–5215.

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

Waters, J. M. et al. 2013. Founder takes all: density-dependent processes structure biodiversity. Trends in Ecology and Evolution 28: 78-85.

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

Williams, S. L., Heck, K. L. Seagrass Community Ecology. In Marine community ecology 317-338.

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

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

Yamada, K., Tanaka, Y., Era T., Nakaoka, M. 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.

718 **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 patterns of species' abundance and distribution suggest environmental niche filtering, and if so along which environmental axis?	HMSC model: 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?	HMSC model: Correlation matrix of pairwise species co-occurences, after controlling for the effect of space and environment	Fig. 4a & 4b

719

720

721

722

723 **Table 2.** Species names, taxonomic groupings, and proportion of deviance (predictive power).
724 Numbers correspond to the axes in Figure 3. The top twenty species with highest predicted
725 abundances according to the HMSC model are bolded.

726

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

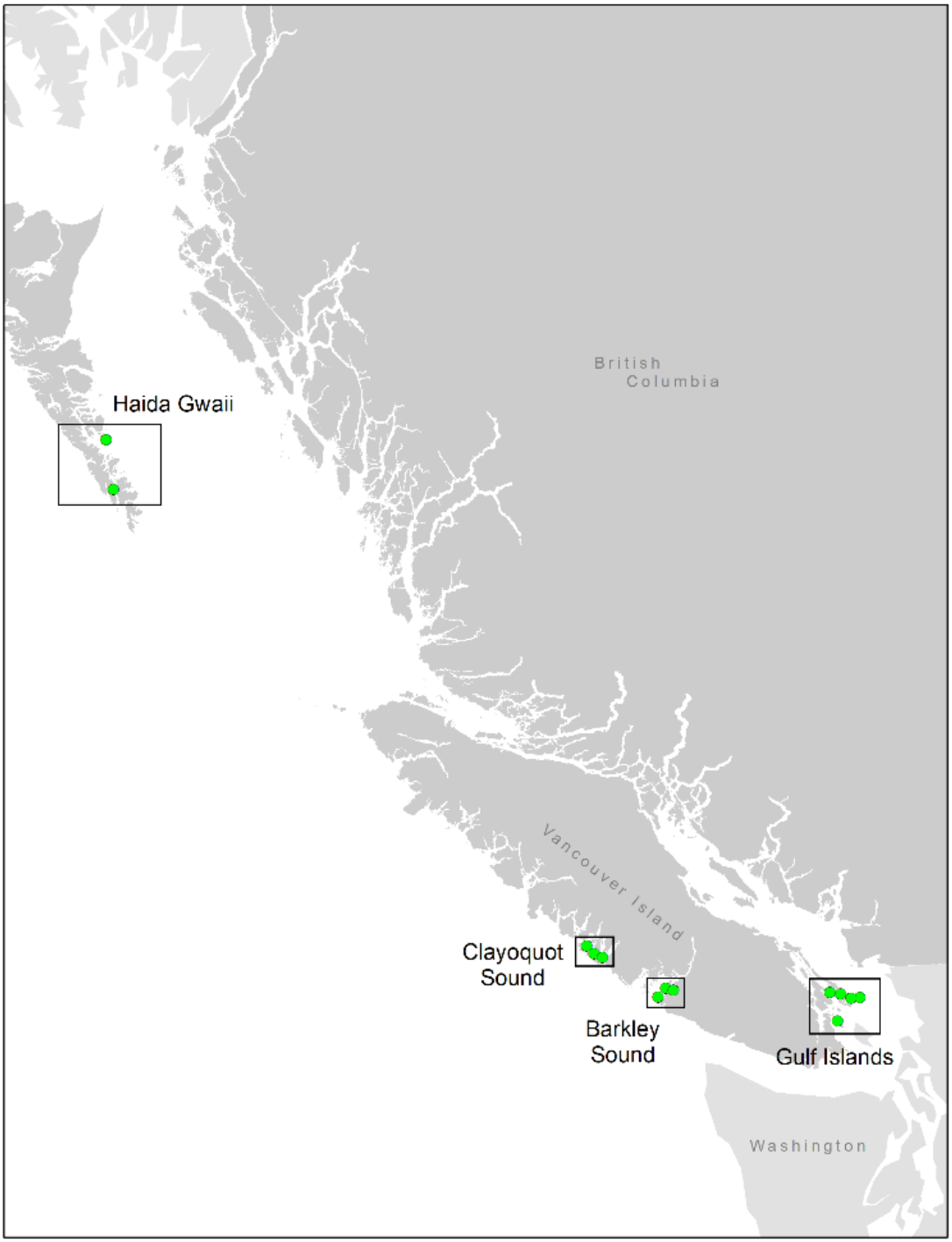


Figure 1. Map of coastal British Columbia showing eelgrass field sites in Haida Gwaii (HL, RA), Clayoquot Sound (DK, EB, IN), Barkley Sound (SA, DC, RB), and the Southern Gulf Islands (GB, JB, CB, LH, SS).

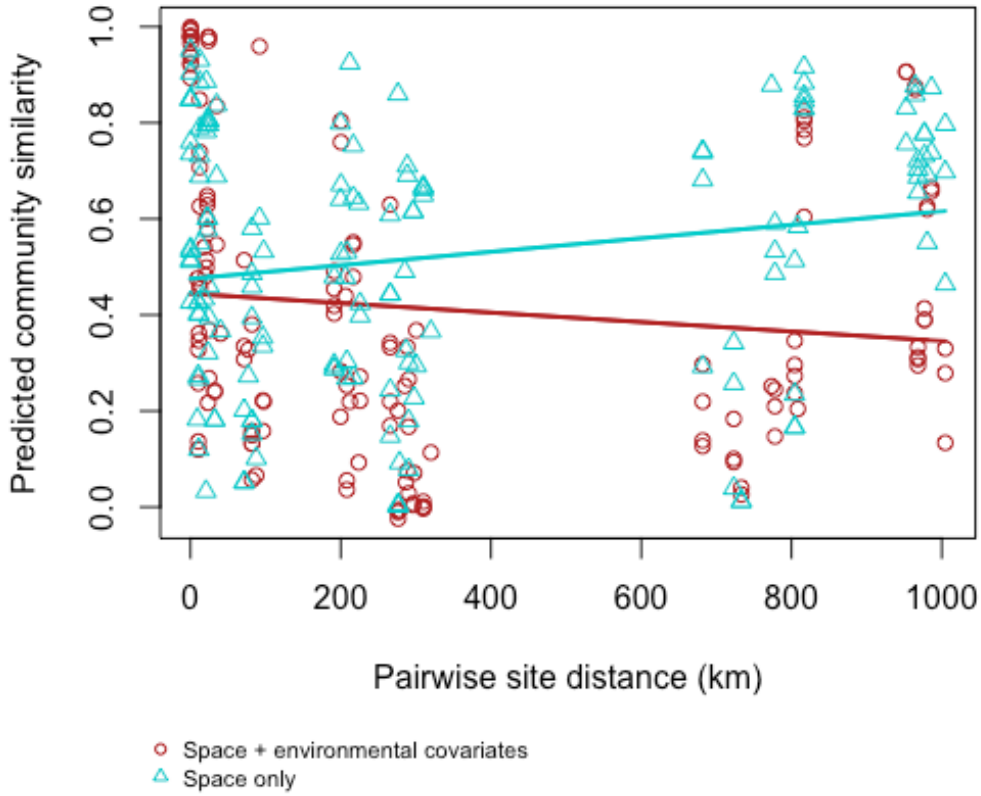


Figure 2. HMSC model predictions of community similarity relative to the spatial distance (Euclidean) between communities. The red circles describe the result based on the full model. The blue triangles show a prediction wherein all environmental parameters (abiotic, biotic, and spatial) have been standardized to their mean values, and thus they show predicted community similarity as a result of spatial distance only.

