

A unified dataset of co-located sewage pollution, periphyton, and benthic macroinvertebrate community and food web structure from Lake Baikal (Siberia)

Journal:	<i>Limnology and Oceanography Letters</i>
Manuscript ID	Draft
Wiley - Manuscript type:	Data Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Meyer, Michael; Washington State University, School of the Environment Ozersky, Ted; University of Minnesota Duluth, Large Lakes Observatory Woo, Kara; Sage Bionetworks Shchapov, Kirill; University of Minnesota Duluth, Large Lakes Observatory Galloway, Aaron; University of Oregon, Schram, Julie; University of Oregon Oregon Institute of Marine Biology Snow, Daniel; University of Nebraska-Lincoln Timofeev, Maxim; Irkutsk State University Karnaughov, Dmitrii; Irkutsk State University Brousil, Matthew; Washington State University, Hampton, Stephanie; Washington State University, Center for Environmental Research, Education and Outreach
Search Terms:	pharmaceuticals, microplastics, fatty acids, stable isotopes, amphipod, spirogyra
Abstract:	Sewage released from lakeside development can introduce nutrients and micropollutants that can restructure aquatic ecosystems. Lake Baikal, the world's most ancient, biodiverse, and voluminous lake, has been experiencing localized sewage pollution from lakeside settlements. Nearby increasing filamentous algal abundance suggests benthic communities are responding to this localized pollution. We surveyed 40-km of Lake Baikal's southwestern shoreline 19-23 August 2015 for sewage indicators, including pharmaceuticals, personal care products, and microplastics with co-located periphyton, macroinvertebrate, stable isotope, and fatty acid samplings. Unique identifiers corresponding to sampling locations are retained throughout all data files to facilitate interoperability among the dataset's 150+ variables. The data are structured in a tidy format (a tabular arrangement familiar to limnologists) to encourage reuse. For Lake Baikal studies, these data can support continued monitoring and research efforts. For global studies of lakes, these data can help characterize sewage prevalence and ecological consequences of anthropogenic disturbance across spatial scales.

SCHOLARONE™
Manuscripts

Scientific Significance Statement

We present a unified dataset of co-located benthic littoral nutrient concentrations, sewage indicators, algal and macroinvertebrate community abundance, stable isotopes, and fatty acids from Lake Baikal (Siberia). While researchers have studied Baikal's exceptionally diverse endemic taxa for centuries, this product is the first publicly available dataset of Baikal benthic amphipod species abundance as well as amphipod fatty acid profiles in a machine-readable format with standardized metadata. Furthermore, with over 150 co-located variables, this dataset is the most extensive, publicly available description of Baikal's nearshore benthic communities and food webs. The data are structured in a `oceanogr` format and incorporate a scripted, sequential workflow, enabling the dataset to either supplement current monitoring efforts or provide data for syntheses across systems.

For Review Only

A unified dataset of co-located sewage pollution, periphyton, and benthic macroinvertebrate community and food web structure from Lake Baikal (Siberia)

4 Michael F. Meyer¹*

5 Ted Ozersky²

6 Kara H. Woo³

7 Kirill Shchapov²

8 Aaron W. E. Galloway⁴

9 Julie B. Schram⁴

10 Daniel D. Snow⁵

11 Maxim A. Timofeyev⁶

12 Dmitry Yu. Karnaukh

13 Matthew R. Brousil³

14 Stephanie E. Hampton³

¹⁶ ^{1.} School of the Environment, Washington State University, Pullman, WA, USA

². Large Lakes Observatory, University of Minnesota - Duluth, Duluth, MN, USA

³ Center for Environmental Research, Education, and Outreach, Washington State University, Pullman, WA, USA

⁴ Oregon Institute of Marine Biology, University of Oregon, Charleston, OR, USA

²¹ 5. School of Natural Resources, University of Nebraska-Lincoln, Lincoln, NE, USA

²² 6. Biological Research Institute, Irkutsk State University, Irkutsk, Irkutsk Oblast, Russia

24 *corresponding author: michael.f.meyer@wsu.edu

Author Contribution Statement

27 Conceptualized the project: MFM, SEH, TO

28 Collected samples in the field: MFM, TO, KHW, SEH

29 Processed samples: MFM, KS, JBS, DDS, TO, AWEG, SEH

30 Wrote and Reviewed R scripts: MFM, MRB, KHW

31 Data management: MFM, MRB

32 Wrote and edited the manuscript: All authors

33 Approved the final manuscript: All authors

35 **Grant sponsor information:**

36 Funding was provided by the National Science Foundation (NSF-DEB-1136637) to S.E.H., a Fulbright
37 Fellowship to M.F.M., a NSF Graduate Research Fellowship to M.F.M. (NSF-DGE-1347973), and the
38 Russian Ministry of Science and Education (N FZZE-2020-0026; N FZZE-2020-0023).

Key Words: pharmaceuticals, microplastics, fatty acids, stables isotopes, amphipod, mollusk, diatom, spirogyra

43 URL of the Dataset and Metadata with permanent identifier:

- 44 • Environmental Data Initiative: doi:10.6073/pasta/76f43144015ec795679bac508efa044b
45 • Open Science Framework: <https://doi.org/10.17605/OSF.IO/9TA8Z>

46

47 **Code URL with permanent identifier:**

- 48 • Environmental Data Initiative: doi:10.6073/pasta/76f43144015ec795679bac508efa044b
49 • Open Science Framework: <https://doi.org/10.17605/OSF.IO/9TA8Z>

50 **Measurement(s):** Chlorophyll a, Fatty Acids, Pharmaceuticals and Personal Care Products,
51 Microplastics, Periphyton community abundance, benthic macroinvertebrate abundance, Stable
52 Isotopes, nitrate, ammonium, total phosphorus

53 **Technology Type(s):** GC/MS, LC/MS, Spectrophotometry

54 **Temporal range:** 19 – 23 August 2015

55 **Frequency or sampling interval:** single snapshot in time

56 **Spatial scale:** site-based

57

58 **Abstract (150 of 150 words)**

59

60 Sewage released from lakeside development can introduce nutrients and micropollutants that can
61 restructure aquatic ecosystems. Lake Baikal, the world's most ancient, biodiverse, and voluminous
62 lake, has been experiencing localized sewage pollution from lakeside settlements.

63 Nearby increasing filamentous algal abundance suggests benthic communities are responding to
64 this localized pollution. We surveyed 40-km of Lake Baikal's southwestern shoreline 19-23 August
65 2015 for sewage indicators, including pharmaceuticals, personal care products, and microplastics
66 with co-located periphyton, macroinvertebrate, stable isotope, and fatty acid samplings. Unique
67 identifiers corresponding to sampling locations are retained throughout all data files to facilitate
68 interoperability among the dataset's 150+ variables. The data are structured in a tidy format (a
69 tabular arrangement familiar to limnologists) to encourage reuse. For Lake Baikal studies, these
70 data can support continued monitoring and research efforts. For global studies of lakes, these data
71 can help characterize sewage prevalence and ecological consequences of anthropogenic disturbance
72 across spatial scales.

73

74 **Background and Motivation**

75

76 Globally, sewage pollution is a common and often concentrated source of nitrogen and phosphorus
77 inputs that can reshape aquatic ecosystems. Sewage inputs are often associated with increased
78 primary production (Edmondson 1970; Moore et al. 2003), which can eventually lead to nuisance
79 algal blooms (Hall et al. 1999; Lapointe et al. 2015). Even in instances where sewage pollution is
80 mitigated, restoring systems can be complicated and necessitate system-specific (Jeppesen et al.
81 2005), long-term mitigation strategies (Hall et al. 1999; Tong et al. 2020). As such, effective
82 sewage monitoring can require merging a suite of chemical, biological, and ecological data to
83 synthesize locations and timing of inputs with associated shifts in ecological communities
84 (Rosenberger et al. 2008; Hampton et al. 2011).

85

86 Definitively identifying sewage as the source of excess nutrients in a system can be challenging.
87 Nutrients can originate from multiple sources, such as agriculture (Powers et al. 2016) or melting
88 permafrost (Turetsky et al. 2000), which can obfuscate wastewater signals. Unlike nutrients,
89 sewage-specific indicators, such as enhanced $\delta^{15}\text{N}$ stable isotope signatures (Costanzo et al. 2001;

90 Camilleri and Ozersky 2019), pharmaceuticals and personal care products (PPCPs) (Bendz et al.
91 2005; Rosi-Marshall and Royer 2012; Meyer et al. 2019) and microplastics (Barnes et al. 2009), can
92 be highly specific to human wastewater. Accordingly, sewage-associated micropollutants have
93 garnered global attention for their usefulness in identifying presence and quantifying magnitude of
94 wastewater inputs. While indicators may accumulate differentially in certain taxa (Gartner et al.
95 2002; Green 2016; Vendel et al. 2017; Richmond et al. 2018), acutely dangerous concentrations are
96 not common in most aquatic systems (Kolpin et al. 2002; Focazio et al. 2008; Yang et al. 2018).
97 However, chronic exposure to microplastics and PPCPs at even minute concentrations (e.g., µg/L)
98 can still disrupt ecological processes (Richmond et al. 2017). For example, oxazepam can increase
99 feeding rate and decrease sociability of river perch (Brodin et al. 2013), and microplastics can
100 release dissolved organic carbon, thereby altering microbial communities (Romera-Castillo et al.
101 2018). The pervasiveness and diversity of sewage-associated micropollutants in tandem with their
102 potency as ecologically disrupting compounds necessitates investigation within and across systems,
103 thereby enabling synthesis of how micropollutants alter ecosystems.

104

105 When assessing biological responses to increased nutrient loading, littoral benthic algal, and
106 macroinvertebrate communities often respond most markedly, as their physical proximity to the
107 shoreline puts them in the path of sewage pollution entering the lake (Rosenberger et al. 2008;
108 Hampton et al. 2011). Filamentous algae, for example, can quickly increase in abundance near
109 sewage sources (Rosenberger et al. 2008; Hampton et al. 2011). As algal communities change, food
110 webs can also restructure. For example, change in algal communities can alter the nutritional value
111 of primary producers or cause changes in the relative abundance of different feeding groups (e.g.,
112 increased representation of detritivores). Among the suite of food quality metrics, availability of
113 essential fatty acids (EFAs) offers a nuanced understanding of food quality as primary producers
114 usually maintain consistent EFA signatures (Taipale et al. 2013) and consumers acquire EFAs by
115 grazing (Dalsgaard et al. 2003) or trophic upgrading (Sargent and Falk-Petersen 1988; Dalsgaard et
116 al. 2003).

117

118 Together, food web structure, community composition, and sewage indicator data can be powerful
119 tools to assess biological impacts of sewage pollution. Despite their utility, these data are not often
120 available for many limnological systems. PPCPs, for example, have historically been less measured
121 in lake environments (Meyer et al. 2019). In instances where data are available, efficiently merging
122 disparate data into a single, analytically-friendly format can be challenging and require relatively
123 complex, computationally intensive workflows (Meyer et al. 2020a).

124

125 To offer a template for harmonizing sewage indicator and biological data, we present a unified data
126 product, which contains disparate data collected from 14 littoral and 3 pelagic sites at Lake Baikal
127 from 19 through 23 August 2015 (Figure 1). Located in Siberia, Lake Baikal is the oldest, most
128 voluminous, and deepest freshwater lake in the world (Hampton et al. 2018). Lake Baikal also has
129 the global distinction of being the most biodiverse lake, with the highest endemism (Moore et al.
130 2009). The lake is experiencing rapid warming associated with climate change, including decrease
131 in ice cover duration (Moore et al. 2009), and it exhibits offshore plankton community changes
132 associated with warming (Hampton et al. 2008; Katz et al. 2015; Izmest'eva et al. 2016). Less is
133 known of the change occurring in the nearshore of Lake Baikal, where not only climatic changes
134 (Swann et al. 2020) but also human activity (Timoshkin et al. 2018) alter the environment.
135 Nearshore change is particularly important to understand in Lake Baikal, since the majority of the

136 lake's biodiversity and endemic species occur in the littoral zone (Kozhova and Izmest'eva 1998).
137 While Lake Baikal's pelagic zone is generally ultra-oligotrophic (Yoshida et al. 2003; O'Donnell et
138 al. 2017), littoral areas abutting lakeside settlements have recently shown distinct signs of
139 eutrophication (Timoshkin et al. 2016; Volkova et al. 2018).

