Beyond a single patch: local and regional processes explain diversity 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: 241 words
Main text: 4756 words
Number of references: 56
Number of figures: 3

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 jointly 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 dispersal. Here, we use the Hierarchical Modelling of Species Communities (HMSC) statistical framework to assess the contributions of metacommunity processes to community structure in an important coastal habitat, seagrass meadows, across approximately 1000 km along the coast of British Columbia. We parametrized a hierarchical joint species distribution model describing the relationship between environmental conditions, spatial distance between sites, and random effects on epifaunal invertebrate species distributions across thirteen eelgrass (*Zostera marina* (L.)) meadows. Compositional similarity did not decrease at large spatial distances suggesting sites are well-connected by dispersal, but observed species distribution patterns suggested 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 and the contribution of processes other than niche fitering. Strong structuring in pairwise species co-occurrences 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, which in turn can inform targeted experiments, monitoring, and management decisions for preserving diversity in this important ecosystem.

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

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, Leibold and Chase 2017). Determining the relative importance of these assembly processes is critical to 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 abundance within metacommunities.

Many marine ecosystems that host high biodiversity occupy spatially structured habitats (e.g. coral reefs, kelp forests, seagrass meadows), and marine species have dispersing life histories that connect populations in distinct habitat patches. Marine systems are characterized by strong species interactions (Berlow 1999, Sala and Graham 2002), and the importance of dispersal-driven population dynamics has long been recognized (Levin and Paine 1974, Gaines and Roughgarden 1985). These combined roles of dispersal, the local environment, and biotic

92 interactions suggest that marine diversity patterns may reflect processes characteristic of
93 metacommunities (Guichard et al. 2003). Despite this, metacommunity dynamics have only been
94 explicitly investigated in a few marine systems (Guichard et al. 2003, Yamada et al. 2014,
95 Valanko et al. 2015).

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

109 To date, evidence for the importance of metacommunity dyamics in seagrass-associated
110 communities is based on two different quantitative approaches (Yamada et al. 2014, Whippo et
111 al. 2018). Using variation partitioning on constrained ordinations (Borcard et al. 1992, Cottenie
112 2005), Yamada et al. (2014) found that the relative contribution of space and environment varied
113 among invertebrate functional groups. However, the variation partitioning could not distinguish
114 between the possible metacommunity dynamics at play owing to the ambiguous statistical

115 significance of the results (Yamada et al. 2014). Simulation studies have shown that this method
116 can over- or under-estimate spatial and environmental components of variation, and produce
117 inflated R^2 values depending on the user's statistical choices (Gilbert and Bennett 2010,
118 Tuomisto et al. 2012). The other approach, known as elements of metacommunity structure
119 (EMS), seeks idealized structures in species occurrence patterns (Leibold and Mikkelson 2002,
120 Presley et al. 2010). While useful for visualizing species distribution patterns across a region,
121 EMS does not robustly elucidate the processes contributing to these patterns (Brown et al. 2017).
122 These limitations, along with the revised goal of individually assessing the contribution of
123 dispersal, niche filtering, and biotic interactions to diversity patterns, invite novel methodological
124 approaches. The Hierarchical Modelling of Species Communities statistical framework (HMSC,
125 Ovaskainen et al. 2017), shows promise for achieving this goal. HMSC uses Bayesian inference
126 to parametrize a hierarchical joint species distribution model, and allows the partitioning of
127 variation in species distributions due to the environment, spatial distance, and random effects,
128 allowing inference of patterns at the whole-community level as well as in individual species.
129 Additionally, it specifies a way to visualize positive and negative pairwise species co-
130 occurences, which can inform hypotheses on possible species interactions. For these reasons,
131 HMSC achieves what researchers aimed to achieve with the previous methods, and much more.
132 Here we test the hypothesis that seagrass-associated epifaunal invertebrate communities are
133 jointly structured by the spatial distances between meadows (a proxy for inferring dispersal
134 dynamics), local environmental conditions, and biotic interactions. We follow the HMSC
135 framework to answer the following questions: (1) Is the metacommunity characterized by
136 dispersal limitation, intermediate dispersal, or high dispersal? (2) Do species distribution patterns
137 suggest environmental niche filtering, and if so along which environmental axes? (3) Do species

138 co-variance patterns suggest possible biotic interactions influencing community assembly? The
139 order in which these questions are posed is significant. For example, evidence for dispersal
140 limitation would affect how we interpret observed species co-occurrence patterns.

141 We address these three questions with a biodiversity survey of invertebrates in 13 *Zostera*
142 *marina* (L.) meadows in British Columbia, Canada. We find evidence of intermediate dispersal
143 rates. We also find variable species responses to environmental conditions, and large overall
144 variation unexplained by environment and space in our model. Strong species' covariance
145 patterns led us to hypothesize possible biotic interactions influencing community structure.

146 **Methods**

147 *Study sites and field sampling of eelgrass and associated epifauna*

148 We sampled epifaunal biodiversity in 13 eelgrass *Zostera marina* (L.) meadows across a
149 1000 km spatial scale from late June through early August in 2017. The sites were located within
150 four coastal areas: Haida Gwaii within Gwaii Haanas National Park Reserve, National Marine
151 Conservation Area Reserve and Haida Heritage Site; southern Clayoquot Sound (one site was
152 within Pacific Rim National Park Reserve); Barkley Sound on the West coast of Vancouver
153 Island; and the Southern Gulf Islands within Gulf Islands National Park Reserve (Fig. 1a).

154 We sampled three sites in Barkley Sound (SA, DC, RB) with SCUBA because the
155 meadows reached a depth of 5 m. All other sites were sampled by wading or snorkeling at low
156 tide. We used a nested sampling regime with 0.25 m x 0.25 m quadrats to collect eelgrass,
157 epifaunal invertebrates, detritus, and macroalgae. Six quadrats were set in a 15 m x 30 m array
158 (Appendix S1: Fig. 1); the array was placed in the middle of the meadow to avoid edge effects
159 where possible. We removed all above-ground biomass (eelgrass, detritus, macroalgae) within
160 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-ORACLE, including surface chlorophyll-*a*, dissolved oxygen, nitrates, silicates, phosphates, pH, salinity, and temperature. Bio-ORACLE is a database of marine environmental 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, and thus had distinct values for the aforementioned variables. We also calculated fetch, which quantifies site exposure.

For each quadrat sampled, we measured five biotic attributes summarizing habitat structure and food availability: eelgrass leaf area index (LAI) and shoot density, and eelgrass dry mass, algal dry mass, and detritus dry mass. To quantify LAI, we measured leaf length, width, and blade number in five haphazardly-chosen shoots from each quadrat, and multiplied the average blade area (m^2) per shoot by the shoot density (number of shoots per m^2). We separated and dried eelgrass, detritus, and macroalgae in a desiccator oven (60° C for 48 hours).

Quantifying epifaunal diversity

We rinsed the eelgrass blades to remove epifaunal invertebrates, and passed samples through a 500 μm sieve. Invertebrates were preserved in 95% EtOH for identification with light microscopy (10x magnification). We identified epifauna to the lowest taxonomic level possible using keys in the Light and Smith manual (Carlton 2007) and Kozloff (1996).