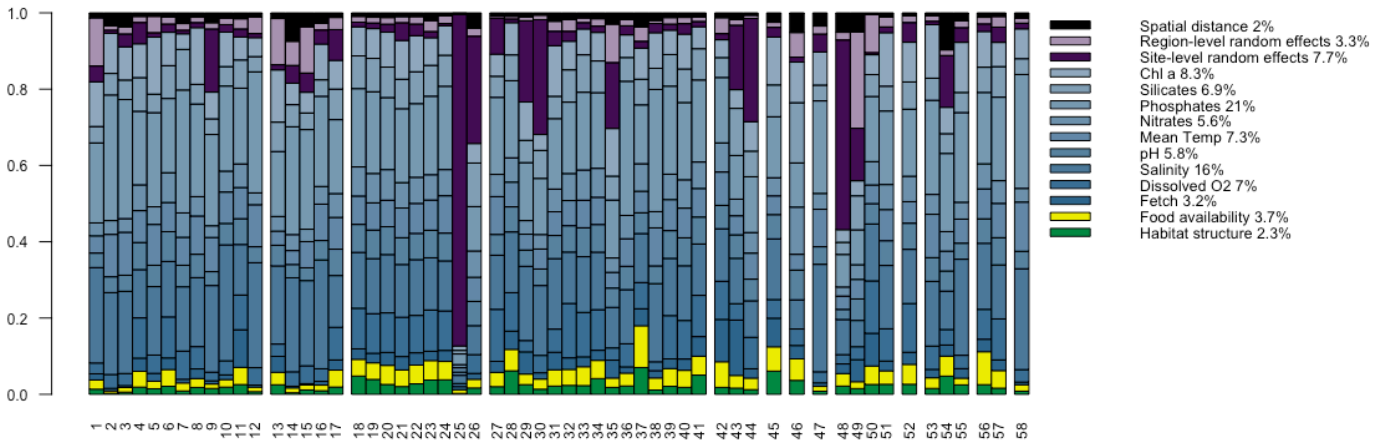


Figure 3. Variance partitioning of fixed (environmental variables, spatial distance) and random effects across all species (numbers correspond to Table 2) in the HMSC model. Purple cells represent site and region-level random effects. Blue cells represent abiotic water quality parameters. Yellow cells 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 Table 2. 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 model fits varied (see Table 2). The mean D^2_{adj} across all species was 0.29.