140
141 As a means of identifying sewage from small, concentrated lakeside towns and the associated
142 ecological responses, we assembled a dataset consisting of over 150 variables collected at 14 littoral
143 and 3 pelagic sampling sites. We structured the dataset in a tidy format, where each row is a
144 sample, each column is a variable, and each CSV file is an observable unit, where more similar
145 variables are contained within an individual file (Wickham 2014). Independent CSV files can be
146 merged using unique locational identifiers as relational keys, enabling future researchers to
147 customize analyses around a particular suite of variables. As a result of the dataset's
148 interoperability, reproducibility, and extensive variable content, it is well poised for future reuse as
149 supporting evidence of sewage pollution in Lake Baikal. Additionally, the data's flexibility and
150 consistent structure enable it to be merged with similar datasets, so as to synthesize biological
151 responses to sewage across systems and scales.

152
153 To our knowledge, no raw data on Lake Baikal macroinvertebrates, periphyton, or nearshore water
154 quality are public in a machine-readable format, for any variable (i.e. abundance, fatty acid content,
155 stable isotopes, nutrient and pollutant concentration), and no georeferenced data on pharmaceuticals
156 and personal care products or microplastics appear to be publicly available for any boreal, subarctic,
157 or arctic lakes or rivers in Siberia. Thus, the dataset fills a substantial gap for future studies,
158 providing a window into nearshore biotic assemblages and water quality in a unique, ancient
159 ecosystem that holds 20% of the world's liquid surface water (Moore et al. 2009).

160 161 Data Description

162
163 The final, replicate-level data products are available on the Environmental Data Initiative (EDI),
164 where they can be freely accessed without potential barriers such as paywalls or account
165 registrations. The final data are provided as 11 separate CSV files, each structured in a tabular
166 format and containing a "site" column that can be used to merge tables. The repository also
167 contains a compressed folder of R scripts (scripts.tar.gz), which were used in the main analysis of
168 the dataset (Meyer et al., Under Review).

169
170 chlorophylla.csv

171
172 This file contains chlorophyll a concentrations as well as fluorometric corrections for each littoral
173 and pelagic sampling location.

174
175 *site*
176 Unique alphanumeric identifier for a sampling location.

177
178 *replicate*
179 Replicate number.

180
181 *filtered_volume_ml*

182 Lake water volume filtered in milliliters for a given replicate.
183
184 *sample_volume_ml*
185 Sample volume filtered for chlorophyll a extraction.
186
187 *raw_fluo*
188 Raw, uncorrected fluorometric reading for chlorophyll analysis.
189
190 *adjusted_raw*
191 Corrected fluorometric reading for chlorophyll analysis.
192
193 *chl_conc*
194 Chlorophyll a concentration in milligrams per liter.
195
196 distance_weighted_population_metrics.csv
197
198 This file contains human population data for each sampled location. Although the majority of sites
199 do not have adjacent shoreline human developments, we calculated inverse distance weighted
200 (IDW) population for each sampling location. IDW population is a generalized representation of the
201 size of and proximity to a sampling location's neighboring human settlements. A full description of
202 the methods used to calculate IDW population can be found in the companion manuscript Meyer et
203 al. (Under Review).
204
205 *site*
206 Unique alphanumeric identifier for a sampling location.
207
208 *distance_weighted_population*
209 Inverse distance weighted population for a given sampling location and estimated as number of
210 people. Because this interpolation process is a function of the size of and proximity to neighboring
211 developed sites, values can contain decimal values.
212
213 fatty_acid.csv
214
215 This file contains fatty acid concentrations for various benthic macroinvertebrate genera,
216 periphyton, and endemic *Draparnaldia* spp. benthic algae collected from the 14 littoral sampling
217 locations.
218
219 *site*
220 Unique alphanumeric identifier for a sampling location.
221
222 *Genus*
223 Genus of the analyzed organism.
224
225 *Species*
226 Species of the analyzed organism. When organism was identified solely to genus, the Species value
227 is NA.

228
229 *c12_0*
230 Concentration of 12:0 fatty acid as micrograms of fatty acid per milligram of tissue.
231
232 *i_14_0*
233 Concentration of i-14:0 fatty acid as micrograms of fatty acid per milligram of tissue.
234
235 *c14_0*
236 Concentration of 14:0 fatty acid as micrograms of fatty acid per milligram of tissue.
237
238 *c14_4w5*
239 Concentration of 14:4n-5 fatty acid as micrograms of fatty acid per milligram of tissue.
240
241 *i_15_0*
242 Concentration of i-15:0 fatty acid as micrograms of fatty acid per milligram of tissue.
243
244 *a_15_0*
245 Concentration of a-15:0 fatty acid as micrograms of fatty acid per milligram of tissue.
246
247 *c15_0*
248 Concentration of 15:0 fatty acid as micrograms of fatty acid per milligram of tissue.
249
250 *c15_1w7*
251 Concentration of 15:1 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
252
253 *i_16_0*
254 Concentration of i-16:0 fatty acid as micrograms of fatty acid per milligram of tissue.
255
256 *c16_0*
257 Concentration of 16:0 fatty acid as micrograms of fatty acid per milligram of tissue.
258
259 *c16_1w9*
260 Concentration of 16:1 ω 9 fatty acid as micrograms of fatty acid per milligram of tissue.
261
262 *c16_1w8*
263 Concentration of 16:1 ω 8 fatty acid as micrograms of fatty acid per milligram of tissue.
264
265 *c16_1w7*
266 Concentration of 16:1 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
267
268 *c16_1w6*
269 Concentration of 16:1 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
270
271 *c16_1w5*
272 Concentration of 16:1 ω 5 fatty acid as micrograms of fatty acid per milligram of tissue.
273

274 *i_17_0*
275 Concentration of i-17:0 fatty acid as micrograms of fatty acid per milligram of tissue.
276
277 *a_17_0*
278 Concentration of a-17:0 fatty acid as micrograms of fatty acid per milligram of tissue.
279
280 *c17_0*
281 Concentration of 17:0 fatty acid as micrograms of fatty acid per milligram of tissue.
282
283 *c17_1w7*
284 Concentration of 17:1 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
285
286 *c16_2w7*
287 Concentration of 16:2 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
288
289 *c16_2w6*
290 Concentration of 16:2 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
291
292 *c16_2w4*
293 Concentration of 16:2 ω 4 fatty acid as micrograms of fatty acid per milligram of tissue.
294
295 *c16_3w6*
296 Concentration of 16:3 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
297
298 *c16_3w4*
299 Concentration of 16:3 ω 4 fatty acid as micrograms of fatty acid per milligram of tissue.
300
301 *c16_3w3*
302 Concentration of 16:3 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
303
304 *c16_4w3*
305 Concentration of 16:4 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
306
307 *c16_4w1*
308 Concentration of 16:4 ω 1 fatty acid as micrograms of fatty acid per milligram of tissue.
309
310 *c18_0*
311 Concentration of 18:0 fatty acid as micrograms of fatty acid per milligram of tissue.
312
313 *c18_1w9*
314 Concentration of 18:1 ω 9 fatty acid as micrograms of fatty acid per milligram of tissue.
315
316 *c18_1w7*
317 Concentration of 18:1 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
318
319 *c18_2w6t*

- 320 Concentration of 18:2 ω 6t fatty acid as micrograms of fatty acid per milligram of tissue.
321
322 *c18_2w6*
323 Concentration of 18:2 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
324
325 *c18_3w6*
326 Concentration of 18:3 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
327
328 *c18_3w3*
329 Concentration of 18:3 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
330
331 *c18_4w4*
332 Concentration of 18:4 ω 4 fatty acid as micrograms of fatty acid per milligram of tissue.
333
334 *c18_4w3*
335 Concentration of 18:4 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
336
337 *c18_5w3*
338 Concentration of 18:5 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
339
340 *c20_0*
341 Concentration of 20:0 fatty acid as micrograms of fatty acid per milligram of tissue.
342
343 *c20_1w9*
344 Concentration of 20:1 ω 9 fatty acid as micrograms of fatty acid per milligram of tissue.
345
346 *c20_1w7*
347 Concentration of 20:1 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
348
349 *c20_2w5_11*
350 Concentration of 20:2-5-11 fatty acid as micrograms of fatty acid per milligram of tissue.
351
352 *c20_2w5_13*
353 Concentration of 20:2-5-13 fatty acid as micrograms of fatty acid per milligram of tissue.
354
355 *c20_2w6*
356 Concentration of 20:2 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
357
358 *c20_3w6*
359 Concentration of 20:3 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
360
361 *c20_4w6*
362 Concentration of 20:4 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
363
364 *c20_3w3*
365 Concentration of 20:3 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.

366
367 *c20_4w3*
368 Concentration of 20:4 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
369
370 *c20_5w3*
371 Concentration of 20:5 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
372
373 *c22_0*
374 Concentration of 22:0 fatty acid as micrograms of fatty acid per milligram of tissue.
375
376 *c22_1w9*
377 Concentration of 22:1 ω 9 fatty acid as micrograms of fatty acid per milligram of tissue.
378
379 *c22_1w7*
380 Concentration of 22:1 ω 7 fatty acid as micrograms of fatty acid per milligram of tissue.
381
382 *c22_2w6*
383 Concentration of 22:2 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
384
385 *c22_4w6*
386 Concentration of 22:4 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
387
388 *c22_5w6*
389 Concentration of 22:5 ω 6 fatty acid as micrograms of fatty acid per milligram of tissue.
390
391 *c22_3w3*
392 Concentration of 22:3 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
393
394 *c22_4w3*
395 Concentration of 22:4 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
396
397 *c22_5w3*
398 Concentration of 22:5 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
399
400 *c22_6w3*
401 Concentration of 22:6 ω 3 fatty acid as micrograms of fatty acid per milligram of tissue.
402
403 *c24_0*
404 Concentration of 24:0 fatty acid as micrograms of fatty acid per milligram of tissue.
405
406 *comments*
407 Quality flag column. Two samples spilled during fatty acid extraction. These samples are flagged as
408 such. Although concentrations are lower than other samples, proportions between fatty acids are
409 consistent.
410
411 invertebrates.csv

412
413 This file contains abundance for benthic macroinvertebrates collected at each of the 14 littoral
414 sampling locations. Only amphipod taxa were identified to species. Mollusks and isopods were
415 identified to genus.
416
417 *site*
418 Unique alphanumeric identifier for a sampling location.
419
420 *replicate*
421 Replicate for sampling location. While three replicates were collected in the field, some samples
422 were poorly preserved, and invertebrates were not enumerated so as to prevent potential errors.
423
424 *Acroloxidae*
425 Mollusk genus
426
427 *Asellidae*
428 Endemic isopod genus
429
430 *Baicaliidae*
431 Mollusk genus, most of which are endemic
432
433 *Benedictidae*
434 Mollusk genus, most of which are endemic
435
436 *Brandtia_latissima*
437 Endemic amphipod species
438
439 *Brandtia_parasitica_parasitica*
440 Endemic amphipod species
441
442 *Caddisflies*
443 General grouping; were not identified to species.
444
445 *Cryptoropus_inflatus*
446 Endemic amphipod species
447
448 *Cryptoropus_pachytus*
449 Endemic amphipod species
450
451 *Cryptoropus_rugosus*
452 Endemic amphipod species
453
454 *Eulimnogammarus_capreolus*
455 Endemic amphipod species
456
457 *Eulimnogammarus_cruentes*

- 458 Endemic amphipod species
459
460 *Eulimnogammarus_cyaneus*
461 Endemic amphipod species
462
463 *Eulimnogammarus_grandimanus*
464 Endemic amphipod species
465
466 *Eulimnogammarus_juveniles*
467 Endemic amphipod genus. Identification kept at genus level so as to prevent misclassification.
468
469 *Eulimnogammarus_maackii*
470 Endemic amphipod species
471
472 *Eulimnogammarus_maritui*
473 Endemic amphipod species
474
475 *Eulimnogammarus_verucossus*
476 Endemic amphipod species
477
478 *Eulimnogammarus_viridis_viridis*
479 Endemic amphipod species
480
481 *Eulimnogammarus_vittatus*
482 Endemic amphipod species
483
484 *Flatworms*
485 Not identified beyond order.
486
487 *Leeches*
488 Not identified beyond order, although 12 endemic species do exist.
489
490 *Maackia*
491 Mollusk genus, most of which are endemic
492
493 *Pallasea_brandtia_brandtia*
494 Endemic amphipod species
495
496 *Pallasea_brandtii_tenera*
497 Endemic amphipod species
498
499 *Pallasea_cancelloides*
500 Endemic amphipod species
501
502 *Pallasea_cancellus*
503 Endemic amphipod species