Analysis

We followed the methodology of Ovaskainen et al. (2017) to construct a hierarchical joint species distribution model describing the relationship between environmental variables,

184 spatial distance, species co-occurrences and species distributions using the HMSC package
185 (Guillaume Blanchet et al. 2018) in R (version 3.4.2; R Development Core Team 2014). The
186 model is comprosed of two linear predictor terms: one describing fixed effects of environment
187 and space, and one describing random effects that cannot be attributed to environment or space
188 (Equations 2 & 3 in Ovaskainen et al. 2017). The random-effects linear predictor term describes
189 random variation in individual species' distributions as well as covariation for every species pair.
190 We used Markov chain Monte Carlo to parametrize the model. Species' responses to
191 environmental variables were assumed to follow a multivariate Gaussian distribution. We
192 assumed a Poisson distribution for species count data and used the same priors as in the
193 examples in Ovaskainen et al. (2017). The Markov Chain included 100 000 iterations, a burn-in
194 length of 1000 iterations, and thinned to retain every 100th sample of the posterior distributions.
195 We confirmed that parameters were well mixed by visually inspecting trace plots. We tested the
196 model's performance by calculating explained deviance (D^2_{adj}) as described in Appendix S1: Eq.
197 S1. D^2_{adj} is comparable to R^2_{adj} in least square regression analysis (Guisan and Zimmermann
198 2000); in this context, it describes the model fit where a value of 1 means the model perfectly
199 predicts the observed data. Our calculations yielded a D^2_{adj} value for each individual species
200 distribution model. The mean D^2_{adj} across all species gave a metacommunity-level value of
201 predictive power.
202 To answer our three main questions about the various metacommunity processes, we
203 visualized the parametrized model in several ways (Appendix S1: Table S1). We investigated
204 three possible dispersal conditions: dispersal limitation, intermediate dispersal, and high
205 dispersal. To investigate dispersal limitation at large spatial scales, we plotted distance-decay of
206 similarity by randomly drawing 200 pairwise distances between sampling plots, and plotting

207 predicted community similarity against distance. Distance-decay of similarity could reflect
208 dispersal limitation due to prohibitively far travel distances, or differences in environmental
209 conditions with increasing spatial scale. To distinguish these possibilities, we compared
210 predicted community similarity as a function of distance and environmental covariates (the full
211 HSMC model), and predicted community similarity using an alternative model in which the
212 effect of the environment was removed (all environmental covariates were set to their mean
213 values). To distinguish between intermediate and high dispersal, we visually inspected the
214 occurrence data for signatures of regional homogenization by high dispersal. We used niche
215 filtering strength (estimated by D^2_{adj}) as another line of evidence to distinguish intermediate and
216 high dispersal: theory predicts that intermediate dispersal (dispersal sufficiency) allows species
217 to colonize sites where local environmental conditions are optimal for growth and reproduction,
218 whereas high dispersal (dispersal surplus) overrides this (Mouquet and Loreau 2003).
219 To estimate niche filtering strength, we conducted variation partitioning of the fixed and
220 random effects (Ovaskainen et al. 2017). This summarized the relative contribution of
221 environmental variables, spatial distance, and random effects to each species distribution model;
222 we visualized this in a stacked bar plot. We then estimated the pure and shared fractions of
223 variation explained by environment and space, as well as the unexplained fraction using type III
224 sum of squares (Borcard et al. 1992), but using the HMSC model instead of constrained
225 ordination. To do so, we followed the equations in Appendix S1: Eq. S2.
226 To identify species co-occurrence patterns resulting from possible biotic interactions, we
227 represented modelled pairwise associations in a correlation matrix describing the extent to which
228 species pairs co-occurred more negatively or positively than by random chance, controlling for
229 the effect of the environment and space. We then conducted hierarchical cluster analysis on sites

230 according to predicted species abundances from the model using Ward's criterion (Ward 1963).
231 We mapped the cluster dendrogram onto an abundance heatmap of the 20 species in our study
232 with the highest mean abundance, and colour coded the cells according to the species co-
233 occurrence groupings (groups of species that tended to positively co-occur amongst themselves).
234 Full versions of these figures with all species are provided in Appendix S1: Fig. S4 and Fig. S5.

235 **Results**

236 *Species abundance and distribution patterns*

237 We counted and identified 39 644 individuals representing at least 58 species across the
238 region (Appendix 1: Table S2). The represented taxa span multiple phyla, diet types (herbivores,
239 detritivores, suspension feeders, carnivores), and life history strategies (brooding, broadcast
240 spawning). Mean quadrat-level species richness was 13 species, and mean site-level richness was
241 16 species. A full summary of diversity metrics can be found in Appendix S1: Table S3.

242 Several taxa were present at all sites (Appendix S1: Fig. S2). These include harpacticoid
243 copepod *Porcellidium* sp., bivalves *Saxidomus gigantea* and *Clinocardium nutallii*, snails
244 *Alvania compacta*, *Lacuna* spp., and the tanaid *Leptochelia* sp. Genera *Mytilus* (a mussel) and
245 *Nereis* (a polychaete worm) were also represented at all sites, however we were unable to
246 confidently identify these organisms to species. Ten species were present in at least one site in
247 every region: pycnogonid *Anoplodactylus viridintestinalis,* isopod *Idotea resecata,* gammarid
248 amphipods *Monocorophium insidiosum, Ampithoe valida, Pontogeneia rostrata, Ampithoe dalli,*
249 *Aoroides* spp*.,* snails *Amphissa columbiana* and *Alia carinata,* and limpet *Lottia pelta*.

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

251 In the full HMSC model summarizing species' responses to environment and spatial
252 distance, predicted community similarity decreased with increasing distance between sites (Fig.

253 1b, red points and line). However, community similarity showed a slight increase with distance
254 when the effect of environmental variables was removed from the model (all environmental
255 variables set to their mean value), suggesting that spatial distance alone did not correspond with
256 differences in invertebrate compositional similarity (Fig. 1b, blue points and line).

257 *Environmental niche filtering: variation partitioning*

258 Patterns in abundance and distribution were consistent with weak environmental niche
259 filtering along an abiotic axis; however, the importance of environmental conditions varied
260 across species. The relative contributions of all 14 environmental variables, spatial distance and
261 random effects in the HMSC model can be seen in Fig. 2. Out of all environmental variables, the
262 water quality variables (blue bars in Fig. 2) explained the largest proportion of variation across
263 nearly all species distributions. Herbivore food availability (seagrass, algae and detritus biomass)
264 and habitat structure (eelgrass shoot density and LAI) had smaller influences on species
265 distributions (Fig. 2). While the mean deviance explained (D^2_{adj}) across all species distribution
266 models was 0.29, individual values varied (0 to 0.82, Appendix 1: Table S2), implying high
267 variation in model fit. Accounting for this, the true fraction of variation explained by
268 environment was 22%, space explained -3%, the shared fraction between environment and space
269 was 4%, and the residual fraction of variation was 76% (Appendix S1: Fig. S3).

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

271 We identified, post hoc, three main site-level species co-occurrence groupings after
272 controlling for fixed effects. Members of the first group, including *Nereis* sp. through *Caprella*
273 *laeviuscula* along the vertical axis in Fig. 3a, tended to positively co-occur with each other, and
274 negatively co-occur members of the second group. This first group is henceforth known as the
275 *Nereis* group. The second group, henceforth known as the *C. californica* group, included

276 *Caprella californica* through *Ampithoe valida,* and Harpacticoid 2 through *Lacuna* spp. in Fig.
277 3a. *Alvania compacta* and *Lacuna* spp. showed co-occurrence patterns consistent with the *C.*
278 *californica* group, but their co-occurrence values were generally weaker. The third group was
279 comprised of species that did not show strong positive or negative co-occurrence values with
280 others; *Spirorbis* sp. was the only representative in this category among the 20 most abundant
281 species.