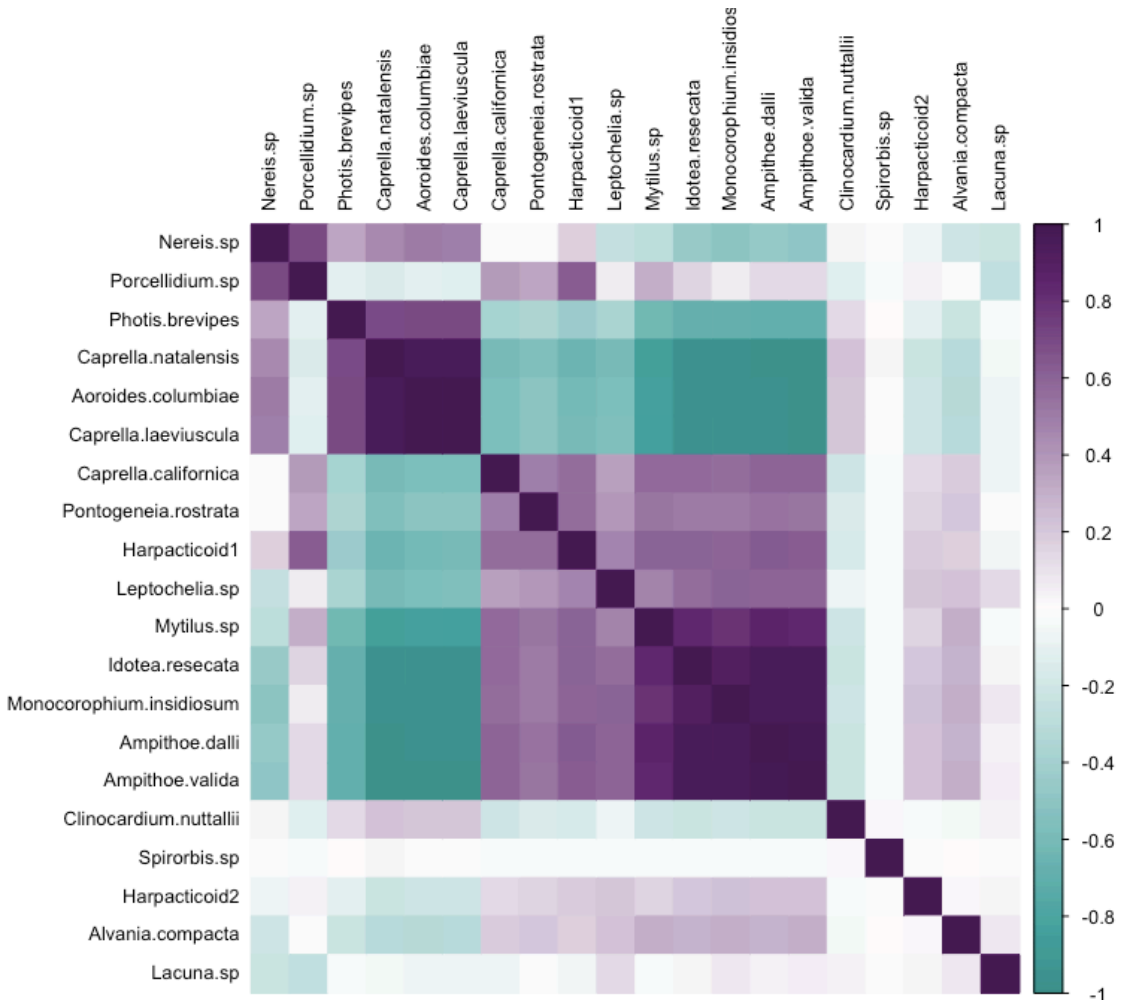


Figure 4a. A correlation plot showing modelled site-level co-occurrence of species pairs in the 20 most abundant species. Purple cells represent positively co-occurring species pairs, and turquoise cells represent negatively co-occurring species pairs. Species names along both axes are ordered according to the output of hierarchical clustering with Ward's criterion on pairwise co-occurrence values. A full version of this figure with all 58 species is shown in S6, Supplementary Material.