504
505 *Pallasea_viridis*
506 Endemic amphipod species
507
508 *Planorbidae*
509 Mollusk genus, most of which are endemic
510
511 *Poekilogammarus_crassimus*
512 Endemic amphipod species
513
514 *Poekilogammarus_ephippiatus*
515 Endemic amphipod species
516
517 *Poekilogammarus_juveniles*
518 Endemic amphipod genus. Identifying to species introduced risk of misclassification.
519
520 *Poekilogammarus_megonychus_perpolitus*
521 Endemic amphipod species
522
523 *Poekilogammarus_pictus*
524 Endemic amphipod species
525
526 *Valvatidae*
527 Mollusk genus, most of which are endemic
528
529 metadata.csv
530
531 This file contains metadata for each of the pelagic and littoral sampling locations. Missing data are
532 assigned as NA.
533
534 *year*
535 Year sampling occurred.
536
537 *month*
538 Month sampling occurred.
539
540 *day*
541 Day sampling occurred.
542
543 *time*
544 Time sampling occurred as Hours:Minutes.
545
546 *site*
547 Unique alphanumeric identifier for a sampling location.
548
549 *lat*

550 Latitude of sampling location in decimal degrees.
551
552 *long*
553 Longitude of sampling location in decimal degrees.
554
555 *site_description*
556 Researchers' description of sampling location at the time of sampling.
557
558 *distance_to_shore_m*
559 Distance from *in situ* sampled location to the shoreline in meters.
560
561 *depth_m*
562 Depth at *in situ* sampling location in meters.
563
564 *air_temp_celsius*
565 Temperature of air at sampling location in Celsius.
566
567 *surface_temp_celsius*
568 Temperature of water's surface at sampling location in Celsius.
569
570 *mid_temp_celsius*
571 Temperature of water midway between surface and bottom at sampling location in Celsius.
572
573 *bottom_temp_celsius*
574 Temperature of water near sediment at sampling location in Celsius.
575
576 *comments*
577 Notes in the field describing sampling conditions.
578
579 *shore_photo*
580 Whether or not photos of the shoreline were taken. Photos are available on the project's Open
581 Science Framework page (Meyer et al. 2015).
582
583 *substrate_photo*
584 Whether or not photos of the substrate were taken.
585
586 *sponges*
587 Whether or not sponges were present at a sampling location.
588
589 *brandtia*
590 Whether or not *Brandtia spp.* (endemic amphipod) was present at a sampling location.
591
592 *microplastics.csv*
593
594 This file contains microplastics counts for each of the pelagic and littoral sampling locations.
595

596 *site*
597 Unique alphanumeric identifier for a sampling location.
598
599 *replicate*
600 Replicate for a given sampling location. Replicate values of "C" indicate a control.
601
602 *fragments*
603 Number of microplastic fragments observed.
604
605 *fibers*
606 Number of microplastic fibers observed.
607
608 *beads*
609 Number of microplastic beads observed.
610
611 *comments*
612 Observer comments while enumerating microplastics
613
614 *volume_filtered_ml*
615 Volume in milliliters for a given replicate filtered.
616
617 nutrients.csv
618
619 This file contains nutrient concentrations for each of the associated sampling locations. Nutrient samples were not filtered prior to analysis, meaning that nitrogen concentrations have the potential to include intracellular nitrogen. Therefore, nitrogenous species' concentrations may be spurious.
620
621
622 *site*
623 Unique alphanumeric identifier for a sampling location.
624
625 *replicate*
626 Replicate for a given sampling location.
627
628
629 *nh4_mg_dm3*
630 Ammonium concentration in milligrams of ammonium per cubic decimeter.
631
632 *no3_mg_dm3*
633 Nitrate concentration in milligrams of nitrate per cubic decimeter
634
635 *tp_mg_dm3*
636 Total phosphorus concentration in milligrams of phosphorus per cubic decimeter.
637
638 *tpo43_mg_dm3*
639 Total phosphate concentration as phosphate in milligrams per cubic decimeter.
640
641 periphyton.csv

642
643 This file contains periphyton abundance data for each of the sampled littoral locations. For poorly
644 preserved samples, counts are listed as NA for each taxonomic grouping, and a note in the
645 "comments" column is provided.

646

647 *site*

648 Unique alphanumeric identifier for a sampling location.

649

650 *replicate*

651 Replicate number for a given sampling site.

652

653 *subsamples_counted*

654 Number of 10 microliter subsamples counted for a given replicate.

655

656 *diatom*

657 Number of diatom cells counted for a given replicate.

658

659 *spirogyra*

660 Number of *Spirogyra spp.* cells counted for a given replicate.

661

662 *spirogyra_filament*

663 Number of *Spirogyra spp.* filaments counted for a given replicate.

664

665 *ulothrix*

666 Number of *Ulothrix spp.* cells counted for a given replicate.

667

668 *ulothrix_filament*

669 Number of *Ulothrix spp.* filaments counted for a given replicate.

670

671 *tetrasporales*

672 Number of *Tetrasporales spp.* cells counted for a given replicate

673

674 *pediastrum*

675 Number of *Pediastrum spp.* cells counted for a given replicate.

676

677 *desmidales*

678 Number of *Desmidales spp.* cells counted for a given replicate.

679

680 *comments*

681 Notes from the observer.

682

683 ppcp.csv

684

685 This file contains Pharmaceutical and Personal Care Product (PPCP) concentrations for each littoral
686 and pelagic sampling location. Detection limits are estimated to be 0.001 µg/L based on a 500 mL
687 sample volume.

- 688
689 *site*
690 Unique alphanumeric identifier for a sampling location.
691
692 *paraxanthine*
693 Concentration of paraxanthine, also known as 1,7-dimethylxanthine, in micrograms per liter.
694 Paraxanthine is the main human metabolite of caffeine
695
696 *acetaminophen*
697 Concentration of acetaminophen, also known as paracetamol, in micrograms per liter.
698
699 *amphetamine*
700 Concentration of amphetamine in micrograms per liter.
701
702 *caffeine*
703 Concentration of caffeine in micrograms per liter.
704
705 *carbamazepine*
706 Concentration of carbamazepine in micrograms per liter.
707
708 *cimetidine*
709 Concentration of cimetidine in micrograms per liter.
710
711 *cotinine*
712 Concentration of cotinine, which is the main human metabolite of nicotine, in micrograms per liter.
713
714 *diphenhydramine*
715 Concentration of diphenhydramine in micrograms per liter.
716
717 *mda*
718 Concentration of methylenedioxymethamphetamine in micrograms per liter.
719
720 *mdma*
721 Concentration of methylenedioxymethamphetamine in micrograms per liter.
722
723 *methamphetamine*
724 Concentration of methamphetamine in micrograms per liter.
725
726 *morphine*
727 Concentration of morphine in micrograms per liter.
728
729 *phenazone*
730 Concentration of phenazone in micrograms per liter.
731
732 *sulfachloropyridazine*
733 Concentration of sulfachloropyridazine in micrograms per liter.

734
735 *sulfamethazine*
736 Concentration of *sulfamethazine* in micrograms per liter.
737
738 *sulfamethoxazole*
739 Concentration of sulfamethoxazole in micrograms per liter.
740
741 *thiabendazole*
742 Concentration of thiabendazole in micrograms per liter.
743
744 *trimethoprim*
745 Concentration of trimethoprim in micrograms per liter.
746
747 *collection_year*
748 Year sample was collected in the field.
749
750 *collection_month*
751 Month sample was collected in the field.
752
753 *collection_day*
754 Day sample was collected in the field.
755
756 *analysis_year*
757 Year sample was analyzed.
758
759 *analysis_month*
760 Month sample was analyzed.
761
762 *analysis_day*
763 Day sample was analyzed.
764
765 stable_isotopes.csv
766
767 This file contains carbon ($\delta^{13}\text{C}$) and nitrogen ($\delta^{15}\text{N}$) values for various benthic macroinvertebrate
768 genera and periphyton collected from the 14 littoral sampling locations.
769
770 *C13*
771 Carbon ($\delta^{13}\text{C}$) stable isotope values in parts per thousand.
772
773 *N15*
774 Nitrogen ($\delta^{15}\text{N}$) stable isotope values in parts per thousand.
775
776 *site*
777 Unique alphanumeric identifier for a sampling location.
778
779 *Genus*

780 Genus of the analyzed organism.
781
782 *Species*
783 Species of the analyzed organism. When organism was identified solely to genus, the Species value
784 is NA.
785
786 *comments*
787 Quality flag column where $\delta^{13}\text{C}$ samples were outside of the range of standards.
788
789 total_lipid.csv
790
791 This file contains gravimetry data for each fatty acid sample.
792
793 *site*
794 Unique alphanumeric identifier for a sampling location.
795
796 *Genus*
797 Genus of the analyzed organism.
798
799 *Species*
800 Species of the analyzed organism. When organism was identified solely to genus, the Species value
801 is NA.
802
803 total_lipid_mg_per_g
804 Total amount of lipids in a sample in milligrams of lipid per gram of tissue.
805
806 *deviation*
807 1. Samples were weighed three times and standard deviation in measurement was calculated.
808 All values are reported in milligrams of lipid per gram of tissue.
809
810 *comments*
811 Quality flag column. Two samples spilled during fatty acid extraction. These samples are flagged as
812 such.
813
814 **Methods**
815
816 *Inverse distance weighted (IDW) population calculation for each sampling location*
817
818 We recognized that sewage indicator concentrations at each sampling location may be related to a
819 sampling location's spatial position relative to both the size and proximity of neighboring
820 developed sites. Therefore, we created the inverse distance weighted (IDW) population metric to
821 compress, into a single metric, information about human population size, density, and location
822 along the shoreline as well as distance between developed sites and sampling locations.
823
824 Our workflow for calculating IDW population required five steps. First, we traced polygons of each
825 lakeside development's perimeter and line geometries of each development's shorelines from

826 satellite imagery for each developed site in Google Earth. Polygons were traced for the entire area
 827 of visible development. Similarly, shoreline traces only reflected shoreline length for which there
 828 was visible development. Second, polygon and line geometries were downloaded from Google
 829 Earth as a .kml file. Third, the .kml file was imported into the R statistical environment (R Core
 830 Team 2019), where using the sf package (Pebesma 2018) we calculated shoreline length, polygon
 831 area, and centroid location for each developed site. Fourth, we joined point locations of each
 832 sampling site with the spatial polygons to calculate the distance from each sampling location to
 833 each developed site's centroid. Fifth, we calculated IDW population for each sampling location,
 834 using formula (1).

$$835 \quad (1) \quad I_j = \frac{\frac{P_{LI} * L_{LI}}{A_{LI}}}{D_{j,LI}} + \frac{\frac{P_{BK} * L_{BK}}{A_{BK}}}{D_{j,BK}} + \frac{\frac{P_{BGO} * L_{BGO}}{A_{BGO}}}{D_{j,BGO}}$$

836 where I is the IDW population at sampling location j , P is the population at each of the three
 837 developed sites Listvyanka (LI), Bolshie Koty (BK), Bolshoe Goloustnoe (BGO), A is the area of a
 838 developed site in km^2 , L is the shoreline length at a developed site in km, and D is the distance from
 839 developed site j to each developed site's centroid in km.

840
 841 *Nutrients*
 842
 843 Water samples for nutrient analyses were collected in 150 mL glass jars that had been washed with
 844 phosphate-free soap and rinsed three times with water from the sampling location. Samples were
 845 collected in duplicates and immediately frozen at -20°C until processing at the A.P.Vinogradov
 846 Institute of Geochemistry (Siberian Branch of the Russian Academy of Sciences, Irkutsk). Samples
 847 were not filtered prior to freezing, meaning that nitrogen and ammonium concentrations may
 848 potentially include intracellular nitrogen and overestimate dissolved nitrogenous forms in the water
 849 column.

850
 851 For each water sample, nitrate, ammonium, and total phosphorus concentrations were measured.
 852 For ammonium (2016a) and nitrate (2017) concentrations, samples were analyzed with a
 853 spectrophotometer following the addition of Nessler's reagent and disulfuric acid respectively.
 854 Total phosphorus concentration was measured with a spectrophotometer following the addition of
 855 persulfate (2016b).

856
 857 *Chlorophyll a*
 858
 859 Water samples were collected in 1.5 L plastic bottles from a depth of approximately 0.75 m. Within
 860 12 h of collection, three subsamples (up to 150 mL each) were filtered through 25-mm diameter,
 861 0.2 µm pore size nitrocellulose filters. Filters were then placed in a 35 mm petri dish and frozen in
 862 the dark until processing.