282 Species from each of the three co-occurrence groupings were present at all sites,
283 indicating that members from these groups can be simultaneously present (Fig. 3b). However,
284 the abundances of members from different co-occurrence groups tended to negatively co-vary; at
285 sites JB, HL, CB, RA, EB, DC, and DK, members of the *Nereis* group were predicted to be most
286 abundant, whereas *C. californica* group members were predicted to be most abundant at SS, SA,
287 LH, GB and RB. The site IN was predicted to be strongly dominated by *Spirorbis* sp. The
288 emergent species groupings identified in Fig. 3a do not clearly correspond with spatial distance
289 because the cluster analysis did not group sites by spatial proximity (black symbols in Fig. 3b).

290 **Discussion**

291 By applying a new statistical framework, we have gained new insight into the possible
292 metacommunity processes structuring seagrass-associated epifaunal diversity in coastal British
293 Columbia. We report evidence that both biotic interactions and environmental conditions
294 influence seagrass meadow invertebrate composition. Spatial patterns of species abundance and
295 distribution suggest regional-scale dispersal patterns are possible, such that even distant
296 meadows may host the same species. However, these dispersal rates are clearly not so high as to
297 overwhelm the local meadow scale effects of biotic interactions and environmental conditions.
298 Below, we review these conclusions and the evidence to support them.

299 *Epifaunal communities may be well-connected and experience intermediate dispersal rates*

300 Epifaunal communities appear to be well-connected by dispersal at the present spatial
301 scale and seagrass habitat distribution in coastal British Columbia. This is true for the seven
302 species that were present at all sites, and the additional ten species that were not present at every
303 site but present in all regions (Appendix S1: Fig. S2). These widely distributed groups had
304 representation from across phyla and life history traits, suggesting that no single taxonomic
305 group or dispersal life history strategy had a consistently greater spatial distribution. Many of the
306 meadows were small (< 1 ha), and epifaunal abundances can vary substantially within meadows
307 from year to year such that local extinctions followed by rescue from nearby populations are
308 likely important (Thom et al. 1995). However, the specific pathway the organisms travelled to
309 arrive at these sites remains unknown. We used Euclidean distances in our analysis, which
310 preserved the rank order of distances between sites, but actual distances are likely greater due to
311 oceanographic circulation patterns (Kinlan and Gaines 2003, Mitarai et al. 2008).

312 It is unclear whether distant sites share the same species because they are 1) linked by
313 direct dispersal; 2) indirectly linked by dispersal via unsampled sites; or 3) were colonized by
314 populations of the same species in a historical dispersal event, but have not seen the exchange of
315 individuals since. Seagrass-associated epifauna generally disperse passively between meadows,
316 either by larval transport in currents or by "rafting" on floating pieces of seagrass or macroalgae.
317 Rafts have been observed hosting benthic and epifaunal invertebrates across phyla (echinoderms;
318 peracarids; molluscs; annelids), and contribute to connectivity between coastal ecosystems
319 across 100s of kilometres (Wichmann et al. 2012). Thus, while incidences of direct dispersal
320 between the farthest sites may be rare, they are not impossible.

321 Species that were only found at one site or region (e.g. *Jassa marmorata,* Appendix S1:

322 Fig. S2) may be dispersal limited for reasons other than prohibitively long travel distances.
323 Those with low abundances (less than five individuals per site) may not disperse in appreciable
324 numbers, and therefore cannot establish populations at all sites as readily as abundant species.
325 This is supported by our observation of a generally positive relationship between range size (the
326 number of sites a species was represented at) and abundance (Appendix 1: Fig. S6).

327 We can also conclude that dispersal rates are high enough for several species to be
328 present at most sites, but not so high as to completely overwhelm the effects of environmental
329 conditions. Thus, dispersal rates are likely intermediate on average. A subset of species was
330 present at all sites, but often at varying abundances (Fig 3b). This may be indicative of weak
331 mass effects, wherein immigration allows populations to persist even if local conditions are not
332 optimal (Schmida and Wilson 1985, Mouquet and Loreau 2003). Metacommunity theory
333 predicts that such weak mass effects are likely to be found whenever dispersal rates are not
334 limiting (Thompson et al. 2017), particularly in organisms that cannot control their own dispersal
335 (Leibold and Chase 2017) such as seagrass-associated epifauna. This may be supported by our
336 result on the contribution of the environment, particularly in ubiquitous species for which the
337 model fit was poor (e.g. *Aoroides* spp., *Spirorbis* sp., $D^2_{adj} = 0$); the poor model fit may indicate
338 that these persist at sites even if local conditions are not optimal, due to sufficient immigration.
339 An alternative explanation for the model's poor predictive power for these species is that these
340 taxa have broader environmental niches, and thus the regional environmental gradient is not
341 sufficient to drive clear presence-absence patterns.

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

343 The abiotic environmental axis had the largest influence on species distributions, with
344 phosphates, salinity, and sea surface temperature explaining the largest fractions of variation on

345 average in the joint species distribution model (Fig. 2). However, after accounting for the model
346 fit, the total variation explained by the environment was small (22%). In some species for which
347 the model had high predictive power (e.g. *Ampithoe dalli* and *Caprella natalensis*, Appendix 1:
348 Table S2*)*, niche filtering may be important, but overall there remains a large amount of variation
349 unexplained by the environment.

350 Environmental heterogeneity may influence species distributions and abundances across
351 the metacommunity in one of three ways. First, the most important abiotic environmental
352 parameters may influence food availability through primary productivity. This is possible despite
353 the fact that primary biomass explained less variation than abiotic parameters (Figure 2). While
354 the primary biomass measurements were taken at a single time point, the abiotic parameters
355 represent a four-year average (Assis et al. 2018) and may therefore better represent the
356 conditions these organisms experience over their lifetimes. Second, sea surface temperature,
357 dissolved oxygen, and salinity may affect species' distributions by via environmental tolerance
358 ranges. For example, isopod *Idotea baltica* (a relative of *I. resecata* and *I. montereyensis* in our
359 study) experiences significantly lower survival in prolonged exposure to higher temperatures and
360 lower salinities than those experienced in their home habitat (Rugiu et al. 2018). Third,
361 temperature may influence epifaunal metabolic demands (many of which are herbivores),
362 thereby influencing grazing rates and altering the available primary biomass (O'Connor 2009).
363 Decreased food resources may increase competition, thus influencing the number and identity of
364 species a seagrass patch can host. Determining whether individual species' responses to
365 competitive stress influence whole-community structure may provide key information on how
366 the environment and species interactions combine to influence community assembly.

367 Environmental conditions in temperate eelgrass meadows fluctuate seasonally, driving

368 changes in epifaunal community structure throughout the year (Wlodarska-Kowalczuk et al.
369 2014). Because these species may not show the same distributions year-round, sampling several
370 times in a year to capture these temporal changes would give a clearer picture of species'
371 responses to all metacommunity processes.

372 *Species co-variance patterns may suggest priority effects and trophic interactions*

373 Three main co-occurrence groupings emerged showing antagonistic abundances across
374 sites (Fig. 3b). Stochasticity alone is unlikely to have produced such a strong pattern, especially
375 given deterministic processes tend to be more important in systems that are not dispersal limited
376 (Lee and Bruno 2009, Ron et al. 2018). While there are multiple possible explanations for non-
377 random species co-occurrence patterns (Connor and Simberloff 1979), space and environmental
378 heterogeneity were ruled out as co-occurrence values represented pairwise covariation
379 unexplained by environment and space (Ovaskainen et al. 2017). This, along with the high
380 unexplained fraction in the variation partitioning leads us to suggest the importance of biotic
381 processes. Several types of interactions may be occurring in the metacommunity; the majority of
382 species are herbivores or detritivores, and thus may compete for primary biomass. Laboratory
383 experiments have shown that behaviours such as grazing rate and habitat selection in amphipods
384 are altered in the presence of morphologically and functionally similar intrerspecific competitors
385 (Beermann et al. 2018). Other species in our study are filter feeders, and thus may not be
386 competing for food resources but possibly predator-free space or substrate on the eelgrass blades.