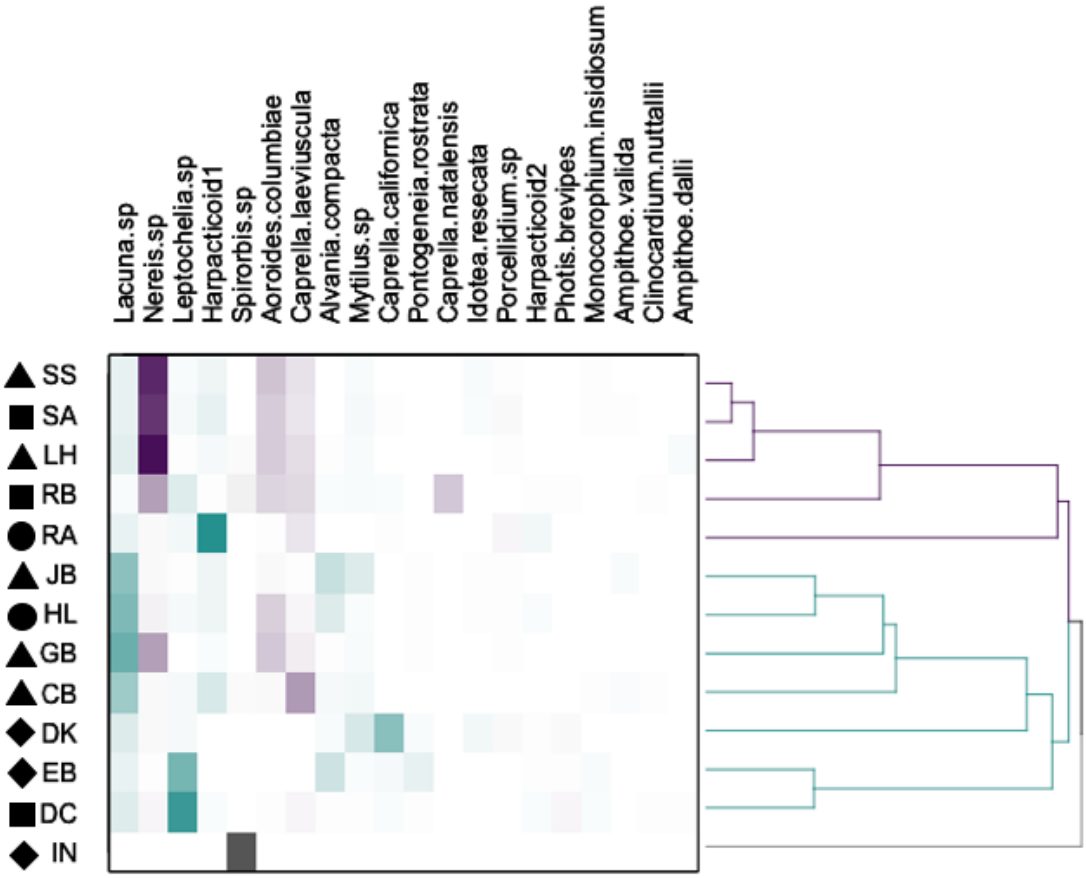


Figure 4b. Heatmap and cluster dendrogram depicting modelled (HMSC) species abundance and compositional similarity across sites. The dendrogram was produced using hierarchical cluster analysis with Ward's criterion on predicted species proportional abundances. The heat map shows the twenty species with 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: turquoise group, purple group, and 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. A version

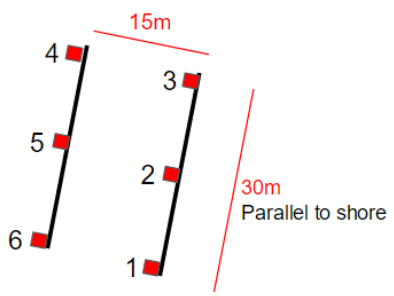
788 of this figure included all species is shown in S7 (Supplementary Material).

789

790

Regional connectivity, local environmental conditions, and biotic interactions jointly structure a temperate seagrass epifaunal metacommunity

Supplementary material



S1. Schematic diagram of eelgrass field sampling regime. Black lines represent transect lines, and red boxes represent 0.25 x 0.25 m quadrats.

S2. Equations for D^2_{adj} calculations

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

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

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

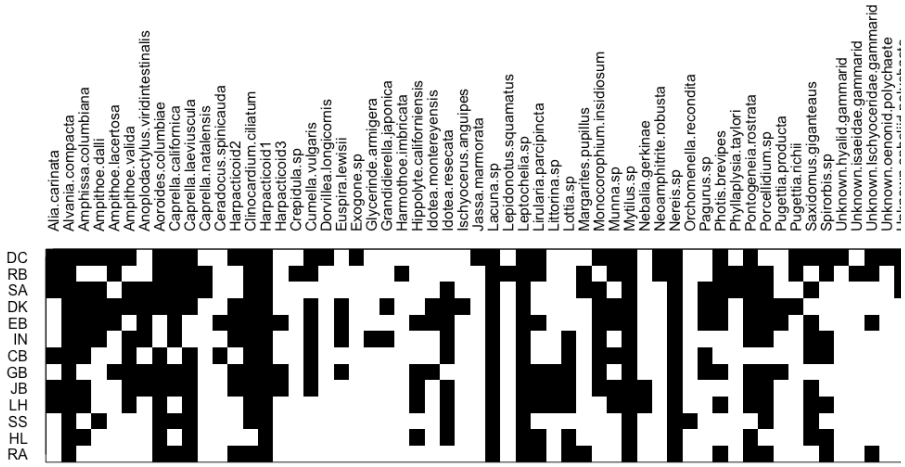
Adjusted D^2 (Guisan and Zimmerman 2000) was calculated by the equation:

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

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 Supplementary Material). Negative D^2_{adj} values (those which had greater residual deviance than null deviance) were set to 0.

S3. Table summarizing species richness and relative abundance at all sites

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



S4. Site-by-species presence-absence matrix of all 58 invertebrate species in the study listed in alphabetical order. Black cells indicate species presences.

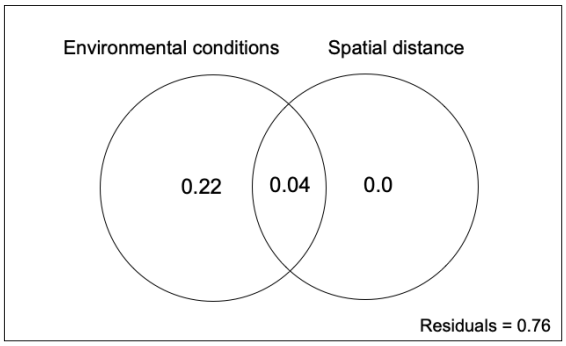
S5.a Calculation for type III variation partitioning result