863
 864 Chlorophyll samples were processed in a manner similar to that of Parsons and Strickland (1963)
 865 and Lorenzen (1967). Nitrocellulose filters were ground in 90% acetone, in which chlorophyll
 866 extraction was allowed to proceed overnight. Samples were then centrifuged for 15-20 minutes.
 867 After centrifugation, absorbance of the chlorophyll extract was measured in a spectrophotometer at
 868 630, 645, 665, and 750 nm. Concentrations were calculated using the formula: $C = 11.64(A_{665} -$
 869 $A_{750}) - 2.16(A_{645} - A_{750}) - 0.1(A_{630} - A_{750}) / (V_2/V_1)$; where A is the absorbance value of a

870 particular wavelength, V_1 is the volume of the filtered water, and V_2 is the volume of extract.
871 Concentrations are reported as mg/L.

872
873 *Pharmaceuticals and Personal Care Products (PPCPs)*
874

875 Water samples for PPCP analysis were collected in 250 mL amber glass bottles that were rinsed
876 with either methanol or acetone and then three times with sample water prior to collections.
877 Following collection, samples were refrigerated and kept in the dark until solid phase extraction
878 (SPE).

879 Within 12 h of collection, samples were filtered directly from the amber glass bottle using an in-line
880 Teflon filter holder with glass microfiber GMF (1.0 μm pore size, WhatmanGrad 934-AH) in
881 tandem with a solid phase extraction (SPE) cartridge (200 mg HLB, Waters Corporation, Milford,
882 MA) connected to a 1-liter vacuum flask. Lab personnel wore gloves and face masks to minimize
883 contamination. Prior to filtration, SPE cartridges were primed with at least 5 mL of either methanol
884 or acetone and then washed with at least 5 mL of sample water. Rate of extraction was maintained
885 at approximately 1 drop per second. Extraction proceeded until water could no longer pass through
886 the SPE cartridge or until all collected water was filtered. Cartridges were stored in Whirlpacks at -
887 20°C until analysis for 18 PPCP residues using liquid chromatography tandem mass spectrometry
888 (LC-MS-MS) following methods of Lee et al. (2016) and D'Alessio et al (2018). Concentrations are
889 reported in $\mu\text{g}/\text{L}$.
890

891
892 *Microplastics*
893

894 At each location, samples were collected in triplicate using 1.5 L clear plastic bottles that were
895 washed thoroughly with sample water before each collection. Samples were collected by hand for
896 each littoral site and with a metal bucket from aboard the ship for pelagic sites.
897

898 For processing, each sample was vacuum filtered on to a 47-mm diameter GF/F filter. During
899 filtration, aluminum foil was used to cover the filtration funnel to prevent contamination from
900 airborne microplastic particles. After filtration, filters were dried under vacuum pressure and then
901 stored in 50-mm petri dishes. Following filtration of all three replicates, the filtrate was collected
902 and then re-filtered through a GF/F filter as a control for contamination from the plastic vacuum
903 funnel or potentially airborne microplastics.
904

905 Microplastic counting involved visual inspection of the entire GF/F in a similar manner to methods
906 described in Hanvey et al. (2017). Visual enumeration was conducted under a stereo microscope
907 with $\sim 100x$ magnification, and microplastics were classified into one of three categories: fibers,
908 fragments, or beads. For all categories, plastics were defined as observed objects with apparent
909 artificial colors, so as to not enumerate plastics potentially contributed from the sampling bottle
910 itself. Fibers were defined as smooth, long plastics with consistent diameters. Fragments were
911 defined as plastics with irregularly sharp or jagged edges. Beads were defined as spherical plastics.
912 Although we did not measure microplastic size, this technique likely allowed us to reliably quantify
913 microplastics as small as $\sim 300 \mu\text{m}$ (Hanvey et al. 2017). During enumeration, GF/Fs remained
914 covered in the petri dish to minimize potential for contamination from the air.
915

916 It is worth noting that since the time of our field sampling, evidence has accumulated that our
917 methods likely dramatically underestimated microplastic abundance (Wang and Wang 2018;
918 Brandon et al. 2020). Recent investigations of microplastics in Lake Baikal near Bolshie Koty (BK)
919 used analogous methods and measured similar microplastic concentrations (Karnaukhov et al.
920 2020). Future studies aiming to use these data for comparison or supplementing potential data gaps
921 should consider the minimum microplastic size that could be reliably detected by the method, so as
922 to ensure data are comparable across methods.

923

924 *Periphyton abundance*

925

926 At each littoral site, we haphazardly selected three rocks representative of local substrate. A plastic
927 stencil was used to define a surface area of each rock from which we scraped a standardized 14.5
928 cm² patch of periphyton. Samples were preserved with Lugol's solution and stored in plastic
929 scintillation vials. Additional periphyton was collected in composite from each site for fatty acid
930 and stable isotope analysis.

931

932 Periphyton taxonomic identification and enumeration was performed by subsampling 10 µL
933 aliquots from each preserved sample. For all 10 µL aliquots, cells, filaments, and colonies were
934 counted, for the entire subsample, until at least 300 cells were identified for a given sampling
935 replicate. If the first aliquot contained less than 300 cells, we counted additional subsamples until
936 we reached at least 300 cells in total. In instances when 300 cells were counted before finishing a
937 subsample, we still counted the entire aliquot. Taxa were classified into broad categories consistent
938 with Baikal algal taxonomy (Izhboldina 2007), using coarse groupings to capture general patterns in
939 relative algal abundance. As a result, algal groups consisted of diatoms, *Ulothrix*, *Spirogyra*, and
940 the green algal Order Tetrasporales.

941

942 *Benthic macroinvertebrate abundance*

943

944 Three kick-net samples were collected for assessment of benthic community composition and
945 abundance. Using a D-net, we collected macroinvertebrates by flipping over 1-3 rocks, and then
946 sweeping five times in a left-to-right motion across approximately 1 m. After the series of sweeps,
947 the catch was rinsed into a plastic bucket. For each replicate, bucket contents were concentrated
948 using a 64-µm mesh and placed in glass jars with 40% ethanol (vodka; the only preservative
949 available to us at the time) for preservation and refrigerated at 4°C aboard the research vessel. The
950 40% ethanol preservative was replaced with ~80% ethanol upon return to the lab within 24 to 48
951 hours, and samples were stored at ~4°C.

952

953 Invertebrate taxonomic identification and enumeration were performed under a stereo microscope.
954 All invertebrates were identified to species with the exception of juveniles (Takhteev and
955 Didorenko (2015) for amphipods; Sitnikova (2012) for mollusks; Table 2). All samples contained
956 oligochaetes and polychaetes, but due to poor preservation, these taxa were not counted. Six
957 samples of the 42 collected were not well-preserved and were excluded from further analyses, in
958 order to reduce errors in identification. KD-1 and LI-1 were the only sites with 1 sample counted.
959 BK-2 and KD-2 each had two samples counted.

960

961 *Stable Isotope Analysis*

962
963 Measurements of $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ were performed on an elemental analyzer-isotope ratio mass
964 spectrometer (EA-IRMS; Finnigan DELTAplus XP, Thermo Scientific) at the Large Lakes
965 Observatory, University of Minnesota Duluth.
966
967 *Fatty Acid Analysis*
968
969 Following freeze-drying, samples were transferred to 10 mL glass centrifuge vials, and 2 mL of
970 100% chloroform was added to each under nitrogen gas. Samples were allowed to sit in chloroform
971 overnight at -80°C. Fatty acid extractions generally involved three phases: (1) 100% chloroform
972 extraction, (2) chloroform-methanol extraction, and (3) fatty acid methylation. Fatty acid extraction
973 methods were adapted from Schram et al. (2018).
974
975 After overnight chloroform extraction, samples underwent a chloroform-methanol extraction three
976 times. To each sample, we added 1 mL cooled 100% methanol, 1 mL chloroform:methanol solution
977 (2:1), and 0.8 mL 0.9% NaCl solution. Samples were inverted three times and sonicated on ice for
978 10 minutes. Next, samples were vortexed for 1 minute, and centrifuged for 5 minutes (3,000 rpm) at
979 4°C. Using a double pipette technique, the lower organic layer was removed and kept under
980 nitrogen. After the third extraction, samples were evaporated under nitrogen flow, and resuspended
981 in 1.5 mL chloroform and stored at -20°C overnight.
982
983 Once resuspended in chloroform, 1 mL of chloroform extract was transferred to a glass centrifuge
984 tube with a glass syringe as well as an internal standard of 4 μL of 19-carbon fatty acid. Samples
985 were then evaporated under nitrogen, and then 1 mL of toluene and 2 mL of 1% sulfuric acid-
986 methanol was added. The vial was closed under nitrogen gas and then incubated in 50°C water bath
987 for 16 hours. After incubation, samples were removed from the bath, allowed to reach room
988 temperature and stored on ice. Next, we performed a potassium carbonate-hexane extraction twice.
989 To each sample, we added 2 mL of 2% potassium bicarbonate and 5 mL of 100% hexane, inverting
990 the capped vial so as to mix the solution. Samples were centrifuged for 3 minutes (1,500 rpm) at
991 4°C. The upper hexane layer was then removed and placed in a vial to evaporate under nitrogen
992 flow. Once almost evaporated, 1 mL of 100% hexane was added and stored in a glass amber
993 autosampler vial for GC/MS quantification. GC/MS quantification was performed with a Shimadzu
994 QP2020 GC/MS following Schram et al. (2018).
995
996 **Technical Validation**
997
998 The dataset had three main validation procedures: taxonomic, analytical, and reproducible.
999
1000 For taxonomic validation, all phylogenetic groupings were based off most recent identification
1001 keys. Amphipods were identified according to Takhteev & Didorenko (2015). Mollusks were
1002 identified according to Sitnikova (2012). Algal taxa were identified according to Izhboldina (2007).
1003 For consistency, all taxa were identified by one person (Michael F. Meyer), who was trained by
1004 experts in Baikal algal and macroinvertebrate taxonomy.
1005
1006 For analytical validation, internal standards were used for all mass-spectroscopy analyses. PPCP
1007 analyses involved labeled internal standards ($^{13}\text{C}_3$ -caffeine, methamphetamine-d8, MDMA-d8,

1008 morphine-d3, and $^{13}\text{C}_6$ -sulfamethazine). Stable isotope values were calibrated against certified
1009 reference materials including L-glutamic acid (NIST SRM 8574), low organic soil and sorghum
1010 flour (standards B-2153 and B-2159 from Elemental Micro-analysis Ltd., Okehampton, UK) and
1011 in-house standards (acetanilide and caffeine). Replicate analyses of external standards showed a
1012 mean standard deviation of 0.06 ‰ and 0.09 ‰, for $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$, respectively. Finally, fatty acid
1013 estimations used an internal 19:0 standard to assess oxidation of fatty acids during extraction,
1014 methylation, and quantification.

1015
1016 For data reproducibility, data aggregation and harmonization procedures were conducted in the R
1017 statistical environment (R Core Team 2019), using the tidyverse (Wickham et al. 2019) packages.
1018 As part of the data aggregation, an initial cleaning script (00_disaggregated_data_cleaning.R)
1019 removed incorrect spellings, erroneous data values, and inconsistent column names from raw data.
1020 This step created the standardized CSV files detailed above, which are available on the EDI
1021 repository (Meyer et al. 2020b). Raw data files are available on the project's Open Science
1022 Framework portal (Meyer et al. 2015) but are not included in the EDI repository to prevent
1023 confusion or incorrect usage. Data hosted on EDI are at the replicate-level but can be aggregated to
1024 the sampling-site-level using script "01_data_cleaning.R". In addition to aggregation scripts, six R
1025 scripts used for analyses in Meyer et al. (*Under Review*) are also available on the EDI repository
1026 within the compressed entity "scripts.tar.gz". All R code for data aggregation was written by one
1027 person (Michael F. Meyer) and then independently reviewed by two others (Matthew R. Brousil
1028 and Kara H. Woo) to confirm that code performed as intended, was well documented, and
1029 annotations were complete.

1030
1031 *A commitment to FAIR and TRUST principles*
1032

1033 Throughout the dataset's development, we strove to incorporate both FAIR (Findable, Accessible,
1034 Interoperable, and Reproducible) and TRUST (Transparency, Responsibility, User Focus,
1035 Sustainability, and Technology) principles where applicable.