387 The community structure of a given site with certain local environmental conditions and
388 access to a certain regional pool may have multiple possible compositional states depending on
389 the arrival order of species; this phenomenon is known as priority effects (Fukami et al. 2016, Ke
390 and Letten 2018) and has been documented in marine fouling communities (Vieira et al. 2018).

The antagonistic co-occurrence groupings in our analysis may suggest priority effects (Fig. 3a and 3b). We hypothesize that following the establishment of a new meadow or a disturbance event, the first few species to colonize a seagrass meadow determine the success of subsequent colonizers, regardless of local environmental conditions. As an illustrative example, *Lacuna* spp. might have been among the first to colonize sites JB, HL, GB, and CB (Fig. 3b), thereby gaining a numerical advantage over species in the alternative cohort. Distibution patterns in the sessile calcifying polychaete *Spirorbis* sp. may also indicate priority effects. While it did not strongly negatively or positively co-occur with other species, it dominated the HL site; 16 300 individuals were found, several orders of magnitude higher than its abundance elsewhere (Data S1). Site HL also had the lowest species richness of all 13 sites (9 species). Eelgrass shoots there were completely covered by these *Spirorbis* sp., possibly preventing grazers in the community from accessing epiphytic algae or the seagrass itself. The age, colonization history, and disturbance regime of the meadows likely matter if priority effects are occurring.

The negatively correlated abundance patterns could also reflect the presence of a predator such as *Nereis* at some sites and not others (Fig. 3b). Field experiments have shown that changes in predation pressure by fish, shorebirds, and predatory invertebrates can shift seagrass-associated epifaunal assemblages in a matter of weeks (Amundrud et al. 2015; Huang et al. 2015). Differential predation pressure could also lead to the pattern of diverging community composition that we have witnessed (Guzman et al. 2018). A more comprehensive sampling effort of fish and bird predators would be required to determine the degree to which trophic interactions structure these communities.

Positive co-occurrence patterns between certain species may additionally be a result of positive biotic interactions. An example is seen in *Orchomenella recondita* - a gammarid

414 amphipod that lives in the gastrovascular cavity of the anemone *Anthopleura elegans* (Carlton
415 2007). This species was only recorded at the SS site, and specifically only found in quadrats
416 where *A. elegans* was collected with the seagrass shoots.

417 **Conclusion**

418 Using a new statistical approach to characterizing dispersal, environmental niche
419 filtering, and biotic interactions, we have developed refined hypotheses about the
420 metacommunity processes driving temperate seagrass epifaunal diversity on a regional scale. Our
421 study also gives a baseline spatial scale (1000 km) for comparing the contribution of these
422 processes across spatial scales and other ecological systems (Viana and Chase 2018).
423 Manipulative experiments and monitoring are now needed to test our refined hypotheses. One
424 could test the hypothesis that species sorting is relatively weak by measuring fitness differences
425 along the most important environmental axes. Given our hypothesis on the importance of priority
426 effects, one might experimentally manipulate the colonization order of key species in our study
427 to determine if resulting compositional states differ (Fukami et al. 2016). Testing our hypothesis
428 that populations within the metacommunity are not dispersal limited is more challenging, but
429 could be done using particle tracking models similar to those used to estimate larval dispersal
430 between coral reefs across thousands of kilometres (Treml et al. 2008). A population genetic
431 analysis could also ground-truth the extent of dispersal between allopatric populations. All
432 together, we suggest that a modern metacommunity perspective combined with new advances in
433 statistical methods provide a powerful and relevant framework for studying natural ecosystems.

434 Our findings are particularly relevant to conservation initiatives that aim to maintain
435 diversity in seagrass meadows; identifying the importance of local and regional processes allows
436 the conservation practitioner to better predict how habitat fragmentation and climate change

might influence diversity. Seagrass epifaunal communities may be quite resilient, given our findings that environmental conditions minimally contribute to species distributions, and several species across 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 IN, DK, and EB. LL provided samples from RA and HL. KAS collected and processed samples, 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 analyses. Matt Whalen pulled the Bio-ORACLE dataset, and John Cristiani made the map in Fig. 1a. Special thanks goes to Michelle Paleczny and Gulf Islands National Park Reserve for boat and laboratory space. We acknowledge staff, students, and volunteers from Pacific Rim National Park Reserve for their assistance collecting eelgrass in Clayoquot Sound. We thank Rob Underhill, Abbie Sherwood, Emily Adamczyk, Minako Ito, Ariane Comeau, Giordano Bua for assistance with sample collection, processing, and ID, and Coreen Forbes, Matt Whalen, and Jeff Groh for feedback on the manuscript.

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, 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, E. L., 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. Patterns 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. The American Naturalist 161:889–-904.

Guillaume Blanchet, F., with contributions from G. G. Tikhonov and A. 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.

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 2:1691–1695.1–7.

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.

Logue, J., et al. 2011. Empirical approaches to metacommunities: a review and comparison with theory. Trends in Ecology and Evolution 26:482–491.

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. The 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. Tikhono, 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.

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.

Ward, J. H. 1963. Hierarchical grouping to optimize an objective ounction. 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, E. 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.

