

1
2 Received Date : 19-Nov-2015

3 Revised Date : 20-May-2016

4 Accepted Date : 28-Jun-2016

5 Article type : Standard Paper

6 Editor : Angela Strecker

7
8
9 **A fresh approach reveals how dispersal shapes metacommunity**
10 **structure in a human-altered landscape**

11
12
13
14 **Barbara J. Downes, Jill Lancaster, Alena Glaister and William D. Bovill**

15 *School of Geography, The University of Melbourne, Victoria, Australia*

16
17 Contact author:

18
19 Prof. Barbara Downes, School of Geography, 221 Bouverie St, The University of Melbourne, Vic.
20 3010, Australia

21
22 e-mail: barbarad@unimelb.edu.au

23
24 Tel: 61-3-83449152

25
26
**This is the author manuscript accepted for publication and has undergone full peer review
but has not been through the copyediting, typesetting, pagination and proofreading process,
which may lead to differences between this version and the [Version of Record](#). Please cite this
article as [doi: 10.1111/1365-2664.12759](https://doi.org/10.1111/1365-2664.12759)**

This article is protected by copyright. All rights reserved

27 *Running title: Dispersal and metacommunity structure*

28

29

30

31

32

33 **Summary**

34 1. To understand species losses from disturbed landscapes, it is important to distinguish the effects of
35 degraded environmental conditions from those caused by barriers to dispersal between habitat
36 patches. To assess the relative importance of these effects, we developed a new approach using
37 permutation and association tests applied to rank abundance data, using the invertebrate fauna of two
38 rivers in two seasons.

39 2. Our study streams were Hughes Creek and Seven Creeks, in south-eastern Australia, which have
40 both been degraded by agriculture in downstream sections. We collected benthic invertebrates and
41 also dispersing individuals (drift, terrestrial adults) during two seasons in 2007–2008. Study sites
42 spanned strong environmental gradients as well as the main dispersal route (up- and down-channel).
43 Environmental data were analysed to set up permutation tests on rank abundances. Survey and
44 disperser data were contrasted using contingency table analyses.

45 3. The results suggest dispersal plays a strong role in community structure. Environmental effects
46 were evident and strongest upstream, but evidence of environmental effects was weak over much of
47 the gradient. Many species had different distributions in different data sets or dispersers that were
48 abundant at locations distant from centres of benthic distribution.

49 4. Our results differ from many studies, but few have been able to evaluate dispersal effects directly.
50 Our method provides a practical approach for evaluating the role dispersal plays in driving species
51 abundance patterns across landscapes, thus bridging a gap between theory and practice.

52 5. *Synthesis and applications.* Managers typically use indices of ecosystem health that assume
53 environmental conditions largely determine species diversity and abundance. Dispersal between
54 habitat patches is known to be important, but there are no reliable methods to assess the role dispersal
55 may play. We provide an approach that allows both dispersal and environmental effects on species
56 distributions to be evaluated from survey data. This may open the way for dispersal information to be
57 incorporated into management actions. Additionally, the approach should allow improved siting of
58 restoration projects that depend greatly on successful dispersal of individuals for successful outcomes.

59

60 **Key-words:** combinatorics, dispersal, disturbance, environmental gradients, land clearance,
61 restoration, south-eastern Australia, species diversity, stream invertebrates, variation partitioning

62 **Introduction**

63 As human impacts on the environment continue to grow, rates of species extinction are expected to
64 climb, but we understand little about the direct causes of species losses when human impacts increase
65 gradually (Scheffers *et al.* 2012). Losses could be due to a gradual degradation of environmental
66 conditions (e.g. increased edge effects as habitat patches shrink in size, etc.) or because dispersal of
67 individuals is slowly cut off as patches become increasingly isolated in a matrix of disturbed
68 environment (Fahrig 2007). It is imperative to distinguish between these two causes but managers
69 typically have little information about dispersal of species across landscapes. Most indices of
70 ecosystem “health” are derived from variables measuring environmental condition, which are used as
71 a surrogate for biological diversity (e.g. Kwok, Eldridge & Oliver 2011; Rodgers *et al.* 2012; Dobbie
72 & Clifford 2015). The strength of the relations between environmental condition and species
73 abundances can be tested using basic survey data (sites X species abundances X environmental
74 variables: e.g. Kwok, Eldridge & Oliver 2011) but there are few ways for highlighting the role that
75 dispersal may be playing from such basic survey data.

76 One under-used framework for addressing these two prospective causal explanations for species
77 losses is that of metacommunity dynamics (Montoya, Rogers & Memmott 2012). The
78 metacommunity approach recognises that communities occur in patches of suitable habitat distributed
79 across landscapes. Environmental conditions vary between places and have a strong impact on local
80 species composition, but dispersal of individuals between communities is a key aspect. Species that
81 are good dispersers can theoretically remain extant regionally even if they are periodically excluded
82 from some communities or are disadvantaged in some localities but sustained by high rates of
83 immigration from patches elsewhere. Thus, a metacommunity framework applied in the context of
84 landscapes disturbed by human activities can help with identifying management actions (Bengtsson
85 2010).

86 Most metacommunity research has also used basic survey data (i.e. sites X species abundances X
87 environmental variables) to separate environmental from dispersal effects on natural communities. A
88 common method is variation partitioning, in which the collective effect of environmental variables on
89 species abundances is separated from “pure” spatial variation – i.e. variation that is related to
90 geographic location and independent of environmental conditions. Pure spatial variation is interpreted
91 to reflect the effects of dispersal constraints on abundances (Cottenie 2005), and hence the approach
92 contrasts the relative roles of dispersal and environmental gradients in structuring communities.
93 However, variation partitioning is effective mostly when environmental and pure spatial variation are

94 largely uncorrelated (Legendre & Legendre 2012). In many ecosystems (e.g. rivers, coastal marine
95 habitats, corridor vegetation across landscapes) dispersal routes can be strongly directional. If
96 environmental conditions change along such dispersal routes, then environmental and pure spatial
97 variation will be correlated, in which case using variation partitioning to separate environmental and
98 dispersal effects is problematic (see also Gilbert & Bennett 2010). Another approach uses
99 presence/absence data to test for patterns of distribution that are consistent with different community
100 models (e.g. Leibold & Mikkelsen 2002; Presley, Higgins & Willig 2010) but these do not
101 specifically test for the effects of dispersal.

102 Here, we present a different analytical approach, and we illustrate it using invertebrate species
103 inhabiting rivers that are partly degraded by human impacts, which may have caused losses of species
104 in downstream areas. Many riverine invertebrates use stream currents to disperse downstream (“the
105 drift”) and, while some species are fully aquatic, many are insects that have a terrestrial, winged adult,
106 meaning that dispersal can also take place during this life-cycle stage. Although they may move
107 between catchments, adult insects commonly use river corridors for dispersal (Lancaster & Downes
108 2013). Thus environmental gradients along channels and the effects of any dispersal constraints are
109 likely to be correlated.