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,

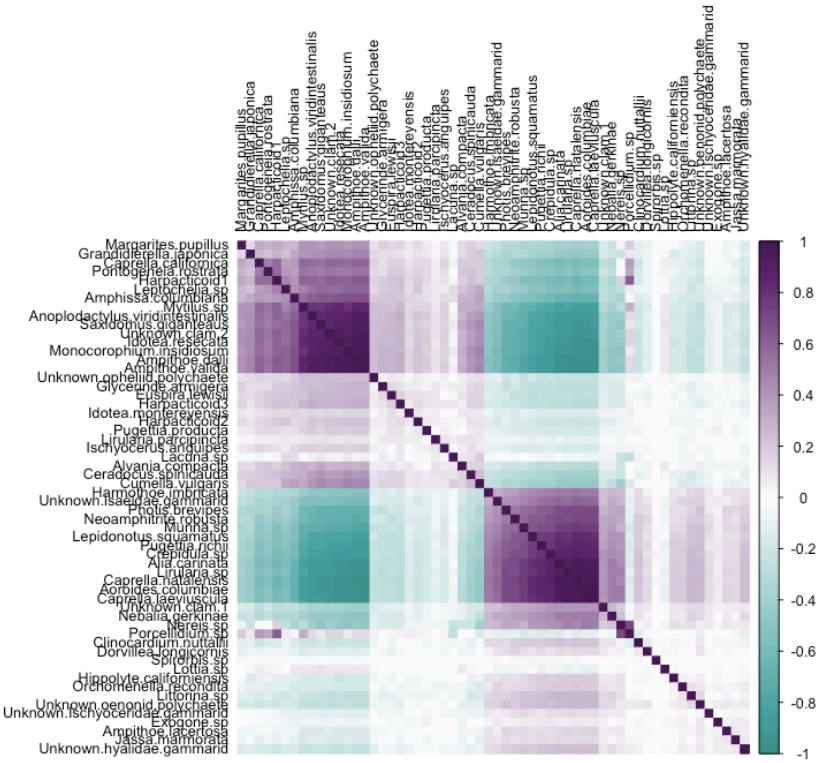
Fraction *ab* (pure environment + shared fraction) = D^2_{adj} *m1*
Fraction *bc* (pure space + shared fraction) = D^2_{adj} *m2*
Fraction *abc* (pure environment + shared fraction + pure space) = D^2_{adj} *m3*
Fraction *a* (pure environment) = *abc* - *bc*
Fraction *c* (pure space) = *abc* - *ab*
Fraction *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 S5.b.

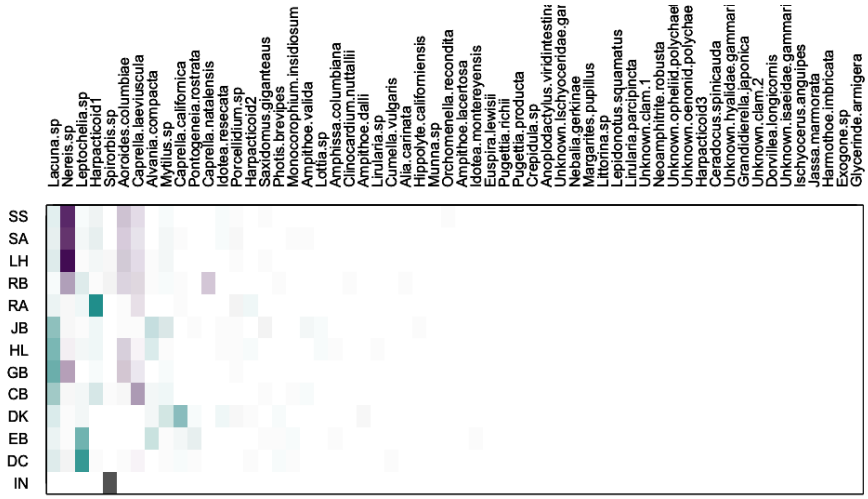
S5.b Venn diagram summarizing the fractions of variation explained by environment only, space only, and the shared fraction explained by environment and space.



Formatted: Font:(Default) Times New Roman



S6. 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 on pairwise co-occurrence values. This figure is a supplement to Fig. 4a in the main text.



S7. Heatmap 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. 3b, with turquoise cells representing members of the turquoise assemblage, purple cells indicating members of the purple 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.