1036 With respect to FAIR principles (Wilkinson et al. 2016), the data are openly accessible in a
1037 standardized, replicate-level format on the EDI portal. The 10 CSV files contained within the
1038 dataset are entirely interoperable using the "site" column, enabling all variables to efficiently be
1039 merged together. Finally, all analytical and some data wrangling scripts are available on the EDI
1040 portal in a compressed format, such that future users can reproduce data manipulation and analyses
1041 described in Meyer et al. (*Under Review*).
1042

1043 With respect to TRUST principles (Lin et al. 2020), we strove to document additional metadata and
1044 data-cleaning practices in a public Open Science Framework (OSF) repository (Meyer et al. 2015).
1045 These steps are not necessarily critical to the core EDI dataset, but provide increased transparency
1046 for future users wishing recreate the dataset de novo. All "raw" data are provided in the OSF portal,
1047 including an initial cleaning script (00_disaggregated_data_cleaning.R) to remove incorrect
1048 spellings, erroneous data values, and inconsistent column names. This repository also includes
1049 photographs of both field notes as well as photographs of shoreline and substrate from sampling
1050 locations. To empower and expedite future reuse, all directories are accompanied with
1051 documentation that details directory contents, and all associated scripts are documented and
1052 annotated. While many of the files are redundant from the EDI repository, the OSF repository is
1053

1054 meant to supplement the EDI repository, so as to enable sustainable, user-focused transparency of
1055 how data were collected and cleaned from their raw formats.

1056
1057 **Data Use and Recommendations for Reuse**
1058

1059 Recognizing the potential for continued low-level, sewage pollution at Lake Baikal (Timoshkin et
1060 al. 2016, 2018; Volkova et al. 2018) and lakes worldwide (Yang et al. 2018; Meyer et al. 2019), the
1061 final dataset can be applied to a suite of research questions pertaining to ecological responses to
1062 human disturbance. We highlight two main areas for immediate application.

1063
1064 First, the final data products can be harmonized with other littoral sampling efforts throughout Lake
1065 Baikal, so as to enhance spatial coverage and data diversity. Since 2010, Lake Baikal has
1066 experienced increasing filamentous algal abundance, especially near larger lakeside developments
1067 (Kravtsova et al. 2014; Timoshkin et al. 2016, 2018; Volkova et al. 2018). Recent benthic algal
1068 surveys throughout Lake Baikal's entirety have suggested that cosmopolitan filamentous algae,
1069 such as *Spirogyra spp.*, tend to be more abundant near larger lakeside developments (Timoshkin et
1070 al. 2016; Volkova et al. 2018). For example, Listvyanka is a small town located at the beginning of
1071 the Angara River, Lake Baikal's only surface outflow. While Listvyanka's permanent population is
1072 approximately 2,000 persons, the town is a growing tourism hub, and hosts over 1.2 million tourists
1073 per year (Interfax-Tourism 2018). Surveys conducted near Listvyanka have suggested increased
1074 *Spirogyra spp.* abundance is associated with wastewater release (Timoshkin et al. 2016). Although
1075 wastewater inputs are likely low and are diluted to negligible concentrations offshore (Meyer et al.,
1076 Under Review), combining monitoring efforts across spatial and temporal scales are necessary to
1077 evaluate the spatial and temporal extent of wastewater entering Lake Baikal. As such, our data
1078 could complement previous, current, and future monitoring efforts, where observations may be
1079 missing.

1080
1081 Second, the final data products are useful to expanding freshwater PPCP, microplastic, and
1082 associated biological responses across large spatial scales. Recent syntheses of the PPCP literature
1083 have reported that studies involving lakes are less abundant relative to those focused on lotic
1084 systems (Meyer et al. 2019). Likewise, microplastic studies have noted that freshwater
1085 environments are less represented in the literature relative to marine ecosystems (Horton et al.
1086 2017). For both PPCPs and microplastics, toxic responses to even minute concentrations can be
1087 uncertain and differ between ecosystem types (e.g., Rosi-Marshall et al. 2013 for lotic and Shaw et
1088 al. 2015 for lentic). As a result of PPCPs and microplastics garnering increasing attention
1089 worldwide, sampling of PPCPs and microplastics with co-located biological data across multiple
1090 spatial and temporal scales would be necessary to synthesize biotic responses to micropollutants
1091 across systems. Although our data constitute a limited sample number of PPCP and microplastic
1092 data that exist globally, our final data products are highly structured and flexible for merging with
1093 similar datasets. Additionally, our dataset's sequential harmonization workflow could be adopted
1094 by similar monitoring efforts, thereby facilitating data interoperability. Through integration with
1095 similar monitoring efforts, our dataset can contribute to global synthesis of emerging contaminant
1096 consequences, especially in a region of the world that is often not easily accessible to many
1097 researchers.

1098
1099

1100 **References**

- 1101
1102 Barnes, D. K. A., F. Galgani, R. C. Thompson, and M. Barlaz. 2009. Accumulation and
1103 fragmentation of plastic debris in global environments. *Philos Trans R Soc Lond B Biol Sci*
1104 **364**: 1985–1998. doi:10.1098/rstb.2008.0205
- 1105 Bendz, D., N. A. Paxéus, T. R. Ginn, and F. J. Loge. 2005. Occurrence and fate of pharmaceutically
1106 active compounds in the environment, a case study: Höje River in Sweden. *Journal of*
1107 *Hazardous Materials* **122**: 195–204. doi:10.1016/j.jhazmat.2005.03.012
- 1108 Brandon, J. A., A. Freibott, and L. M. Sala. 2020. Patterns of suspended and salp-ingested
1109 microplastic debris in the North Pacific investigated with epifluorescence microscopy.
1110 *Limnology and Oceanography Letters* **5**: 46–53. doi:10.1002/lol2.10127
- 1111 Brodin, T., J. Fick, M. Jonsson, and J. Klaminder. 2013. Dilute Concentrations of a Psychiatric
1112 Drug Alter Behavior of Fish from Natural Populations. *Science* **339**: 814–815.
1113 doi:10.1126/science.1226850
- 1114 Camilleri, A. C., and T. Ozersky. 2019. Large variation in periphyton $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in the
1115 upper Great Lakes: Correlates and implications. *Journal of Great Lakes Research* **45**: 986–
1116 990. doi:10.1016/j.jglr.2019.06.003
- 1117 Costanzo, S. D., M. J. O'Donohue, W. C. Dennison, N. R. Loneragan, and M. Thomas. 2001. A
1118 New Approach for Detecting and Mapping Sewage Impacts. *Marine Pollution Bulletin* **42**:
1119 149–156. doi:10.1016/S0025-326X(00)00125-9
- 1120 D'Alessio, M., S. Onanong, D. D. Snow, and C. Ray. 2018. Occurrence and removal of
1121 pharmaceutical compounds and steroids at four wastewater treatment plants in Hawai'i and
1122 their environmental fate. *Science of The Total Environment* **631–632**: 1360–1370.
1123 doi:10.1016/j.scitotenv.2018.03.100

- 1124 Dalsgaard, J., M. St. John, G. Kattner, D. Müller-Navarra, and W. Hagen. 2003. Fatty acid trophic
1125 markers in the pelagic marine environment, p. 225–340. In *Advances in Marine Biology*.
1126 Elsevier.
- 1127 Edmondson, W. T. 1970. Phosphorus, Nitrogen, and Algae in Lake Washington after Diversion of
1128 Sewage. *Science* **169**: 690–691.
- 1129 Fellows, I., and using the Jm. library by J. P. Stotz. 2019. OpenStreetMap: Access to Open Street
1130 Map Raster Images.,
- 1131 Focazio, M. J., D. W. Kolpin, K. K. Barnes, E. T. Furlong, M. T. Meyer, S. D. Zaugg, L. B. Barber,
1132 and M. E. Thurman. 2008. A national reconnaissance for pharmaceuticals and other organic
1133 wastewater contaminants in the United States - II) Untreated drinking water sources.
1134 *SCIENCE OF THE TOTAL ENVIRONMENT* **402**: 201–216.
1135 doi:10.1016/j.scitotenv.2008.02.021
- 1136 Gartner, A., P. Lavery, and A. J. Smit. 2002. Use of delta N-15 signatures of different functional
1137 forms of macroalgae and filter-feeders to reveal temporal and spatial patterns in sewage
1138 dispersal. *Mar. Ecol.-Prog. Ser.* **235**: 63–73. doi:10.3354/meps235063
- 1139 Green, D. S. 2016. Effects of microplastics on European flat oysters, *Ostrea edulis* and their
1140 associated benthic communities. *Environmental Pollution* **216**: 95–103.
1141 doi:10.1016/j.envpol.2016.05.043
- 1142 Hall, R. I., P. R. Leavitt, R. Quinlan, A. S. Dixit, and J. P. Smol. 1999. Effects of agriculture,
1143 urbanization, and climate on water quality in the northern Great Plains. *Limnology and*
1144 *Oceanography* **44**: 739–756. doi:10.4319/lo.1999.44.3_part_2.0739

- 1145 Hampton, S. E., S. C. Fradkin, P. R. Leavitt, and E. E. Rosenberger. 2011. Disproportionate
1146 importance of nearshore habitat for the food web of a deep oligotrophic lake. *Marine and*
1147 *Freshwater Research* **62**: 350. doi:10.1071/MF10229
- 1148 Hampton, S. E., L. R. Izmest'eva, M. V. Moore, S. L. Katz, B. Dennis, and E. A. Silow. 2008.
1149 Sixty years of environmental change in the world's largest freshwater lake - Lake Baikal,
1150 Siberia. *Global Change Biology* **14**: 1947–1958. doi:10.1111/j.1365-2486.2008.01616.x
- 1151 Hampton, S. E., S. McGowan, T. Ozersky, and others. 2018. Recent ecological change in ancient
1152 lakes. *Limnology and Oceanography* **63**: 2277–2304. doi:10.1002/lno.10938
- 1153 Hanvey, J. S., P. J. Lewis, J. L. Lavers, N. D. Crosbie, K. Pozo, and B. O. Clarke. 2017. A review
1154 of analytical techniques for quantifying microplastics in sediments. *Anal. Methods* **9**: 1369–
1155 1383. doi:10.1039/C6AY02707E
- 1156 Horton, A. A., A. Walton, D. J. Spurgeon, E. Lahive, and C. Svendsen. 2017. Microplastics in
1157 freshwater and terrestrial environments: Evaluating the current understanding to identify the
1158 knowledge gaps and future research priorities. *Science of The Total Environment* **586**: 127–
1159 141. doi:10.1016/j.scitotenv.2017.01.190
- 1160 Interfax-Tourism. 2018. Байкал с января по август 2018 года посетили 1,2 миллиона туристов
1161 (1.2 million tourists visted Baikal from January through August 2018). Interfax-Tourism,
1162 October 25
- 1163 Izhboldina, L. A. 2007. Guide and Key to Benthic and Periphyton Algae of Lake Baikal (meio- and
1164 macrophytes) with Brief Notes on Their Ecology, Nauka-Centre.
- 1165 Izmest'eva, L. R., M. V. Moore, S. E. Hampton, and others. 2016. Lake-wide physical and
1166 biological trends associated with warming in Lake Baikal. *Journal of Great Lakes Research*
1167 **42**: 6–17. doi:10.1016/j.jglr.2015.11.006

- 1168 Jeppesen, E., M. Søndergaard, J. P. Jensen, and others. 2005. Lake responses to reduced nutrient
1169 loading – an analysis of contemporary long-term data from 35 case studies. *Freshwater
1170 Biology* **50**: 1747–1771. doi:10.1111/j.1365-2427.2005.01415.x
- 1171 Karnaukhov, D., S. Biritskaya, E. Dolinskaya, M. Teplykh, N. Silenko, Y. Ermolaeva, and E.
1172 Silow. 2020. POLLUTION BY MACRO- AND MICROPLASTIC OF LARGE
1173 LACUSTRINE ECOSYSTEMS IN EASTERN ASIA. *Pollution Research* **2**: 353–355.
- 1174 Kassambara, A. 2019. *ggpubr: “ggplot2” Based Publication Ready Plots,*,
1175 Katz, S. L., L. R. Izmest’eva, S. E. Hampton, T. Ozersky, K. Shchapov, M. V. Moore, S. V.
1176 Shimaraeva, and E. A. Silow. 2015. The “Melosira years” of Lake Baikal: Winter
1177 environmental conditions at ice onset predict under-ice algal blooms in spring. *Limnology
1178 and Oceanography* **60**: 1950–1964. doi:10.1002/lo.10143
- 1179 Kolpin, D. W., E. T. Furlong, M. T. Meyer, E. M. Thurman, S. D. Zaugg, L. B. Barber, and H. T.
1180 Buxton. 2002. Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in
1181 U.S. Streams, 1999–2000: A National Reconnaissance. *Environmental Science &
1182 Technology* **36**: 1202–1211. doi:10.1021/es011055j
- 1183 Kozhova, O. M., and L. R. Izmest’eva. 1998. *Lake Baikal: Evolution and Biodiversity*, Backhuys
1184 Publishers.
- 1185 Kravtsova, L. S., L. A. Izhboldina, I. V. Khanaev, and others. 2014. Nearshore benthic blooms of
1186 filamentous green algae in Lake Baikal. *Journal of Great Lakes Research* **40**: 441–448.
1187 doi:10.1016/j.jglr.2014.02.019
- 1188 Lapointe, B. E., L. W. Herren, D. D. Debortoli, and M. A. Vogel. 2015. Evidence of sewage-driven
1189 eutrophication and harmful algal blooms in Florida’s Indian River Lagoon. *Harmful Algae*
1190 **43**: 82–102. doi:10.1016/j.hal.2015.01.004