110 We collected benthic invertebrates as well as dispersing individuals (drifters and winged terrestrial
111 adults) from multiple sites in two rivers and at two times and evaluated a set of predictions (Table 1,
112 Fig. 1), as explained further below. Unlike most studies relying on survey data, our alternative
113 approach suggests that dispersal played a prominent role in determining the distributions of species,
114 even though environmental gradients were strong.

115 **Materials and methods**

116 APPROACH FOR TESTING HYPOTHESES

117 The first logical element of our approach uses survey data and focuses upon those locations where
118 a species is relatively abundant. Such locations signal the environment is suitable for that species,
119 whereas locations where a species is relatively uncommon provide ambiguous information
120 (uncommonness could be caused by either environmental unsuitability or dispersal constraints; note
121 that we do not consider naturally rare species that are uncommon everywhere). If species distributions
122 are largely determined by environmental conditions, then we expect the highest abundances to occur
123 in locations having similar environments. Using this logic, if we order sites by their relative
124 environmental similarity, we can use this order to arrange the rank abundances of species. We expect
125 high ranks to cluster together when the environment determines distributions. The alternative
126 hypothesis is that species have relatively wide tolerances for environmental conditions and so

127 distributions are largely determined by successful dispersal (i.e. dispersal where species are able to
 128 establish). Because dispersal can vary greatly in space and time, high abundances will sometimes
 129 occur in well-separated locations having dissimilar environmental conditions. Thus when dispersal
 130 plays a strong role, sites ordered by relative environmental similarity should sometimes have highly
 131 ranked species abundances that are well-separated (“split”) or at least not clustered.

132 Formulae from combinatorics can provide tests to detect whether such clusters (or splits) of high
 133 ranks occur more or less often than expected by chance. Here, we consider only linear arrays of
 134 abundances because this information is relevant to the data we have collected, but the logic explained
 135 below can be extended to consider grid-based survey data (i.e. data collected across landscapes).

136 Combinatorics provide the formulae for calculating the probability that certain numbers occur
 137 together (a “cluster”) by chance (Allenby & Slomson 2011). If we have a linear array of n items and c
 138 of these items must occur adjacent to each other but can be in any order, the number of permutations
 139 fulfilling this requirement is:

$$(n - (c - 1))! c!$$

(Equation 1)

140 The total number of possible permutations is $n!$ Thus, the probability of gaining a cluster of size c by
 141 chance is:

$$P_c = \frac{(n - (c - 1))! c!}{n!}$$

(Equation 2)

143 (see Appendix S1 in Supporting Information for full explanation).

144 Because species abundances at n sites can be converted to ranks, Equation 2 permits us to calculate
 145 the probability that the highest ranks occur in a cluster of size c by chance. The value of c is
 146 determined by analysing environmental data, using tests designed *a priori* to identify the cluster size
 147 that captures significant differences in environmental conditions (explained further below). If there
 148 are N species in a data set, then NP_c is the expected number of species that are clustered by chance.
 149 The difference between expected and observed numbers of clusters can be tested using chi-square
 150 (Sokal & Rohlf 1995). If we have M multiple, independent data sets, we can also test whether
 151 individual species are consistently clustered, as expected if they have strong responses to
 152 environmental conditions. The probability that clustering occurs in multiple independent data sets by
 153 chance is P_c^M . The latter allows us to calculate the expected number of species that will cluster
 154 repeatedly in independent data sets by chance, which we can compare to observed numbers using chi-
 155 square.

156 Equation 2 can also be used to calculate the probability that a cluster of size c contains one or more
 157 numbers but not others (Appendix S1), which means we can also calculate the probability that the
 158 highest ranks are not within a cluster, i.e. are “split” (dispersal is frequent, environmental effects are

160 weak). The ability to calculate an expected probability of splits allows further tests. We expect the
161 frequency of splits to be significantly lower than expected by chance if environmental effects
162 predominate and significantly greater if dispersal is common.

163 This approach differs from many other methods (e.g. direct gradient analyses) in which relations
164 between species abundances and environmental variables are tested first. The effects of dispersal are
165 discerned secondarily from outcomes that have already largely been shaped by environment-
166 abundance relations; this therefore seems likely to privilege the environmental explanations. The
167 benefit of our approach is that environmental data are used to set up separate predictions for each
168 hypothesis (i.e. predominantly environment effects, predominantly dispersal effects). Each set of
169 predictions is then evaluated independently and with potentially equal statistical power. Although
170 variation partitioning evaluates environmental and dispersal effects simultaneously, results are
171 difficult to interpret if environmental and spatial variation overlap greatly, i.e. are autocorrelated
172 (Legendre & Legendre 2012). For illustration, we analysed our data with variation partitioning (see
173 Appendix S2) to demonstrate that, as expected, spatial and environmental variation were highly
174 overlapped. Moreover, spatially-structured environmental variation was higher than either spatial or
175 environmental variation, and this precludes a confident assessment of environmental and dispersal
176 effects on community data (Legendre & Legendre 2012). Note that there are more complex methods
177 of analysis than we used, such as Moran's eigenvector maps (MEM) (or asymmetric eigenvector maps
178 – although the latter can produce similar results to MEM: Legendre & Legendre 2012). However,
179 MEM have been found also to produce unreliable estimates of variation (Gilbert & Bennett 2010).

180 A second logical element of our approach makes use of individuals sampled at each site in the act
181 of dispersing. The ability to sample dispersers may be infeasible for some species but is possible for
182 many, especially those that disperse advectively (i.e. by wind or water currents), where nets or traps
183 can be specifically deployed to capture them. Most studies demonstrate that many dispersers do not
184 travel far from centres of high abundance (Lowe & McPeck 2014). Indeed, if environmental
185 conditions determine species abundances, we expect dispersers to be most abundant at the same sites
186 as settled individuals, because dispersal away from suitable environments would be maladaptive.
187 Thus the classifications above (clustered, split, etc.) can be applied to the rank abundances of
188 dispersing individuals. Classifications of dispersers can be compared to those of surveyed individuals
189 of the same species using contingency table analysis (G or chi-square: Sokal & Rohlf 1995). If
190 environmental conditions are key, we expect dispersers and surveyed species to be similarly and
191 consistently clustered, i.e. classifications should be strongly associated. Alternatively, if dispersal is
192 frequent, then we expect mismatches, and dispersing species may show different classifications in
193 different data sets.

194 In a third logical element, outcomes for tests are compiled for each species. Those taxa for which
195 distributions are largely determined by environmental conditions are expected to be consistently

196 clustered and to have disperser distributions that consistently match those of surveyed species. Those
197 taxa where dispersal strongly affects distributions should not be clustered, have inconsistent
198 distributions in different data sets and mismatches between surveyed and disperser distributions.