592 **Figure legends**

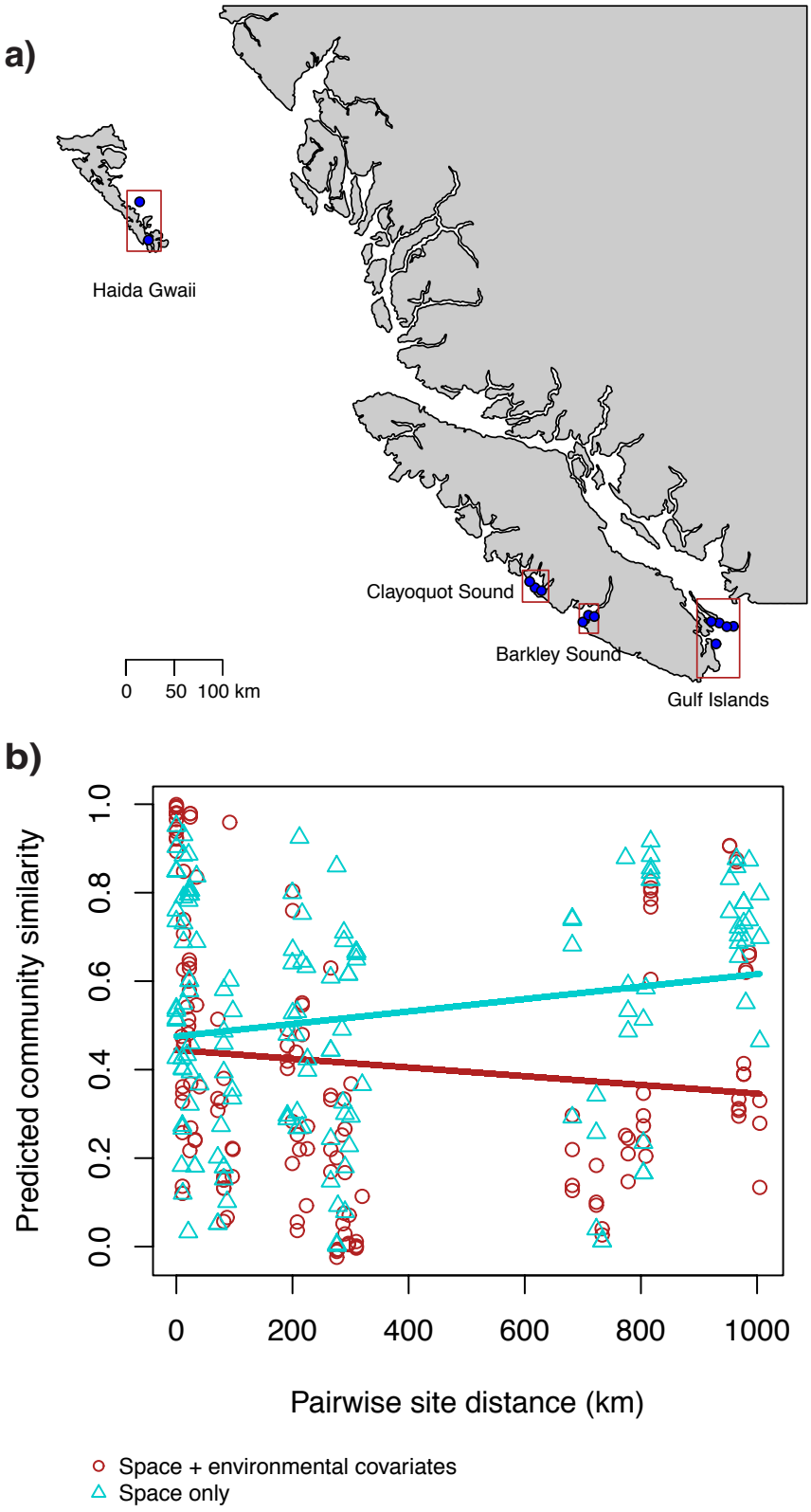
593 **Figure 1. a)** Map of coastal British Columbia showing eelgrass field sites in Haida Gwaii (HL,
594 RA), Clayoquot Sound (DK, EB, IN), Barkley Sound (SA, DC, RB), and the Southern Gulf
595 Islands (GB, JB, CB, LH, SS). **b**) HMSC model predictions of community similarity relative to
596 the spatial distance (Euclidean) between communities. The red circles show similarity
597 predictions based on the full model. The blue triangles show predictions wherein the effect of all
598 environmental variables has been removed (standardized to their mean values).

599 **Figure 2.** Variation partitioning of fixed and random effects across all HMSC species
600 distribution models (numbers correspond to species names in Appendix 1: Table S2). Black cells
601 represent the contribution of spatial distance. Purple cells represent random effects. Blue cells
602 represent the contribution of water quality parameters. Yellow cells represent food availability
603 (eelgrass, algae and detritus biomass). Green cells represent habitat structure, which is the sum of
604 variation explained by eelgrass LAI and shoot density. Bars are organized according to broad
605 taxonomic groupings as shown in Appendix 1: Table S2. The percentages next to the legend
606 labels indicate mean variation explained by that variable across all species. Adjusted deviance
607 explained (D^2_{adj}) varied (Appendix 1: Table S2). The mean D^2_{adj} across all species was 0.29.

608 **Figure 3. a.)** A correlation plot showing modelled site-level pairwise co-occurrences in the 20
609 most abundant species. Purple cells represent positively co-occurring species pairs, and turquoise
610 cells represent negatively co-occurring species pairs. Species names are ordered according to the
611 output of hierarchical clustering with Ward's criterion on pairwise co-occurrence values. A
612 version of this figure with all 58 species is shown in Appendix S1: Fig. S4. **b)** Heatmap and
613 cluster dendrogram depicting modelled species abundance and compositional similarity across
614 sites. The heat map shows the twenty species with the highest predicted mean proportional

615 abundances ordered from highest to lowest. Cell colours correspond to the HMSC co-occurrence
616 groups shown in Fig. 3a: turquoise or *Nereis* group, purple or *C. californica* group, and grey
617 *Spirorbis* group. Cell shade strength indicates proportional abundance at a given site (darker
618 means higher abundance). The symbols to the left of the site abbreviations indicate region
619 membership; circles show Haida Gwaii sites, squares show Barkley Sound sites, diamonds show
620 Clayoquot Sound sites, and triangles show Southern Gulf Islands sites. A version of this figure
621 with all 58 species is shown in Appendix S1: Fig. S5.

622 **Figure 1**



623

Figure 2.

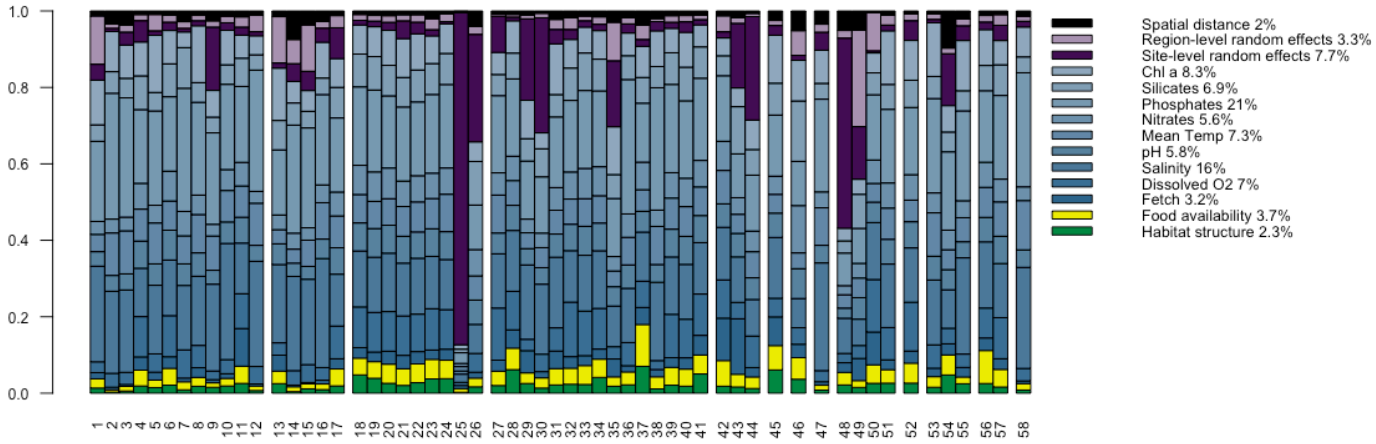
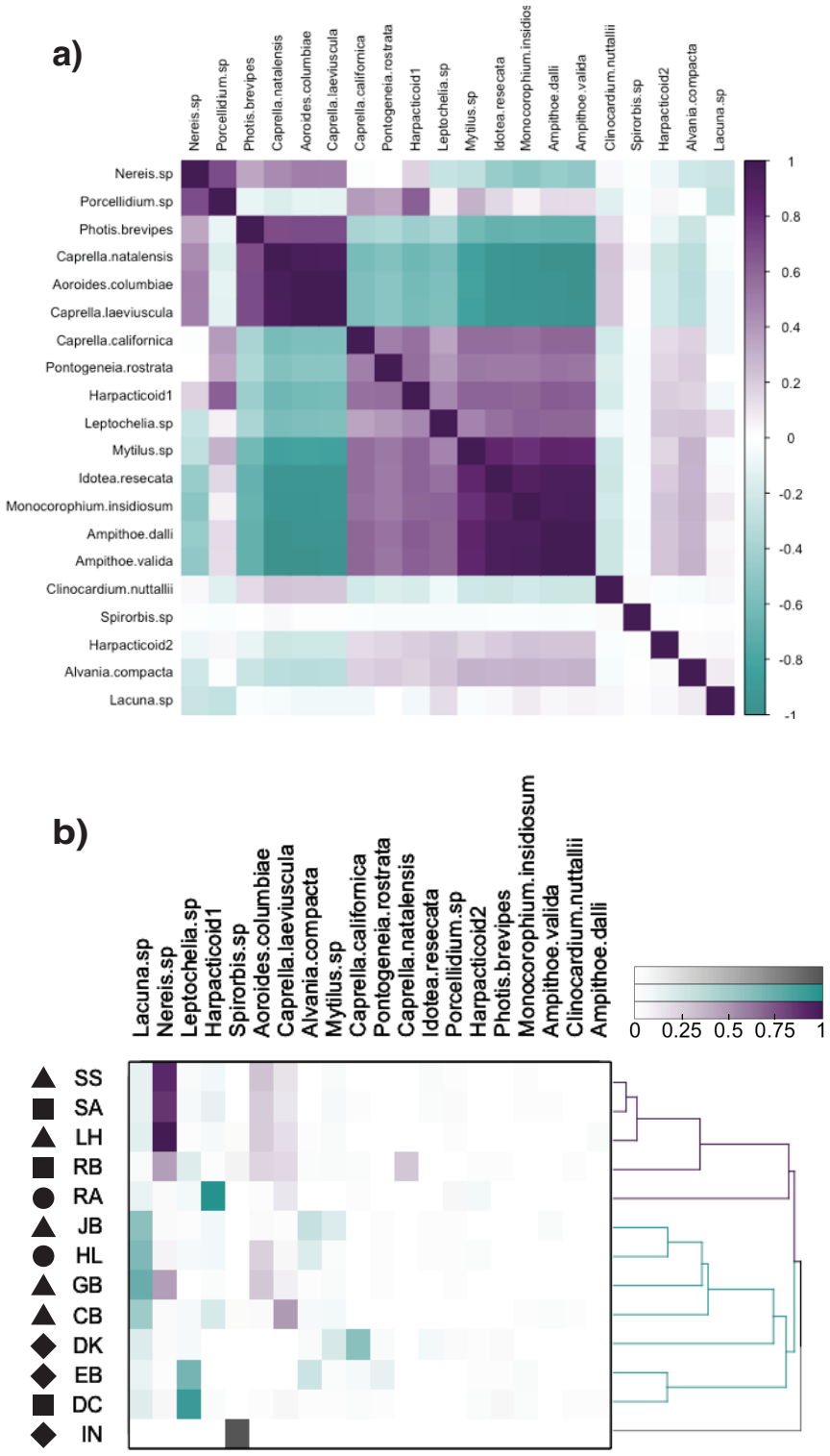