- 1191 Lee, S. S., A. M. Paspalof, D. D. Snow, E. K. Richmond, E. J. Rosi-Marshall, and J. J. Kelly. 2016.
1192 Occurrence and Potential Biological Effects of Amphetamine on Stream Communities.
1193 Environmental Science & Technology **50**: 9727–9735. doi:10.1021/acs.est.6b03717
- 1194 Lin, D., J. Crabtree, I. Dillo, and others. 2020. The TRUST Principles for digital repositories.
1195 Scientific Data **7**: 144. doi:10.1038/s41597-020-0486-7
- 1196 Lorenzen, C. J. 1967. Determination of Chlorophyll and Pheo-Pigments: Spectrophotometric
1197 Equations1. Limnology and Oceanography **12**: 343–346.
1198 doi:<https://doi.org/10.4319/lo.1967.12.2.0343>
- 1199 Meyer, M. F., S. G. Labou, A. N. Cramer, M. R. Brousil, and B. T. Luff. 2020a. The global lake
1200 area, climate, and population dataset. Scientific Data **7**: 174. doi:10.1038/s41597-020-0517-
1201 4
- 1202 Meyer, M. F., T. Ozersky, K. H. Woo, and others. 2020b. A unified dataset of co-located sewage
1203 pollution, periphyton, and benthic macroinvertebrate community and food web structure
1204 from Lake Baikal (Siberia).doi:10.6073/PASTA/76F43144015EC795679BAC508EFA044B
- 1205 Meyer, M. F., T. Ozersky, K. H. Woo, and others. Effects of spatially heterogeneous lakeside
1206 development on nearshore biotic communities in a large, deep, oligotrophic lake (Lake
1207 Baikal, Siberia).
- 1208 Meyer, M. F., S. M. Powers, and S. E. Hampton. 2019. An Evidence Synthesis of Pharmaceuticals
1209 and Personal Care Products (PPCPs) in the Environment: Imbalances among Compounds,
1210 Sewage Treatment Techniques, and Ecosystem Types. Environ. Sci. Technol. **53**: 12961–
1211 12973. doi:10.1021/acs.est.9b02966
- 1212 Meyer, M., T. Ozersky, K. Woo, A. W. E. Galloway, M. R. Brousil, and S. Hampton. 2015. Baikal
1213 Food Webs.doi:[10.17605/OSF.IO/9TA8Z](https://doi.org/10.17605/OSF.IO/9TA8Z)

- 1214 Moore, J. W., D. E. Schindler, M. D. Scheuerell, D. Smith, and J. Frogge. 2003. Lake
1215 eutrophication at the urban fringe, Seattle region, USA. *AMBIO: A Journal of the Human*
1216 *Environment* **32**: 13–18.
- 1217 Moore, M. V., S. E. Hampton, L. R. Izmest'eva, E. A. Silow, E. V. Peshkova, and B. K. Pavlov.
1218 2009. Climate Change and the World's "Sacred Sea"-Lake Baikal, Siberia. *Bioscience* **59**:
1219 405–417. doi:10.1525/bio.2009.59.5.8
- 1220 O'Donnell, D. R., P. Wilburn, E. A. Silow, L. Y. Yampolsky, and E. Litchman. 2017. Nitrogen and
1221 phosphorus colimitation of phytoplankton in Lake Baikal: Insights from a spatial survey and
1222 nutrient enrichment experiments. *Limnology and Oceanography* **62**: 1383–1392.
1223 doi:10.1002/lno.10505
- 1224 Parsons, T. R., and J. D. H. Strickland. 1963. Discussion of spectrophotometric determination of
1225 marine-plant pigments, with revised equations for ascertaining chlorophylls and carotenoids.
1226 *Journal of Marine Research*.
- 1227 Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R*
1228 *Journal* **10**: 439–446. doi:10.32614/RJ-2018-009
- 1229 Powers, S. M., T. W. Bruulsema, T. P. Burt, and others. 2016. Long-term accumulation and
1230 transport of anthropogenic phosphorus in three river basins. *Nature Geoscience* **9**: 353–356.
1231 doi:10.1038/ngeo2693
- 1232 R Core Team. 2019. R: A Language and Environment for Statistical Computing.,
- 1233 Richmond, E. K., M. R. Grace, J. J. Kelly, A. J. Reisinger, E. J. Rosi, and D. M. Walters. 2017.
1234 Pharmaceuticals and personal care products (PPCPs) are ecological disrupting compounds
1235 (EcoDC). *Elem Sci Anth* **5**: 66. doi:10.1525/elementa.252

- 1236 Richmond, E. K., E. J. Rosi, D. M. Walters, J. Fick, S. K. Hamilton, T. Brodin, A. Sundelin, and M.
1237 R. Grace. 2018. A diverse suite of pharmaceuticals contaminates stream and riparian food
1238 webs. *Nature Communications* **9**: 4491. doi:10.1038/s41467-018-06822-w
- 1239 Romera-Castillo, C., M. Pinto, T. M. Langer, X. A. Álvarez-Salgado, and G. J. Herndl. 2018.
1240 Dissolved organic carbon leaching from plastics stimulates microbial activity in the ocean.
1241 *Nat Commun* **9**: 1–7. doi:10.1038/s41467-018-03798-5
- 1242 Rosenberger, E. E., S. E. Hampton, S. C. Fradkin, and B. P. Kennedy. 2008. Effects of shoreline
1243 development on the nearshore environment in large deep oligotrophic lakes. *Freshwater
1244 Biology* **53**: 1673–1691. doi:10.1111/j.1365-2427.2008.01990.x
- 1245 Rosi-Marshall, E. J., D. W. Kincaid, H. A. Bechtold, T. V. Royer, M. Rojas, and J. J. Kelly. 2013.
1246 Pharmaceuticals suppress algal growth and microbial respiration and alter bacterial
1247 communities in stream biofilms. *Ecological Applications* **23**: 583–593. doi:10.1890/12-
1248 0491.1
- 1249 Rosi-Marshall, E. J., and T. V. Royer. 2012. Pharmaceutical Compounds and Ecosystem Function:
1250 An Emerging Research Challenge for Aquatic Ecologists. *Ecosystems* **15**: 867–880.
1251 doi:10.1007/s10021-012-9553-z
- 1252 Sargent, J. R., and S. Falk-Petersen. 1988. The lipid biochemistry of calanoid copepods.
1253 *Hydrobiologia* **167–168**: 101–114. doi:10.1007/BF00026297
- 1254 Schram, J. B., J. N. Kobelt, M. N. Dethier, and A. W. E. Galloway. 2018. Trophic Transfer of
1255 Macroalgal Fatty Acids in Two Urchin Species: Digestion, Egestion, and Tissue Building.
1256 *Front. Ecol. Evol.* **6**. doi:10.3389/fevo.2018.00083
- 1257 Shaw, L., C. Phung, and M. Grace. 2015. Pharmaceuticals and personal care products alter growth
1258 and function in lentic biofilms. *Environmental Chemistry* **12**: 301. doi:10.1071/EN14141

- 1259 Sitnikova, T. Ya. 2012. Определитель брюхоногих моллюсков бухты Большие Коты (юго-
1260 западное побережье озера Байкал) [Key of the Gastropod Molluscs in the Bay of Bolshie
1261 Koty (South-West shoreline of Lake Baikal)], Irkutsk State University.
- 1262 Slowikowski, K. 2019. ggrepel: Automatically Position Non-Overlapping Text Labels with
1263 "ggplot2,"
- 1264 Swann, G. E. A., V. N. Panizzo, S. Piccolroaz, and others. 2020. Changing nutrient cycling in Lake
1265 Baikal, the world's oldest lake. PNAS **117**: 27211–27217. doi:10.1073/pnas.2013181117
- 1266 Taipale, S., U. Strandberg, E. Peltomaa, A. W. E. Galloway, A. Ojala, and M. T. Brett. 2013. Fatty
1267 acid composition as biomarkers of freshwater microalgae: analysis of 37 strains of
1268 microalgae in 22 genera and in seven classes. Aquatic Microbial Ecology **71**: 165–178.
1269 doi:10.3354/ame01671
- 1270 Takhteev, V. V., and D. I. Didorenko. 2015. Fauna and ecology of amphipods of Lake Baikal: A
1271 Training manual, V.B. Sochava Institute of Geography SB RAS.
- 1272 Timoshkin, O. A., M. V. Moore, N. N. Kulikova, and others. 2018. Groundwater contamination by
1273 sewage causes benthic algal outbreaks in the littoral zone of Lake Baikal (East Siberia).
1274 Journal of Great Lakes Research. doi:10.1016/j.jglr.2018.01.008
- 1275 Timoshkin, O. A., D. P. Samsonov, M. Yamamoto, and others. 2016. Rapid ecological change in
1276 the coastal zone of Lake Baikal (East Siberia): Is the site of the world's greatest freshwater
1277 biodiversity in danger? Journal of Great Lakes Research **42**: 487–497.
1278 doi:10.1016/j.jglr.2016.02.011
- 1279 Tong, Y., M. Wang, J. Peñuelas, and others. 2020. Improvement in municipal wastewater treatment
1280 alters lake nitrogen to phosphorus ratios in populated regions. Proc Natl Acad Sci USA **117**:
1281 11566–11572. doi:10.1073/pnas.1920759117

- 1282 Turetsky, M. R., R. K. Wieder, C. J. Williams, and D. H. Vitt. 2000. Organic matter accumulation,
1283 peat chemistry, and permafrost melting in peatlands of boreal Alberta. *Écoscience* **7**: 115–
1284 122. doi:10.1080/11956860.2000.11682608
- 1285 Vendel, A. L., F. Bessa, V. E. N. Alves, A. L. A. Amorim, J. Patrício, and A. R. T. Palma. 2017.
1286 Widespread microplastic ingestion by fish assemblages in tropical estuaries subjected to
1287 anthropogenic pressures. *Marine Pollution Bulletin* **117**: 448–455.
1288 doi:10.1016/j.marpolbul.2017.01.081
- 1289 Volkova, E. A., N. A. Bondarenko, and O. A. Timoshkin. 2018. Morphotaxonomy, distribution and
1290 abundance of *Spirogyra* (Zygnematophyceae, Charophyta) in Lake Baikal, East Siberia.
1291 *Phycologia* **57**: 298–308. doi:10.2216/17-69.1
- 1292 Wang, W., and J. Wang. 2018. Investigation of microplastics in aquatic environments: An overview
1293 of the methods used, from field sampling to laboratory analysis. *TrAC Trends in Analytical
1294 Chemistry* **108**: 195–202. doi:10.1016/j.trac.2018.08.026
- 1295 Wickham, H. 2014. Tidy Data. *Journal of Statistical Software* **59**: 1–23. doi:10.18637/jss.v059.i10
- 1296 Wickham, H., M. Averick, J. Bryan, and others. 2019. Welcome to the tidyverse. *Journal of Open
1297 Source Software* **4**: 1686. doi:10.21105/joss.01686
- 1298 Wilke, C. O. 2019. cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2.”
- 1299 Wilkinson, M. D., M. Dumontier, Ij. J. Aalbersberg, and others. 2016. The FAIR Guiding
1300 Principles for scientific data management and stewardship. *Sci Data* **3**.
1301 doi:10.1038/sdata.2016.18
- 1302 Yang, Y., W. Song, H. Lin, W. Wang, L. Du, and W. Xing. 2018. Antibiotics and antibiotic
1303 resistance genes in global lakes: A review and meta-analysis. *Environment International*
1304 **116**: 60–73. doi:10.1016/j.envint.2018.04.011

- 1305 Yoshida, T., T. Sekino, M. Genkai-Kato, and others. 2003. Seasonal dynamics of primary
1306 production in the pelagic zone of southern Lake Baikal. Limnology **4**: 53–62.
1307 doi:10.1007/s10201-002-0089-3
- 1308 2016a. Methods for determination of nitrogen-containing matters (with corrections) (Методы
1309 определения азотсодержащих веществ (с Поправками)).
1310 2016b. Methods for determination of phosphorus-containing matters (with corrections) (Методы
1311 определения фосфорсодержащих веществ).
1312 2017. Nitrate concentration in waters: Photometric methods with Giress reagent following
1313 stabilization in a cadmium reducer (Массовая концентрация нитратного азота в водах:
1314 Методика измерений фотометрическим методом с реагентом Грисса после
1315 восстановления в камиевом редукторе).
1316
1317
1318

1319 **Acknowledgements**

1320
1321 We would like to thank the faculty, students, staff, and mariners of the Irkutsk State University's
1322 Biological Research Institute Biostation for their expert field, taxonomic, and laboratory support;
1323 Marianne Moore and Bart De Stasio for helpful advice; the researchers and students of the Siberian
1324 Branch of the Russian Academy of Sciences Limnological Institute for expert taxonomic and
1325 logistical assistance; Oleg A. Timoshkin, Tatiana Ya. Sitnikova, Irina V. Mekhanikova, Nina A.
1326 Bondorenko, Ekaterina Volkova, and Vadim V. Takhteev for offering insights and taxonomic
1327 training throughout the development of this project. Funding was provided by the National Science
1328 Foundation (NSF-DEB-1136637) to S.E.H., a Fulbright Fellowship to M.F.M., a NSF Graduate
1329 Research Fellowship to M.F.M. (NSF-DGE-1347973), and Russian Ministry of Science and
1330 Education (N FZZE-2020-0026; N FZZE-2020-0023). This work serves as one chapter of M.F.M.'s
1331 doctoral dissertation in Environmental and Natural Resource Sciences at Washington State
1332 University.