199 We used the above reasoning to produce a set of specific hypotheses (Table 1, Fig. 1) to test for
200 samples collected from our two rivers, as described below. The dispersal of stream organisms in a
201 downstream direction using stream currents (the drift) and the fact that both larval and adult stages
202 can disperse allowed us to erect some predictions that were specific for flowing water environments.
203 The practical steps followed during numerical analyses are outlined in Table 2 and discussed below.

204 STUDY SITES AND ENVIRONMENTAL MEASURES

205 We used two rivers, Hughes Creek and Seven Creeks, located within an agricultural region of
206 south-eastern Australia (Downes *et al.* 2011). Both creeks arise on Strathbogie Range but flow in
207 different directions (their closest headwaters are ~ 7-8 km apart, a distance that is traversable by adult
208 insects: unpublished data, Lancaster & Downes 2013) and both have sandy beds. Sand in the creek
209 beds is sourced from erosion caused by natural events (e.g. bushfires) and by widespread clearance of
210 catchment vegetation. Deposition of sand causes an unstable bed and a loss of habitat structure as
211 wood, bark and leaves (the primary source of hard substrate and food in the systems: Lancaster &
212 Downes 2014b) are smothered. Riparian vegetation (mostly *Eucalyptus camaldulensis*) is continuous
213 along both banks but is sparse in downstream locations due to stock access, reducing contributions of
214 detritus to the stream channel and further exacerbating declines of in-stream resources of food and
215 living space (Reid *et al.* 2008).

216 The environmental gradient along each creek was characterised using eight, randomly chosen 40 m
217 long sites (Appendix S3) from sections of each creek that were topographically accessible. These sites
218 span strong environmental gradients (see Results). We deliberately sampled streams with two sites
219 located in upstream areas, two located in downstream areas and four sites in-between, with an
220 expectation that the latter would reflect conditions that were intermediate between these starkly
221 different environments. Densities of detritus in the channel (leaves, bark, twigs) were measured at
222 each site (Appendix S3) on each stream during summer, and in spring in Hughes Creek. Additionally,
223 in summer, the total surface areas of wood and bark at each site were measured. Google Earth
224 photographs were used to estimate riparian width and tree cover (Appendix S3). We took spot
225 measures of water quality (pH, conductivity, dissolved oxygen, turbidity using a Multi 340i WTW
226 82362; MultiLine, Weilheim, Germany), and measured channel widths and water depths on summer
227 dates. Regular measures of water and air temperatures were collected from 2008 to 2009 using data
228 loggers (TruTrack Loggers, Model WT-HR). Discharge data were obtained from government sources;
229 discharges were similar between the two study years (Appendix S3). Collectively these variables span

230 the range of major environmental variation that affects the distribution and abundance of stream
231 invertebrates (Hynes 1970).

232 INVERTEBRATE SAMPLING AND IDENTIFICATION

233 Benthic invertebrates at all sites on both creeks were sampled in summer in 2007, and Hughes
234 Creek sites were sampled again in spring of 2007. The benthic data comprise mostly larval stages and
235 form the survey data on species distributions. The sampling protocols have been described elsewhere
236 (Downes *et al.* 2011) (Appendix S3). Prospective dispersers (drifting invertebrates, terrestrial adult
237 insects) were sampled at pairs of locations on each creek near the ends of environmental gradients
238 only (species that occurred solely in the middle of the gradient proved uncommon - see Results).
239 Drifters were collected using drift nets and were composited into a single sample for each site and
240 time. At the same times and dates, adults were captured using light traps and sticky traps (Appendix
241 S3) and also compiled into a single, composited sample for each site and time. Drift numbers were
242 corrected by the proportion of the cross-sectional area of sites that was intercepted by nets, which
243 estimates total numbers of drifters moving hourly through sites (as per Downes & Lancaster 2010).
244 Adult samples were counted in entirety, whereas drift and benthic samples were sub-sampled
245 (Appendix S3). Drift and adults were sampled a year after benthic sampling. Given the environmental
246 gradients associated with longitudinal channel profiles and discharge hydrographs do not vary greatly
247 between years, sampling dispersers a year after benthic samples is unlikely to be problematic.

248 DATA PREPARATION AND STATISTICAL ANALYSES

249 *Environmental variation along channels (Table 2, Step 2)*

250 Environmental variables collected at each site along Hughes Ck and Seven Cks were log-
251 transformed where necessary, normalised and converted to a dissimilarity matrix using Euclidean
252 distances (using the statistical program PRIMER: Anderson, Gorley & Clarke 2008). We then used
253 PRIMER to carry out non-metric, multi-dimensional scaling (NMDS) to display differences among
254 sites with respect to position along each channel. Stress scores were used to assess the number of axes
255 needed to display the data adequately (Anderson, Gorley & Clarke 2008). As expected, sites that were
256 adjacent within creeks were environmentally similar and hence adjacent also within the NMDS plot
257 (see Results).

258 We chose a cluster size by grouping environmentally similar sites (i.e. that were adjacent within
259 the NMDS plot) and testing whether groups differed significantly. Thus, we first put adjacent sites
260 into pairs (i.e. site 1 and 2 formed the first pair, site 3 and 4 the second pair, etc.), and analysed the
261 differences between creeks and the four pairs (i.e. groups) with a factorial PERMANOVA (Anderson,
262 Gorley & Clarke 2008) with both factors fixed. *A priori* contrasts compared adjacent pairs (i.e. 1 vs 2,
263 2 vs 3, 3 vs 4). The analysis and contrasts were repeated by creating three pairs of sites beginning

264 from site 2 and ending at site 7, for a total of five contrasts using pairs of sites. We then repeated this
265 procedure using two trios of sites (i.e. groups of three), beginning with site 1 (i.e. sites 1, 2 and 3 vs 4,
266 5 and 6), then from site 2, and then from site 3, producing three contrasts. We did not use cluster sizes
267 of four or higher given the total number of sites was only eight and given our sampling design. The
268 proportion of significant contrasts was used to decide choice of cluster size (i.e. either two or three).
269 Cluster size is not a “definitive” measure of environmental variation to which all species are expected
270 to respond. Rather, it is a group size over which environments change significantly and that should
271 therefore be associated with changes in the abundances of environmentally responsive species.
272 Different types of data may require a different approach to identifying cluster sizes (Table 2), and
273 choice of cluster size can affect hypothesis tests. We discuss these issues more broadly in Appendix
274 S4, however the most important point is that decision about cluster size must be made independently
275 of the abundance data.

276 *Evaluating predictions in Table 1*

277 Benthic abundances, drift abundances and adult abundances of each species were converted into
278 ranks across sites in each data set. The analyses for predictions 1-4 used species as replicates and, for
279 some tests, all species in all data sets were combined. Thus the same species could potentially appear
280 multiple times, and we address below whether this occurred and affected the results.

281 Following Step 3 (Table 2) and using a cluster size of three (see Results), species in the benthic
282 data were classified as clustered, split, neither (distributions did not fit either of these extremes) or
283 uncommon (occurred at only one site). Taxa were classified as “split” if the two highest abundances
284 were split between the four most upstream sites and four most downstream sites *and* were not
285 adjacent. This definition of split represents an extreme difference in the location of high abundances,
286 and presents a strong test for this pattern. To test prediction 1, the numbers of taxa in each of the three
287 datasets that were classified as clustered or split were compared to the numbers expected by chance
288 using a chi-square test (Sokal & Rohlf 1995).