1.0
0.8
0.6
0.4
0.2
0.0
Spatial distance 2%
Region-level random effects 3.3%
Site-level random effects 7.7%
Chl a 8.3%
Silicates 6.9%
Phosphates 21%
Nitrates 5.6%
Mean Temp 7.3%
pH 5.8%
Salinity 16%
Dissolved O2 7%
Fetch 3.2%
Food availability 3.7%
Habitat structure 2.3%

Figure 3


a)
Nereis.sp
Porcellidium.sp
Photis.brevipes
Caprella.natalensis
Aoroides.columbiae
Caprella.laeviuscula
Caprella.californica
Pontogeneia.rostrata
Harpacticoid1
Leptochelia.sp
Mytilus.sp
Idotea.resecata
Monocorophium.insidiosum
Ampithoe.dalli
Ampithoe.valida
Clinocardium.nuttallii
Spirorbis.sp
Harpacticoid2
Alvania.compacta
Lacuna.sp
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
b)
Lacuna.sp
Nereis.sp
Leptochelia.sp
Harpacticoid1
Spirorbis.sp
Aoroides.columbiae
Caprella.laeviuscula
Alvania.compacta
Mytilus.sp
Caprella.californica
Pontogeneia.rostrata
Caprella.natalensis
Idotea.resecata
Porcellidium.sp
Harpacticoid2
Photis.brevipes
Monocorophium.insidiosum
Ampithoe.valida
Clinocardium.nuttallii
Ampithoe.dalli
0
0.25
0.5
0.75
1
SS
SA
LH
RB
RA
JB
HL
GB
CB
DK
EB
DC
IN

628 **Appendix S1:** Supporting Information

629

630 **Table S1.** 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. 1b
	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. 3a, 3b
(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. 2
(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. 3a, 3b

631

632

Table S2. Species names, taxonomic groupings, and proportion of deviance (predictive power). Numbers correspond to those in Figure 2. The top twenty species with highest predicted 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

640 **Table S3.** Summary of traditional biodiversity metrics for each site

641

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

642

643

644

645

646

647

648

649

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 the 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*).

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;

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 to calculate fractions *a, b, c,* and *d.*

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.

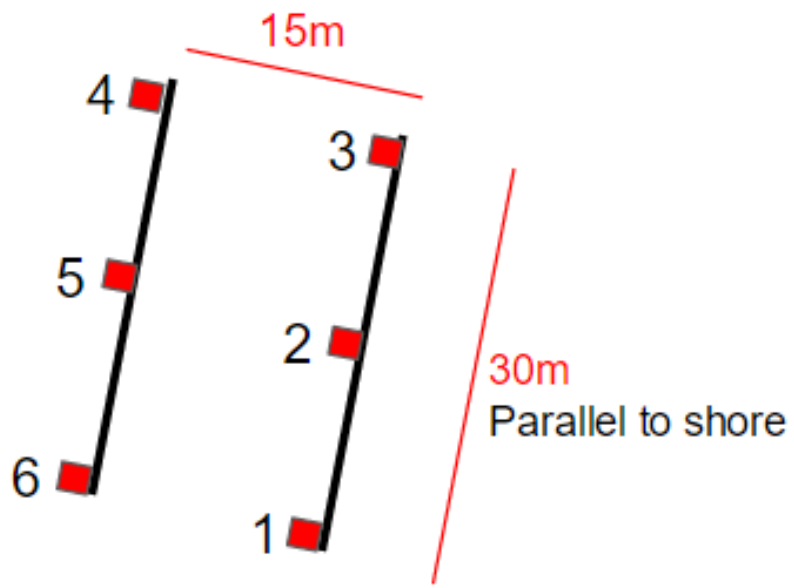


Figure S1. Schematic diagram field sampling regime. Black lines represent transect lines, and red boxes represent 0.25 x 0.25 m quadrats.