For Review Only

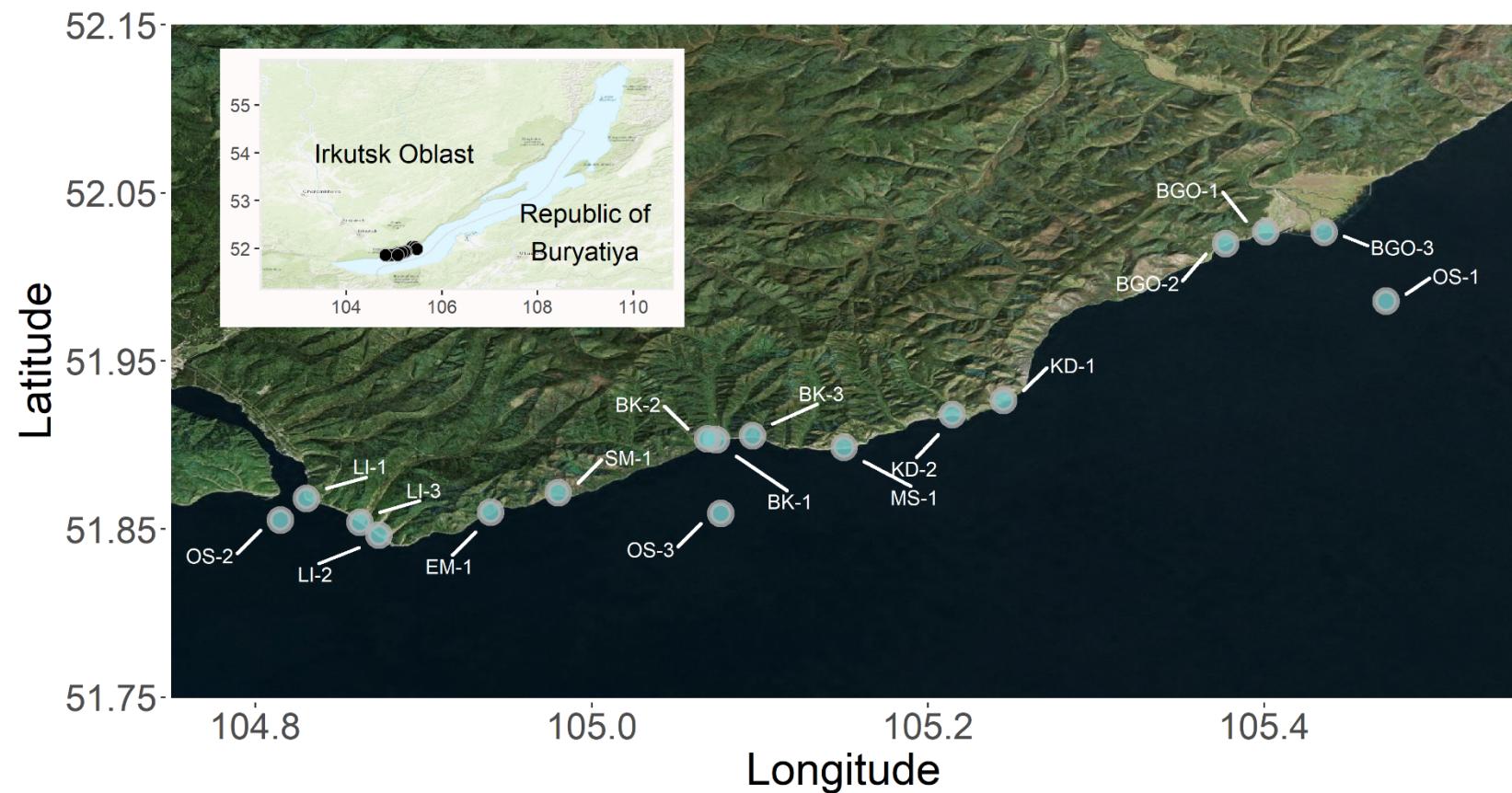


Figure 1: Map of all sampling locations with sites labeled with unique alphanumeric code. The entire transect included three developed sites (i.e., Listvyanka (LI), Bolshie Koty (BK), Bolshoe Goloustnoe (BGO)). Three offshore sites (OS) were also sampled to compare pelagic sewage signals to those in the littoral. Sites without adjacent lakeside development included Emelyanikha Bay (EM), Maloe Kadilnoe (KD), Mys Soboliny (MS), Sredny Mys (SM). Littoral sampling locations were all 8.90–20.75 m from shore and at a depth approximately of 0.75 m, whereas pelagic sites were approximately 2–5 km from shore and ranged in depth from 900 to 1300 m. This map was created using the R statistical environment (R Core Team 2019) and the tidyverse (Wickham et al. 2019), OpenStreetMap (Fellows and Stotz 2019), ggpubr (Kassambara 2019), cowplot (Wilke 2019), and ggrepel (Slowikowski 2019) packages.

Site	Latitude	Longitude	Depth (m)	Distance to shore (m)
BK-1	51.90316	105.074	0.7	10
BK-2	51.90365	105.069	0.9	17.5
BK-3	51.90536	105.0957	0.8	10
BGO-1	52.02693	105.401	0.9	18
BGO-2	52.0197	105.3771	1.1	14
BGO-3	52.02649	105.4358	0.7	21
OS-1	51.98559	105.4724	900	NA
KD-1	51.92646	105.245	0.8	20.75
KD-2	51.91807	105.2146	0.9	14.5
MS-1	51.89863	105.1502	0.6	10.5
SM-1	51.87152	104.9801	0.9	11.5
LI-1	51.86825	104.8304	0.6	8.9
LI-2	51.84626	104.8736	0.8	9.4
LI-3	51.85407	104.8622	0.7	9.25
EM-1	51.86005	104.94	0.7	15.5
OS-2	51.8553	104.8148	1300	NA
OS-3	51.85911	105.0769	1400	5000

Table 1: Locational information for each of the 17 sampling stations. “OS” refers to pelagic locations (i.e., “Offshore”), whereas other site abbreviations refer to littoral sampling locations.



Effects of spatially heterogeneous lakeside development on nearshore biotic communities in a large, deep, oligotrophic lake (Lake Baikal, Siberia)

Journal:	<i>Limnology and Oceanography</i>
Manuscript ID	Draft
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Meyer, Michael; Washington State University, School of the Environment Ozersky, Ted; University of Minnesota Duluth, Large Lakes Observatory Woo, Kara; Sage Bionetworks Shchapov, Kirill; University of Minnesota Duluth, Large Lakes Observatory Galloway, Aaron; University of Oregon Oregon Institute of Marine Biology Schram, Julie; University of Oregon Oregon Institute of Marine Biology Rosi, Emma; Cary Institute of Ecosystem Studies Snow, Daniel; University of Nebraska-Lincoln Timofeev, Maxim; Irkutsk State University Karnaughov, Dmitrii; Irkutsk State University Brousil, Matthew; Washington State University, Hampton, Stephanie; Washington State University, Center for Environmental Research, Education and Outreach
Keywords:	sewage, PPCP, food webs, fatty acids, human disturbance
Abstract:	Sewage released from lakeside development can reshape ecological communities. In particular, nearshore periphyton can rapidly assimilate sewage-associated nutrients, leading to increases of filamentous algal abundance, thus altering both food abundance and quality for grazers. In Lake Baikal, a large, ultra-oligotrophic, remote lake in Siberia, filamentous algal abundance has increased near lakeside developments, and localized sewage input is the suspected cause. These shifts are of particular interest in Lake Baikal, where endemic littoral biodiversity is high, lakeside settlements are mostly small, tourism is relatively high (~1.2 million visitors annually), and settlements are separated by large tracts of undisturbed shoreline, enabling investigation of heterogeneity and gradients of disturbance. We surveyed sites along 40 km of Baikal's southwestern shore for sewage indicators – pharmaceuticals and personal care products (PPCPs) and microplastics – as well as periphyton and macroinvertebrate abundance and indicators of food web structure (stable isotopes and fatty acids). PPCPs, including caffeine and acetaminophen/paracetamol, were spatially related to lakeside development. As predicted, lakeside development was associated with more filamentous algae and lower abundance of sewage-sensitive mollusks. Periphyton and macroinvertebrate stable isotopes and essential fatty acids suggested that food web structure otherwise

remained similar across sites; yet, the invariance of amphipod fatty acid composition, relative to periphyton, suggested that grazers adjust behavior or metabolism to compensate for different periphyton assemblages. Our results demonstrate that even low levels of human disturbance can result in spatial heterogeneity of nearshore ecological responses, with potential for creating less visible effects that propagate through the food web.

SCHOLARONE™
Manuscripts

Scientific Significance Statement Topic

See submitted text

Scientific Significance Statement Outlet

See submitted text

For Review Only

1 **Effects of spatially heterogeneous lakeside development on nearshore biotic communities in a large,**
2 **deep, oligotrophic lake (Lake Baikal, Siberia)**

3 Michael F. Meyer^{1*}

4 Ted Ozersky²

5 Kara H. Woo³

6 Kirill Shchapov²

7 Aaron W. E. Galloway⁴

8 Julie B. Schram⁴

9 Emma J. Rosi⁵

10 Daniel D. Snow⁶

11 Maxim A. Timofeyev⁷

12 Dmitry Yu. Karnaukhov⁷

13 Matthew R. Brousil³

14 Stephanie E. Hampton³

15

16 ¹. School of the Environment, Washington State University, Pullman, WA, USA

17 ². Large Lakes Observatory, University of Minnesota - Duluth, Duluth, MN, USA

18 ³. Center for Environmental Research, Education, and Outreach, Washington State University,
19 Pullman, WA, USA

20 ⁴. Oregon Institute of Marine Biology, University of Oregon, Charleston, OR, USA

21 ⁵. Cary Institute of Ecosystem Studies, Millbrook, NY, USA

22 ⁶. School of Natural Resources, University of Nebraska-Lincoln, Lincoln, NE, USA

23 ⁷. Biological Research Institute, Irkutsk State University, Irkutsk, Irkutsk Oblast, Russia

24 *corresponding author: michael.f.meyer@wsu.edu

25

26 **Keywords:** sewage, PPCP, food webs, fatty acids, human disturbance

27 **Statement of novelty, significance, and breadth of interest of the science presented in the
28 proposed manuscript**

29 We examined food web consequences of spatially heterogeneous disturbance associated with
30 small human settlements along the shoreline of remote oligotrophic Lake Baikal, illustrating that
31 nuanced biotic effects can occur at low levels of nutrient pollution. Using sewage-specific
32 indicators (pharmaceuticals and personal care products - PPCPs), we demonstrated that increased
33 nutrients at three discrete lakeside developments (80-1,963 permanent residents) and the
34 associated increased filamentous benthic algal abundance were consistent with sewage pollution.

35 This is the first study to provide such robust evidence that recent benthic algal blooms are caused
36 by sewage. These changes in benthic algae altered resources and nutrition for grazing
37 invertebrates, and macroinvertebrate composition accordingly differed at disturbed sites. Stable
38 isotope and fatty acid analysis of benthic algae and macroinvertebrates suggested that grazers at
39 sewage disturbed sites compensate for changing resource nutrition through behavior or altered
40 metabolism. This study demonstrates how patchy, low-level eutrophication of oligotrophic
41 systems can cause food webs to respond in less visible ways. Further, Baikal is an iconic ancient
42 lake, holding 20% of the world's fresh surface water and harboring the highest endemism and
43 biodiversity known in a lake; thus, understanding human impacts on this system is of both
44 societal importance and broad scientific interest.