289 To test prediction 2, taxa classified as clustered were divided into the location where the highest
290 abundances occurred (upstream: wholly present in the four most upstream sites; downstream: wholly
291 present in the four most downstream sites; middle: clusters that straddled the transition zone between
292 these two types of environments by occurring within the four middle sites – see Appendix S1). We
293 used Fisher’s exact test (Sokal & Rohlf 1995) to examine whether the frequency of cluster locations
294 varied between the three data sets and chi-square to test whether they differed from the frequencies
295 expected by chance (0.33 for each: Appendix S1). We also used chi-square to test whether species
296 were clustered in multiple data sets more often than expected by chance. Finally, we identified species
297 occurring only within a cluster and not at any other locations (“restricted”) vs those where individuals
298 occurred also at other locations along channels. We tested whether the frequency of restricted clusters
299 differed between locations using Fisher’s exact test.

300 Following Step 4 (Table 2), to test for concordance between benthic and drift distributions
301 (prediction 3) and between benthic and adult distributions (prediction 4), we re-used the above
302 classifications, producing the following classes: Restricted upstream, highest upstream (but occurred
303 downstream), restricted downstream, highest downstream (but occurred upstream), split. Clustered
304 species (other than those in the middle, which were excluded) were spread between the first four
305 classes depending on whether they were restricted or not. Species that were neither split nor clustered
306 were mostly able to be classified into one of the first four classes. These five classes were also applied
307 to the drift and adult data to produce analogous classifications, allowing us to examine whether
308 benthic and drift rank abundances were concordant at gradient ends, as follows.

309 We first determined which benthic species were found in the drift (at comparable seasons in the
310 same creek) and analysed the resulting 2x5 contingency table to test whether drifting species were
311 equally common in each of the five benthic classifications. For species that drifted, we tested for
312 congruence between their benthic and drift classifications (prediction 3), and, for those where adults
313 were sampled, we tested for congruence between larval and adult distributions (prediction 4). For
314 some of these contingency tables, expected numbers of species in some cells were low (<5), which
315 violates an assumption underpinning confidence in probabilities associated with chi-square tests.
316 However, in most such tables, the effect on probabilities of chi-square tests is minor and thus of
317 concern only when probabilities are very near the significance boundary (i.e. $P = 0.05$) (Bradley *et al.*
318 1979).

319 In the final step, we compiled results by species to identify those fulfilling each of the predictions
320 in Table 1 for environmental effects or dispersal effects (prediction 5).

321 Results

322 ENVIRONMENTAL VARIATION

323 NMDS achieved a stress score of 0.08, which is an adequate representation of the environmental
324 data in two dimensions (Clarke & Warwick 2001). NMDS plots showed that both creeks had strong
325 environmental gradients from up- to downstream (Fig. 2). Although the creeks differed from each
326 other (PERMANOVA: Pseudo-F = 6.33, $P < 0.01$), the interaction between creek and group was non-
327 significant in all analyses ($P > 0.05$), showing that the environmental gradient was similar between the
328 creeks. The environmental gradients reflect changes in detrital food resources, provision of hard
329 surfaces from wood and bark, channel morphology, and in water quantities and qualities. Detrital
330 density declined dramatically downstream in accordance with an increasingly narrower and sparser
331 riparian zone (Appendix S3). Additionally, wide channels and shallow water depths meant both

332 creeks often registered water temperatures at downstream locations of up to 36 °C during summer,
333 which are extreme temperatures for most aquatic insects (Lancaster & Downes 2013).

334 Contrasts between adjacent pairs of sites were significant (i.e. $P < 0.05$) in only two of five tests,
335 whereas all three tests between trios of sites were significantly different. Thus, we chose a cluster size
336 of three.

337 PREDICTIONS 1 AND 2: CLUSTERING OF TAXA ALONG CHANNELS

338 Numbers of clustered taxa were significantly higher than expected in each data set (Table 3a). Split
339 taxa (e.g. Fig. 3e vs 3f) were significantly more common than expected in Hughes Ck in summer, and
340 were not different from expected in the other two data sets.

341 Of those species that were clustered, the proportions of up, middle and downstream species did not
342 vary between the three data sets. Overall there were many more upstream and many fewer middle and
343 downstream species than expected (Table 3b). A few species were clustered in two or more data sets,
344 and at a frequency significantly greater than expected by chance (Hughes Ck summer and Seven Cks
345 summer: $\chi^2 = 35.6$, $P < 0.001$; Hughes summer and Hughes spring: $\chi^2 = 105.1$, $P < 0.001$). Nevertheless,
346 70% of the observations in Table 3b were of different species. Species richness was far higher in
347 upstream locations, and over 50% of species clustered upstream were restricted to that zone (e.g. Fig.
348 3a vs Fig. 3b), whereas all middle and downstream species occurred upstream as well as in other
349 locations (e.g. Fig. 3c, d), a statistically significant difference (Table 3c). Thus, the incidence of
350 clustering was strongest upstream and became infrequent in a downstream direction.

351 PREDICTIONS 3 AND 4: BENTHIC CLASSIFICATIONS VS DRIFT AND ADULT CLASSIFICATIONS

352 Most species were found in the drift, but those absent from the drift were disproportionately likely
353 to be restricted upstream (Table 4), with a few having split distributions and virtually none in other
354 categories. The 14 occurrences of non-drifters that were restricted upstream were produced across 12
355 taxa - i.e. only two taxa showed the same pattern twice, whereas all other species showed different
356 patterns in different datasets.

357 Benthic and drift classifications were significantly associated overall (Table 5a), but only two of
358 the categories (restricted upstream, highest downstream) delivered this outcome. Of the 20 cases of
359 concordance for restricted upstream taxa (Table 5a), 19 were unique species, and only one species
360 produced two of these observations; all other species had different outcomes in different data sets.
361 Overall, drift classifications were not strictly concordant with benthic classifications in about one
362 third of cases (Table 5a).

363 Comparisons between juvenile and adult classifications showed a similar picture to the drift data
364 (Table 5b). Classifications were significantly associated but taxa restricted upstream contributed
365 mostly to this pattern. Approximately one third of cases had adult distributions that were incongruent

366 with corresponding benthic juvenile distributions. Caution is required however because the sample
367 size is small and the probability value of the chi-square is relatively close to 0.05.

368 PREDICTION 5: CONSISTENCY OF SPECIES' PATTERNS

369 Four of 43 species had distributions that were consistent with a strong effect of local environment:
370 they were clustered, showed consistent patterns between at least two data sets, and had concordance
371 between benthos and drift and between larvae and adults (Appendix S5). A further six species'
372 abundances were also strongly related to environmental conditions albeit with some information
373 missing. Alternatively, three species' abundances were consistent with strong effects of dispersal. The
374 remaining taxa had a mix of outcomes suggesting both environmental and dispersal effects played
375 significant roles.