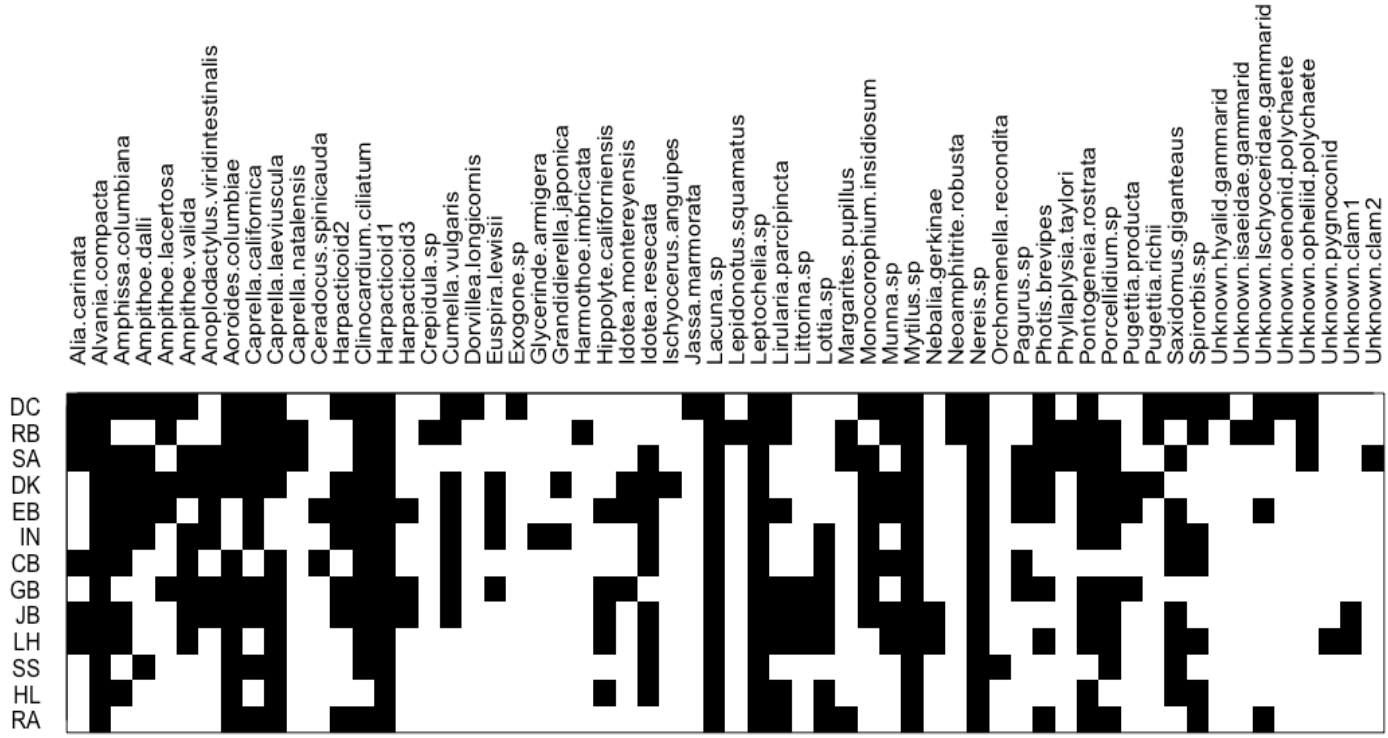


Figure S2. Site-by-species presence-absence matrix of all 58 invertebrate species in the study listed in alphabetical order. Black cells indicate species presences.

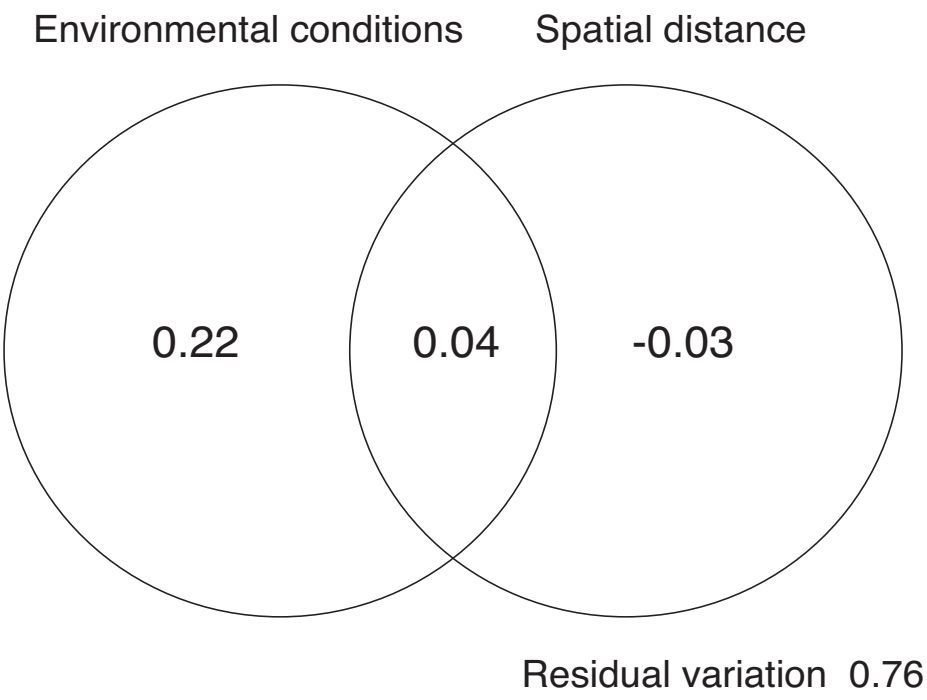


Figure S3. Venn diagram summarizing the fractions of variation explained by environmental covariates only (nine water quality and five biotic variables), spatial distance only, and the shared fraction explained by environment and space.