376 Discussion

377 Separating the effects that environmental conditions and dispersal play in allowing species to
378 remain extant in altered landscapes is critically important. We developed a new approach that uses
379 several lines of evidence to test hypotheses about environmental and dispersal effects. In contrast to
380 most other comparable research (Logue *et al.* 2011), our method highlighted the role that dispersal
381 played in setting distributions, as well as environmental conditions, which we discuss first. We then
382 consider how our method may open the way for dispersal effects to be incorporated into management
383 decision-making.

384 Environmental effects were signalled by significant numbers of clustered rank abundances, and
385 some species occurred only within clusters. There were more species clustered in two or more data
386 sets than expected, and several of these species had consistently concordant benthic and disperser
387 distributions. The results suggest these species are strongly affected by environmental conditions.
388 Interestingly, these species were restricted almost entirely to upstream locations (found also by Brown
389 & Swan 2010). It has been hypothesized such patterns are due to relatively few drifters at upstream
390 locations and because adults fly across catchment boundaries to other upstream locations, not
391 downstream (Brown & Swan 2010). In our data, species were much more likely to be absent from the
392 drift if they were present upstream, which is consistent with this model. However, relatively few
393 species were strict, upstream specialists. Many species that were clustered upstream in one data set
394 occurred at other locations in another data set. Thus the prevalence of environmental effects at
395 upstream sites appeared to be a feature of those locations rather than a result of the identity of the
396 species that were present.

397 In contrast, dispersal clearly had pervasive effects along the whole environmental gradient. Split
398 distributions were not fewer than expected by chance in two data sets and more frequent than

399 expected in one. Clusters were very uncommon at middle and downstream locations rather than being
400 equally prevalent along the entire gradient. This pattern occurred because species that tolerated harsh,
401 downstream conditions also inhabited upstream locations and therefore had wide distributions. Thus,
402 environmental conditions at middle and downstream sites resulted in few clusters in those locations;
403 this pattern is expected also if there is frequent drift along channels (Prediction 2 in Table 2, also
404 found by Brown & Swan 2010). Moreover, many taxa had inconsistent distributions and lacked
405 consistent concordance between benthic and disperser distributions. These results highlight the
406 influence dispersal has on the distributions of many species in these rivers. Our findings are unusual,
407 but most comparable studies have relied on indirect estimates (using variance partitioning) and
408 dispersal effects may have been underestimated (Gilbert & Bennett 2010).

409 The variable outcomes for many species suggest that a trait-based approach – in which we seek
410 species characteristics associated with either a strong environmental or dispersal effect – is unlikely to
411 result in useful predictions. Species traits have been used to provide a second line of evidence about
412 dispersal effects (e.g. Brown & Swan 2010), but such traits are rarely independently tested (Verberk,
413 van Noordwijk & Hildrew 2013) and are often applied at coarse taxonomic levels (e.g. genera,
414 families or orders). For example, larvae of the mayfly *Baetis* are invariably classified as “good
415 dispersers” but larvae do not drift away from natal sites in some studies (e.g. Lancaster & Downes
416 2014a). In contrast to these indirect methods, disperser data permit direct tests. While the linear nature
417 of rivers makes capturing dispersers relatively straightforward, there are many other ecosystems in
418 which dispersal is also highly directional (terrestrial systems affected by prevailing winds Gillespie *et al.*
419 *et al.* 2012; e.g. coastal ecosystems affected by currents Moritz *et al.* 2013) and methods are available to
420 sample them. Having dispersal data means that species absences caused by dispersal constraints can
421 be distinguished from an incapacity of dispersers to colonise new locations following arrival. In the
422 latter situation, we expect to capture high numbers of dispersers in locations where the species is
423 uncommon; in the former we do not expect to capture many dispersing individuals where a species is
424 uncommon. These two hypotheses reflect different community models and have different implications
425 for managing biological diversity but are rarely evaluated (Logue *et al.* 2011).

426 Species interactions also determine distribution and abundance, but effects of interactions cannot
427 be identified using only survey data. Thus, in variation partitioning, patterns caused by interactions
428 may contribute spuriously to environmental or pure spatial variation depending upon whether or not
429 effects of interactions are correlated with particular environmental or spatial variables. Likewise, in
430 our method, species’ distributions determined by interactions may have contributed spuriously to
431 detecting patterns associated with environmental or dispersal effects. All methods using survey data
432 have to assume that species distributions are largely independent of each other, even though it is well-
433 known that species interactions can play strong roles in affecting abundances.

434 This short-coming notwithstanding, our approach provides a practical way of assessing dispersal
435 effects that may deliver benefits for management. Managers usually rely on indicators of ecosystem
436 health that are based almost entirely on measures of environmental condition. Our approach would
437 allow dispersal effects to be considered given that they can be derived from the same kind of data that
438 are typically used to assess the worth of ecosystem-level indices (basic survey data that help verify
439 whether species abundances correlate with environmental variables used to quantify ecosystem health
440 e.g. Kwok, Eldridge & Oliver 2011). A benefit of our method is that it is simpler to apply than many
441 multivariate analyses that rely on advanced statistical expertise to test hypotheses. Apart from the
442 analysis of environmental data to determine cluster size, the approach relies on setting up logical
443 predictions that are evaluated through the use of simple chi-square tests. If it is possible to gather
444 multiple, independent data sets and also to catch dispersing individuals, then the complete set of logic
445 described here can be applied, producing multiple lines of evidence about dispersal effects. Such
446 information could lead to alternative measures of ecosystem condition that capture the role that
447 dispersal plays and could be valuable in assessing whether some species or populations might be
448 particularly vulnerable to extirpation if dispersal routes are cut off.

449 Finally, dispersal of individuals from intact areas is a key aspect of successful restoration or
450 rehabilitation programs (Palmer, Ambrose & Poff 1997). Many restoration projects on rivers have
451 failed to deliver increases in diversity, and lack of dispersal to restored sites is a key potential
452 explanation (Palmer, Hondula & Koch 2014). The capacity to consider background amounts of
453 dispersal means restoration projects could be designed to take advantage of prospective dispersal
454 routes and realistic restoration targets identified. The results reported here, for example, are
455 encouraging for considering restoration works to improve species diversities of these and comparable
456 streams. Both rivers have relatively harsh environments in downstream locations, but upstream intact
457 areas with cool, detritus-rich conditions may contain reservoirs of species that routinely disperse into
458 downstream locales. Our results suggest that species from upstream locations would reach and
459 colonise repairs in downstream locations if suitable resources and conditions are achieved, thus
460 providing the impetus for such restoration efforts.

461 **Acknowledgements**

462 This research was supported by Australian Research Council Discovery grants (DP0772854,
463 DP120103145) awarded to B.J.D. and J.L. Thanks to Ash Macquoen and Al Danger for help in the
464 field and laboratory, and to local landholders for allowing access.

465 **Data accessibility**

466 All relevant data used in this manuscript are publicly available at Dryad Digital Repository: doi:
 467 10.5061/dryad.rt201 (Downes *et al.* 2016).

468 **References**

- 469 Allenby, R.B.J.T. & Slomson, A. (2011) *How To Count: An Introduction to Combinatorics*, 2nd edn.
 470 Chapman & Hall/CRC, Boca Raton, USA.
- 471 Anderson, M.J., Gorley, R.N. & Clarke, K.R. (2008) *PERMANOVA+ for PRIMER: guide to software*
 472 *and statistical methods*. PRIMER-e, Plymouth, UK.
- 473 Bengtsson, J. (2010) Applied (meta)community ecology: diversity and ecosystem services at the
 474 intersection of local and regional processes. *Community ecology: processes, models, and*
 475 *applications* (eds H.A. Verhoef & P.J. Morin), pp. 115-131. Oxford University Press, New York.
- 476 Bradley, D.R., Bradley, T.D., McGrath, S.G. & Cutcomb, S.D. (1979) Type-I error rate of the chi-
 477 square test of independence in R x C tables that have small expected frequencies. *Psychological*
 478 *Bulletin*, **86**, 1290-1297.
- 479 Brown, B.L. & Swan, C.M. (2010) Dendritic network structure constrains metacommunity properties
 480 in riverine ecosystems. *Journal of Animal Ecology*, **79**, 571-580.
- 481 Clarke, K.R. & Warwick, R.M. (2001) *Change in Marine Communities: An Approach to Statistical*
 482 *Analysis and Interpretation*. Plymouth Marine Laboratory, Plymouth.
- 483 Cottenie, K. (2005) Integrating environmental and spatial processes in ecological community
 484 dynamics. *Ecology Letters*, **8**, 1175-1182.
- 485 Dobbie, M.J. & Clifford, D. (2015) Quantifying uncertainty in environmental indices: an application
 486 to an estuarine health index. *Marine and Freshwater Research*, **66**, 95-105.
- 487 Downes, B.J. & Lancaster, J. (2010) Does dispersal control population densities in advection-
 488 dominated systems? A fresh look at critical assumptions and a direct test. *Journal of Animal*
 489 *Ecology*, **79**, 235–248.
- 490 Downes, B.J., Lancaster, J., Glaister, A & Bovill, W.D. (2016) Data from: A fresh approach reveals
 491 how dispersal shapes metacommunity structure in a human-altered landscape. Dryad Digital
 492 Repository: doi: 10.5061/dryad.rt201
- 493 Downes, B.J., Lancaster, J., Hale, R., Glaister, A. & Bovill, W.D. (2011) Plastic and unpredictable
 494 responses of stream invertebrates to leaf pack patches across sandy-bottomed streams. *Marine and*
 495 *Freshwater Research*, **62**, 394-403.
- 496 Fahrig, L. (2007) Non-optimal animal movement in human-altered landscapes. *Functional Ecology*,
 497 **21**, 1003-1015.

- 498 Gilbert, B. & Bennett, J.R. (2010) Partitioning variation in ecological communities: do the numbers
499 add up? *Journal of Applied Ecology*, **47**, 1071-1082.
- 500 Gillespie, R.G., Baldwin, B.G., Waters, J.M., Fraser, C.I., Nikula, R. & Roderick, G.K. (2012) Long-
501 distance dispersal: a framework for hypothesis testing. *Trends in Ecology and Evolution*, **27**, 47-
502 56.
- 503 Hynes, H.B.N. (1970) *The Ecology of Running Waters*. Liverpool University Press, Liverpool, UK.
- 504 Kwok, A.B.C., Eldridge, D.J. & Oliver, I. (2011) Do landscape health indices reflect arthropod
505 biodiversity status in the eucalypt woodlands of eastern Australia? *Austral Ecology*, **36**, 800-813.
- 506 Lancaster, J. & Downes, B.J. (2013) *Aquatic Entomology*. Oxford University Press, Oxford, UK.
- 507 Lancaster, J. & Downes, B.J. (2014a) Maternal behaviours may explain riffle-scale variations in some
508 stream insect populations. *Freshwater Biology*, **59**, 502-513.
- 509 Lancaster, J. & Downes, B.J. (2014b) Population densities and density–area relationships in a
510 community with advective dispersal and variable mosaics of resource patches. *Oecologia*, **176**,
511 985-996.
- 512 Legendre, P. & Legendre, L. (2012) *Numerical ecology*, 3rd English Edition edn. Elsevier,
513 Amsterdam.
- 514 Leibold, M.A. & Mikkelsen, G.M. (2002) Coherence, species turnover, and boundary clumping:
515 elements of meta-community structure. *Oikos*, **97**, 237-250.
- 516 Logue, J.B., Mouquet, N., Peter, H., Hillebrand, H. & The Metacommunity Working Group (2011)
517 Empirical approaches to metacommunities: a review and comparison with theory. *Trends in*
518 *Ecology and Evolution*, **26**, 482-491.
- 519 Lowe, W.H. & McPeck, M.A. (2014) Is dispersal neutral? *Trends in Ecology and Evolution*, **29**, 444-
520 450.
- 521 Montoya, D., Rogers, L. & Memmott, J. (2012) Emerging perspectives in the restoration of
522 biodiversity-based ecosystem services. *Trends in Ecology & Evolution*, **27**, 666-672.
- 523 Moritz, C., Meynard, C.N., Devictor, V., Guizien, K., Labrune, C., Guarini, J.-M. & Mouquet, N.
524 (2013) Disentangling the role of connectivity, environmental filtering, and spatial structure on
525 metacommunity dynamics. *Oikos*, **122**, 1401-1410.
- 526 Palmer, M.A., Ambrose, R.F. & Poff, N.L. (1997) Ecological theory and community restoration
527 ecology. *Restoration Ecology*, **5**, 291-300.
- 528 Palmer, M.A., Hondula, K.L. & Koch, B.J. (2014) Ecological restoration of streams and rivers:
529 Shifting strategies and shifting goals. *Annual Review of Ecology, Evolution and Systematics*, **45**,
530 247-269.
- 531 Presley, S.J., Higgins, C.L. & Willig, M.R. (2010) A comprehensive framework for the evaluation of
532 metacommunity structure. *Oikos*, **119**, 908-917.