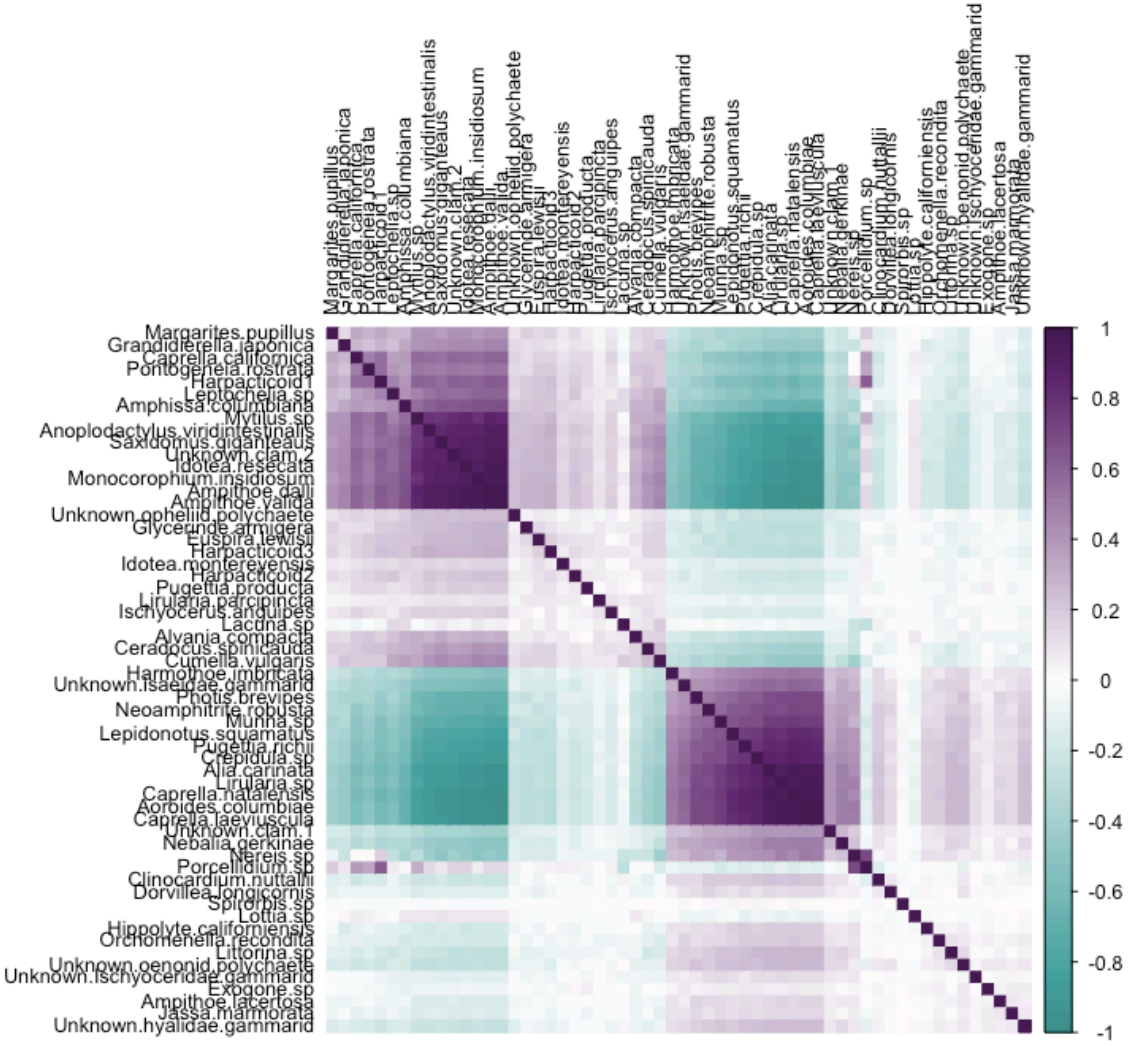


Figure S4. A correlation plot showing modelled site-level co-occurrence of species pairs across all species. Purple cells represent positively co-occurring species pairs, and turquoise cells represent negatively co-occurring species pairs. Species names along both axes are ordered according to the output 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.

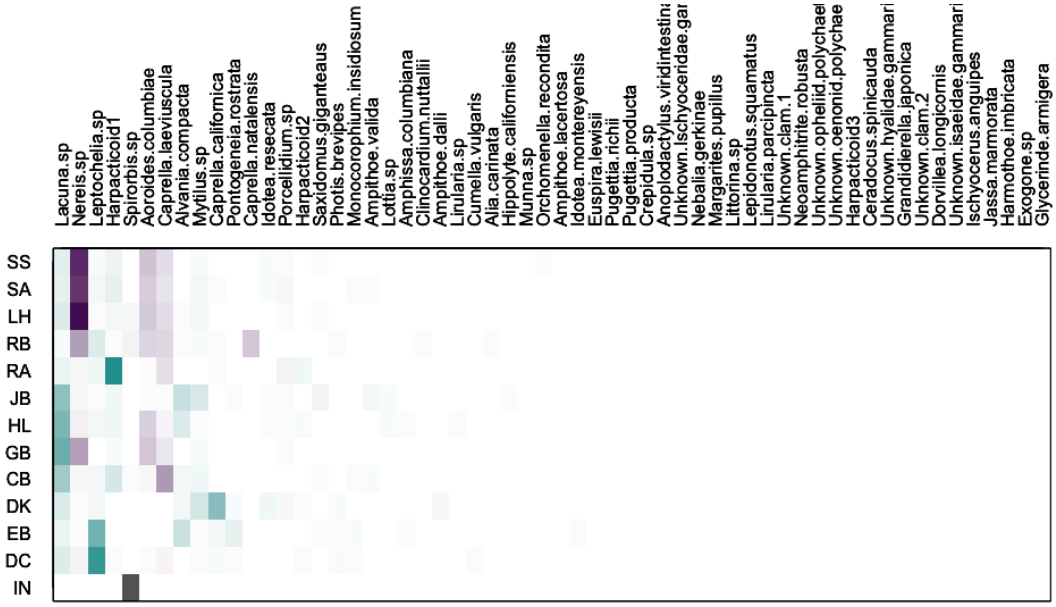


Figure S5. Heat map depicting abundance patterns in all 58 species in the study. Species names are ordered from highest to lowest predicted mean proportional abundance. Cell colours correspond to the co-occurrence groups in Fig. 3a, with purple cells representing members of the *Nereis* assemblage, and turquoise cells indicating 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. 3b in the main text.

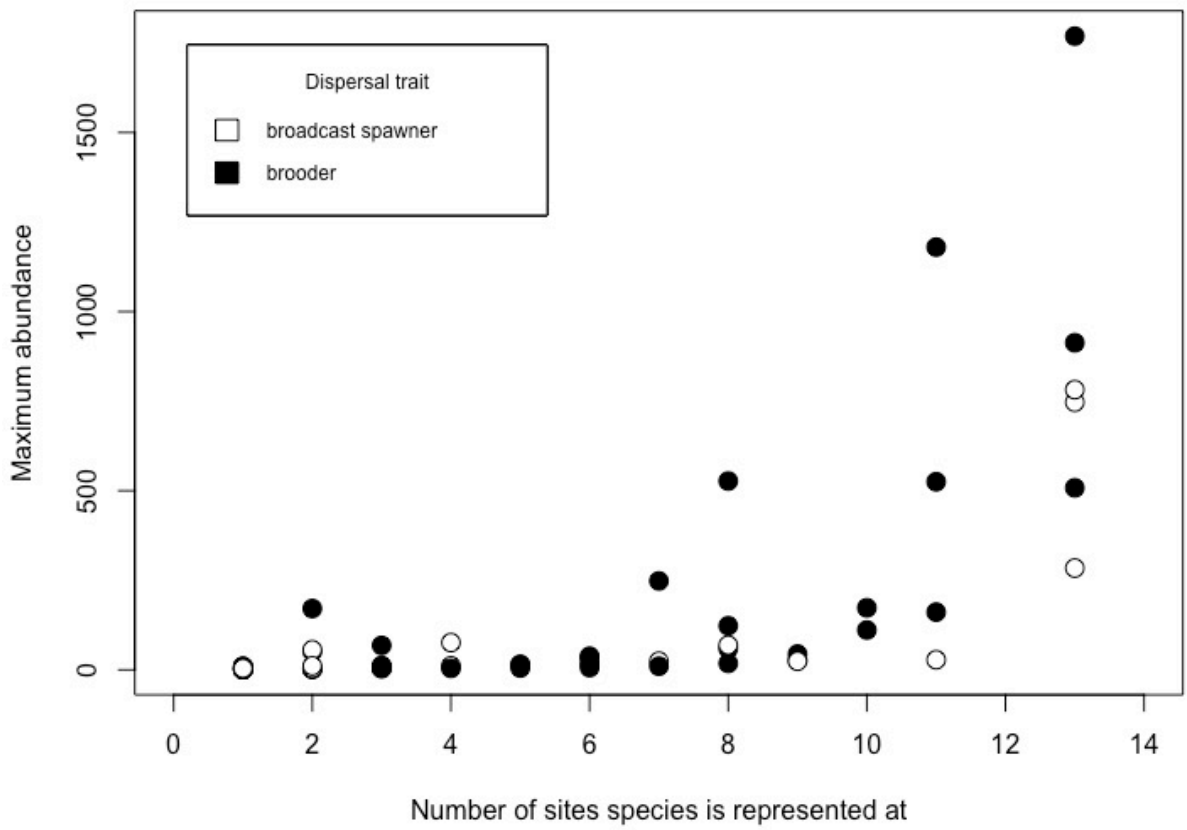


Figure S6. Plot showing relationship between maximum site-level abundance and site occupancy for all species. Black points represent broadcast spawners and white points represent brooders. Spirorbid polychaetes were excluded from this figure due to their unusually high maximum abundance (approximately 16 300 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 ($t = 0.052$, $df = 42.5$, $p = 0.96$).

Literature Cited

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology 73:1045–1055.

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

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.

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