- 533 Reid, D.J., Lake, P.S., Quinn, G.P. & Reich, P. (2008) Association of reduced riparian vegetation
 534 cover in agricultural landscapes with coarse detritus dynamics in lowland streams. *Marine and*
 535 *Freshwater Research*, **59**, 998-1014.
- 536 Rodgers, K.S., Kido, M.H., Jokiel, P.L., Edmonds, T. & Brown, E.K. (2012) Use of integrated
 537 landscape indicators to evaluate the health of linked watersheds and coral reef environments in the
 538 Hawaiian Islands. *Environmental Management*, **50**, 21-30.
- 539 Scheffers, B.R., Joppa, L.N., Pimm, S.L. & Laurence, W.F. (2012) What we know and don't know
 540 about Earth's missing biodiversity. *Trends in Ecology and Evolution*, **27**, 501-510.
- 541 Sokal, R.R. & Rohlf, F.J. (1995) *Biometry: the principles and practice of statistics in biological*
 542 *research*. W.H. Freeman, New York.
- 543 Verberk, W.C.E.P., van Noordwijk, C.G.E. & Hildrew, A.G. (2013) Delivering on a promise:
 544 integrating species traits to transform descriptive community ecology into a predictive science.
 545 *Freshwater Science*, **32**, 531-547.

546

547 **Supporting information**

548 Additional Supporting Information may be found in the online version of this article:

549 **Appendix S1** Calculating the proportion of taxa expected to clustered or split by chance

550 **Appendix S2** Outcome of variation partitioning

551 **Appendix S3** Site locations, measurement of wood, detritus and riparian tree cover, discharge, and
 552 sampling details for drift and adults

553 **Appendix S4** Considerations for deciding on cluster size using environmental data

554 **Appendix S5** Distribution patterns of individual species

555

556 Table 1. Predictions associated with each of two general views of metacommunities for stream
 557 species, the first of which reflects overarching effects of local, environmental conditions within
 558 channels (Fig. 1a), which change in relatively predictable ways from up- to downstream, and the
 559 second of which reflects primarily the effects of dispersal, both in the drift (Fig. 1b) or by adults (Fig.
 560 1c), which disperse in search of conditions suitable for mate location and egg-laying. Material in
 561 italics provides explanation of predictions where necessary.

Type of patterns	<u>Local environmental effects</u>	<u>Dispersal effects</u>
	Local environmental conditions predominately determine benthic abundances; dispersal has few persistent effects	Dispersal effects dominate; local environmental conditions explain comparatively little of species' distributions

1. Incidence of clustering	Most species cluster, i.e. attain highest abundances at sites with similar environments. <i>In streams, these are likely to be sites that are adjacent to each other</i>	Most species show patchy distributions unrelated to environmental gradients and/or highest abundances occur at locations that are environmentally dissimilar. <i>In streams, such sites will be well separated along channels.</i>
2. Spread of clustering	Clustering is apparent along all parts of the channel (up-, mid- and downstream locations). <i>The effects of local environmental conditions should be present along the whole gradient.</i>	If drift dispersal is frequent or often long distance, then clustering, if present at all, becomes less frequent in a downstream direction. <i>Frequency of clustering weakens due to the successful arrival of increasing numbers of species drifting from various upstream locations.</i>
3. Benthic and drift concordance	Most drifting species are consistently and predominantly found in the zone in which they occur benthically. <i>Drifting out of the zone the species inhabits would be maladaptive.</i>	Drift abundances are often not concordant with benthic abundances. <i>Drifters may frequently be in search of new resources, for example.</i>
4. Adult and juvenile concordance	Adult and juvenile distributions consistently show concordance. <i>We expect most adults to be trapped near where their larvae do best.</i>	Adults and juvenile distributions may often be non-concordant. <i>Adult females may lay eggs in places driven by dispersal and choices of oviposition habitat, so larval distribution is not related to local, in-channel environmental conditions.</i>
5. Consistency across data sets	Most species have consistent patterns in both benthic and dispersers distributions with respect to environmental gradients when examined across different rivers	Most species' distributions have little consistency with respect to environmental gradients in different rivers or at different times

and times.

562 Table 2. A summary of practical steps to analyse data, following the logical approach outlined in the
 563 text
 564

Step 1. Develop predictions for species abundance data that relate directly to hypotheses, e.g. under environmental and dispersal explanations (i.e. create Table 1)

- (a) Make predictions for surveyed abundances of species (e.g. Table 1, predictions 1 & 2) and for dispersers (e.g. prediction 3)
- (b) Consider whether further predictions arise due to nature of the ecosystem or species' biology (e.g. dispersal effects in prediction 3, prediction 4)
- (c) If available, make predictions for multiple data sets (e.g. prediction 5)

Step 2. Using the environmental data, decide cluster size, c .¹

- (a) Use NMDS to display differences among sites²
- (b) Order sites by their environmental similarity³
- (c) Group sites that are most similar environmentally⁴
- (d) Use PERMANOVA with planned contrasts to assess the degree of differences among groups; repeat using different numbers of sites⁵
- (e) Identify the group size that delivers the greatest degree of statistical significance among groups and designate as cluster size, c

Step 3. Statistically test predictions about abundance of surveyed species

- (a) Decide on an appropriate definition of a split distribution
- (b) Convert abundances of each species to ranks across sites
- (c) Order each species' ranks according to the order of site similarity (2b, above)
- (d) Classify species into those that are clustered, split, or neither (using *a priori* definitions, above)
- (e) Using combinatorics, calculate the number of species expected to occur in clusters and splits by chance
- (f) Using chi-square tests, compare expected and observed numbers of species in each category

Step 4. When dispersal data are available

- (a) Create classes of distribution derived from clusters and splits (e.g. using location of clusters along environmental gradient, whether species are restricted to clusters or not), which will be applied to both survey data and dispersal data⁶
- (b) Using ranked data (3b above), classify species in surveyed abundance data according to distribution classes (4a, above)
- (c) Repeat steps 3b and 3c for dispersal data to ranks across sites, and then classify species into distribution classes (4a above)

(d) Create a 2-way table having each distributional class for survey abundance data crossed with each distributional class for dispersal data; sum the number of species falling into each cell of the table

(e) Using contingency table analysis, test for association between classifications of abundance and dispersal data

565

566 ¹Any legitimate, analytical approach can be used as long as decisions about cluster size are made
567 independently of species abundance data. Note that our method is designed for community data where
568 dispersal between sites is feasible; it is not designed to address questions about biogeographical
569 patterns or geographic range (i.e. very large spatial scales).

570

571 ²With large data sets, cluster analysis can help assess similarity because it provides a visual display of
572 the major divisions among sites in their degree of similarity.

573

574 ³We consider the simple situation where sites can be ordered along a single dimension; other data sets
575 may be more complex and require patterns (and combinatoric formulae) that consider clusters or other
576 patterns in two or more dimensions.

577

578 ⁴Group sizes (and therefore prospective cluster size) are likely to be related to sampling design
579 relative to environmental gradients.

580

581 ⁵Ecosystems with diffuse or conflicting gradients may not deliver a definitive decision about a single
582 cluster size. See material in Appendix S4.

583

584 ⁶As in this paper, multiple classes of distributions derived from clusters and splits (e.g. locations along
585 gradients where clusters occurred) are likely needed to capture distributions along environmental
586 gradients.

587

588

589

590 Table 3. (a) Numbers of taxa classified as clustered, split, neither, or uncommon in each of the three
591 data sets, with expected values for clustered and split classes given in brackets. Also given are the
592 values for chi-square tests (χ^2) of whether the numbers of clustered (Clust χ^2) or split (Split χ^2) species
593 differ from those expected by chance (see Appendix S1). ***, probability $P < 0.001$; **, $P < 0.01$; *, P
594 < 0.05 ; ns, $P > 0.05$. (b) For clustered species only, the observed numbers of upstream (U), middle
595 (M) and downstream (D) species in each dataset with a chi-square test of whether numbers differ

596 between datasets. Given this is non-significant, numbers were summed and compared to the numbers
 597 expected by chance (as determined by expected proportions 0.33, 0.33, 0.33, respectively, Appendix
 598 S2) using a chi-square test. Symbols as for (a). (c) For clustered species only, numbers of taxa
 599 restricted to a zone (upstream or middle + downstream) vs the numbers found outside of zones and a
 600 Fisher exact test of the differences in proportions. Symbols as for (a).

601 (a)

Data set	Clustered	Split	Neither	Uncommon	Total	Clust. χ^2	Split χ^2
Hughes Ck Summer	14 (5.25)	10 (4.39)	14	3	41	14.58***	7.16**
Seven Cks Summer	10 (4.04)	4 (3.75)	15	6	35	8.81**	0.02ns
Hughes Ck Spring	14 (4.47)	6 (3.75)	10	5	35	20.32***	1.351ns

602

603

604

605

606 (b)

	U	M	D	Fisher's exact
Hughes Creek Summer	11	1	2	P=0.80ns
Seven Creeks Summer	8	1	1	
Hughes Creek Spring	13	0	1	
Observed	32	2	4	χ^2 44.42***
Expected	12.7	12.7	12.7	

607

608

609

610 (c)

	U	M+D	Fisher's exact test

Restricted	18	0	P=0.021*
Not restricted	14	6	

611

612

613 Table 4. The numbers of species (across all three data sets) in different benthic classifications *vs*
 614 presence/absence in the drift at commensurate times/locations. Because only two taxa were restricted
 615 downstream, they were combined with taxa having their highest abundance downstream for this
 616 analysis. A test of independence showed there was a significant association ($\chi^2 = 10.3$, $p < 0.025$). The
 617 standardized residuals were all < 2.0 , indicating that all cells contributed to the result, but the highest
 618 (1.88) was for non-drifting taxa that were restricted upstream, whereas all other cells contained
 619 negative residuals.

620

Benthic classification	Drift presence		TOTAL
	Present	Absent	
Restricted upstream	24	14	38
Highest upstream	16	1	17
Restricted or highest downstream	11	1	12
Split	26	8	34
TOTAL	77	24	101

621

622

623 Table 5. (a) For species found at least once in the drift, benthic classification *vs* drift classification
 624 and (b) juvenile *vs* adult classification across the three data sets. Benthic taxa that were restricted
 625 downstream were added to those found in highest abundances downstream because of low sample
 626 size ($n=1$ only; this species had a split classification in the drift). There is a significant association
 627 between benthic and drift classification ($\chi^2 = 33.4$, $p = 0.001$) and between benthic (i.e. juvenile) and
 628 adult classification ($\chi^2 = 22.1$, $p = 0.035$). Numbers in italics indicate cells where the standardized
 629 deviates were > 2 , indicating a significant contribution to the test outcome. Numbers in bold indicate
 630 cells where classifications were not in strict concordance.

631

632 (a) Benthic classification *vs* drift classification

Benthic	Drift classification					TOTAL
	Restricted upstream	Highest upstream	Split	Highest downstream	Restricted downstream	

Restricted upstream	20	1	1	1	1	24
Highest upstream	7	3	4	1	1	16
Split	8	5	5	5	3	26
Restricted to or highest downstream	0	1	2	5	3	11
TOTAL	35	10	12	12	8	77

633

634

635 (b) Juvenile (benthic) classification vs adult classification

Benthic	Adult classification					TOTAL
	Restricted upstream	Highest upstream	Split	Highest downstream	Restricted downstream	
	Restricted upstream	4	1	0	0	
Highest upstream	4	5	0	0	0	9
Split	0	2	4	1	1	8
Restricted to or highest downstream	0	1	2	1	1	5
TOTAL	8	9	6	2	2	27

636 **Figure captions**

637 **Fig. 1** Diagrams illustrating some ways in which environmental gradients and dispersal could drive
638 community structure, in this case along rivers where upstream, middle and downstream locations have
639 differing environmental conditions reflecting gradients in important environmental variables. The
640 stars, circles and squares, both filled and open, indicate different species. The presence of species and
641 their approximate abundances in the stream benthos (symbols within squares), in the drift and as
642 adults (whether terrestrial or aquatic) are indicated by the number and type of symbols organised
643 under the appropriate headings. Arrow direction and length signify dispersal direction and distances.
644 **A.** Community structure in the benthos reflects strict effects of environmental gradients and dispersal
645 plays no significant role in driving species abundances. Each location or zone (upstream, middle,
646 downstream) comprises mostly species that are restricted to that zone. Species that drift remain within
647 the appropriate zone, and adults, particularly those in the terrestrial environment, disperse only into or
648 within the appropriate zone. **B.** One way in which drift dispersal could affect benthic distributions.
649 Juveniles hatch from eggs laid into particular zones in the benthos and hence species are initially

650 clustered but some individuals drift and settle downstream. Numbers of drifting species increase
651 cumulatively in a downstream direction, and hence clustering is weakened in a downstream direction.
652 C. Ways in which abundances of adults can be mismatched with juvenile distributions: (i) Adults are
653 abundant at sites with low numbers of juveniles compared to other sites (open circles and squares)
654 because adult and juvenile requirements are not correlated across locations. (ii) Adults are in high
655 abundance where juveniles are absent either because eggs have yet to hatch or adults fail to reproduce
656 successfully (filled squares and stars). There are no arrows because adults may potentially disperse
657 from anywhere.

658 **Fig. 2** NMDS plot of eight sites each in Hughes Creek and Seven Creeks based on a resemblance
659 matrix created from environmental variables. The two-dimensional stress value indicates the NMDS
660 plots are a good fit to the data. Numbers above points are site numbers with 1 the most upstream and 8
661 the most downstream.

662 **Fig. 3** Representative examples of types of benthic distribution patterns, taken from the Hughes Creek
663 samples from summer. (a) *Notalina fulva*: restricted to upstream sites, (b) *Triplectides ciuskus*: most
664 abundant upstream but found downstream, (c) *Ecnomus continentalis*: most abundant at middle sites,
665 (d) *Cheumatopsyche* sp. AV2: most abundant downstream but found upstream, (e) *Austrogomphus*
666 sp.: a split distribution. (f) An example of a species with inconsistent patterns - *Austrogomphus* sp. in
667 Seven Creeks in summer, where it was restricted to upstream sites, as compared to Hughes Creek in
668 summer where it had a split distribution (e).