45

46 **Statement indicating why L&O is the best outlet for the work**

47 This manuscript will appeal to L&O readers who are interested in both basic and applied issues.
48 From a basic ecology perspective, we investigate how bottom-up disturbances along a gradient
49 can propagate throughout a food web. From an applied perspective, we highlight how our results

50 can inform monitoring programs, especially as long-term littoral monitoring tends to be
51 uncommon relative to pelagic efforts. In addition, we use a suite of techniques that crosses
52 subdisciplines in a manner appreciated by limnologists and oceanographers, and less familiar to
53 other disciplines, such that L&O seems like the perfect home for this manuscript.

54

55 **Abstract**

56 Sewage released from lakeside development can reshape ecological communities. In particular,
57 nearshore periphyton can rapidly assimilate sewage-associated nutrients, leading to increases of
58 filamentous algal abundance, thus altering both food abundance and quality for grazers. In Lake
59 Baikal, a large, ultra-oligotrophic, remote lake in Siberia, filamentous algal abundance has
60 increased near lakeside developments, and localized sewage input is the suspected cause. These
61 shifts are of particular interest in Lake Baikal, where endemic littoral biodiversity is high,
62 lakeside settlements are mostly small, tourism is relatively high (~1.2 million visitors annually),
63 and settlements are separated by large tracts of undisturbed shoreline, enabling investigation of
64 heterogeneity and gradients of disturbance. We surveyed sites along 40 km of Baikal's
65 southwestern shore for sewage indicators – pharmaceuticals and personal care products (PPCPs)
66 and microplastics – as well as periphyton and macroinvertebrate abundance and indicators of
67 food web structure (stable isotopes and fatty acids). PPCPs, including caffeine and
68 acetaminophen/paracetamol, were spatially related to lakeside development. As predicted,
69 lakeside development was associated with more filamentous algae and lower abundance of
70 sewage-sensitive mollusks. Periphyton and macroinvertebrate stable isotopes and essential fatty
71 acids suggested that food web structure otherwise remained similar across sites; yet, the
72 invariance of amphipod fatty acid composition, relative to periphyton, suggested that grazers

73 adjust behavior or metabolism to compensate for different periphyton assemblages. Our results
74 demonstrate that even low levels of human disturbance can result in spatial heterogeneity of
75 nearshore ecological responses, with potential for creating less visible effects that propagate
76 through the food web.

77

78 **Introduction**

79 The release of treated and untreated wastewater into aquatic ecosystems is a common human
80 disturbance that can introduce pollutants and reshape aquatic ecological communities (Moore et
81 al. 2003). Nitrogen and phosphorus are among the primary pollutants in wastewater and its
82 associated byproducts (Smith et al. 1999), yet these nutrients can also originate from disparate
83 anthropogenic and natural environmental sources, thereby complicating their use as sewage
84 indicators. For example, agriculture (Powers et al. 2016), watershed processes such as melting
85 permafrost (Turetsky et al. 2000), and changes in terrestrial plant communities (Moran et al.
86 2012) can all increase allochthonous nutrient inputs similar to sewage. Regardless of the
87 nutrients' source, biological processes can further confound sewage detection. Benthic primary
88 producers, especially those in oligotrophic systems, can assimilate nutrients quickly from the
89 water column (e.g., hours), such that elevated nutrient concentrations may not be observed
90 (Hadwen and Bunn 2005).

91

92 Because nutrients come from numerous non-sewage sources, indicators consistently associated
93 with human activity, such as enhanced $\delta^{15}\text{N}$ stable isotope signatures (Costanzo et al. 2001;
94 Camilleri and Ozersky 2019), pharmaceuticals and personal care products (PPCPs) (Rosie
95 Marshall and Royer 2012; Meyer et al. 2019) and microplastics (Barnes et al. 2009), have
96 garnered increasing attention for their usefulness as sewage indicators. Stable isotopes, such as

97 $\delta^{15}\text{N}$, have been frequently used to trace sewage pollution (Gartner et al. 2002), yet their
98 potential to indicate sewage can be obfuscated by complex terrestrial (Craine et al. 2018) and
99 aquatic (Guzzo et al. 2011) processes. PPCP studies from continental (Kolpin et al. 2002;
100 Focazio et al. 2008; Yang et al. 2018) to colloidal pore (Yang et al. 2016) scales, have shown
101 that PPCP concentrations tend to be greatest closer to their source. In addition to identifying
102 areas and periods of sewage pollution, PPCPs have also demonstrated robustness in defining
103 gradients of sewage pollution in river systems, with concentrations being directly proportional to
104 population density and inversely proportional to distance from a densely populated area (Bendz
105 et al. 2005). Similar to PPCPs, microplastics (plastic debris up to 5 mm in size) also have been
106 useful to detect sewage pollution (Li et al. 2018) along gradients of increasing human population
107 density (Klein et al. 2015), although they can sometimes originate from non-sewage sources,
108 such as shoreline debris or fishing nets (Free et al. 2014). In contrast to $\delta^{15}\text{N}$ signatures and
109 PPCPs concentrations, microplastics are typically resistant to degradation (Barnes et al. 2009),
110 providing a signal over a longer time frame than many PPCPs and nutrients in sewage. As a
111 result of each pollutant's consistent association with sewage, co-located $\delta^{15}\text{N}$, PPCP, and
112 microplastic measurements can be used to infer the spatial extent and timing of sewage pollution
113 in an ecosystem.

114

115 The effects of sewage pollution are frequently first seen in nearshore benthic communities where
116 increased nutrients alter algal species composition, abundance, nutritional quality, as well as
117 food web trophic structure. Increased filamentous algal abundance, for example, has been
118 frequently observed in areas suspected of sewage pollution (Rosenberger et al. 2008; Hampton et
119 al. 2011), likely due to benthic filamentous algae efficiently removing nutrients from the water

120 column (Hadwen and Bunn 2005; Andersson and Brunberg 2006). With a changing resource
121 base, grazing macroinvertebrate communities may likewise shift to include more detritivores or
122 species capable of consuming filamentous algae (Rosenberger et al. 2008). In addition to some
123 grazers' physical difficulty consuming filamentous algae (Mazzella and Russo 1989), there also
124 may be changes in algal nutritional quality, as filamentous algae tend to contain a different
125 mixture of essential fatty acids (EFAs) in comparison to diatoms (Kelly and Scheibling 2012),
126 which dominate periphyton communities in unimpacted ecosystems. In particular, the EFAs
127 18:3ω3 and 18:2ω6 are commonly associated with green filamentous algae (Taipale et al. 2013),
128 whereas 20:5ω3 is more associated with diatoms (Taipale et al. 2013). All EFAs are largely
129 synthesized by primary producers, and each related group produces strongly differentiated
130 multivariate signatures (Taipale et al. 2013; Galloway and Winder 2015). Consumers can acquire
131 fatty acids by grazing (Dalsgaard et al. 2003) or upgrading fatty acids at their own energetic
132 expense (Sargent and Falk-Petersen 1988; Dalsgaard et al. 2003) and often reflect the fatty acid
133 signatures of their diets. Thus, comparing consumer and producer fatty acid compositions can be
134 used to infer how grazing patterns change in response to increasing sewage pollution.

135
136 To investigate lake littoral community and food web responses to sewage pollution, we surveyed
137 40 km of Lake Baikal's shoreline for indicators of sewage pollution and metrics of benthic
138 community composition and structure. Located in Siberia, Lake Baikal is the oldest, most
139 voluminous, and deepest freshwater lake in the world (Hampton et al. 2018), with the majority of
140 Lake Baikal's biodiversity occurring in the littoral zone (Kozhova and Izmest'eva 1998). While
141 Lake Baikal's pelagic zone is generally ultra-oligotrophic (Yoshida et al. 2003; O'Donnell et al.
142 2017), nearshore areas abutting lakeside settlements have shown distinct signs of eutrophication

143 (Timoshkin et al. 2016). Much of Lake Baikal's shoreline lacks human development, and
144 Baikal's watershed is largely roadless and unpopulated (Moore et al. 2009). Despite low levels of
145 development, uncharacteristic filamentous algal blooms have been occurring throughout the lake
146 since 2010 (Kravtsova et al. 2014; Timoshkin et al. 2016; Volkova et al. 2018). While increased
147 *Ulothrix* spp. abundance historically occurs in late summer (Kozhov 1963; Kozhova and
148 Izmest'eva 1998), recent observations of *Spirogyra* spp. abundance at unprecedented levels are
149 thought to be associated with increased nearshore nutrient concentrations (Volkova et al. 2018;
150 Ozersky et al. 2018). Inadequate wastewater management in lakeside settlements is likely the
151 main driver of these nearshore algal blooms (Timoshkin et al. 2016, 2018), motivating further
152 research to identify the extent to which sewage is altering nearshore communities

153

154 Given the growing evidence that Baikal's nearshore periphyton communities are responding to
155 sewage inputs, our goal was to understand how littoral benthic community composition and
156 interactions may be changing near areas of sewage pollution. This overarching goal was divided
157 into three specific objectives:

- 158 1. identify areas of wastewater pollution using consistent sewage indicators,
- 159 2. assess the relationship between sewage indicators and littoral periphyton and
- 160 macroinvertebrate community composition, and
- 161 3. evaluate how food webs may restructure with increasing sewage pollution.

162 We hypothesized that (1) sewage indicators, such as PPCP concentrations, $\delta^{15}\text{N}$, and
163 microplastic densities, would increase with increasing population density and proximity of
164 lakeside development; (2) an increasing sewage signal would correlate with increased dominance
165 of filamentous benthic algae; and (3) increasing filamentous algae abundance would result in

166 changes in the abundance of different macroinvertebrate feeding guilds, reflected in community
167 composition and dietary tracers such as carbon and nitrogen stable isotopes and fatty acids.

168

169 **Methods**

170 *Site description*

171 The vast majority of Lake Baikal's 2,100-km shoreline lacks lakeside development (Moore et al.
172 2009; Timoshkin et al. 2016). Our study focused on a 40-km section of Baikal's southwestern
173 shoreline, which included three settlements of different sizes (Figure 1; Figure 2). From 19
174 through 23 August 2015, we sampled 14 littoral and 3 pelagic locations along our 40-km
175 transect. Littoral locations were chosen to capture a range of sites with varying degrees of
176 adjacent shoreline development – from “developed” (along the waterfront of human settlements)
177 to “undeveloped” (no adjacent human settlements and complete forest cover; Figure 1; Figure 2;
178 Table 1). Pelagic sites were located 2 to 5 km offshore from each of the developed sites in water
179 depths of 900 to 1300 m (Figure 1; Table 1). All littoral sites were sampled at approximately the
180 same depth (~1.25 m) at a distance of 8.90 to 20.75 m from shore (Table 1). At each site, air
181 temperature was measured with a mercury thermometer, and photographs were taken of the
182 substrate and the shoreline.

183

184 Three discrete lakeside settlements were located along our 40-km transect. The largest,
185 Listvyanka, is primarily a tourist town of approximately 2000 permanent residents, although
186 tourism can contribute significantly to the town's population with approximately 1.2 million
187 annual visitors (Interfax-Tourism 2018). The other two settlements are the villages Bolshie Koty
188 and Bolshoe Goloustnoe, which have approximately 80 and 600 permanent residents,

189 respectively. Bolshie Koty is home to two field research stations and several small tourist
190 accommodations. Bolshoe Goloustnoe has several hotels and tourist camps. Although Bolshie
191 Koty and Bolshoe Goloustnoe are built along small streams that empty into Baikal, there are no
192 upstream developed sites, meaning that any observed sewage indicators in Baikal most likely
193 originated either from Bolshie Koty or Bolshoe Goloustnoe.

194

195 *Inverse distance weighted (IDW) population calculation*

196 We recognized that sewage indicator concentrations at each sampling location may be related to
197 a sampling location's spatial position relative to both the size and proximity of neighboring
198 developed sites. Therefore, we created the inverse distance weighted (IDW) population metric to
199 compress, into a single metric, information about human population size, density, and location
200 along the shoreline as well as distance between developed sites and sampling locations. The
201 IDW metric reflects the idea that sewage pollution should be positively related to increasing
202 human density and inversely related with distance from densely populated areas (sensu Bendz et
203 al., 2005). Additionally, Timoshkin et al. (2018) noted that sewage enters Baikal's nearshore
204 largely through groundwater, implying that locations with more directly adjacent shoreline
205 development should experience higher sewage concentrations in the lake. Acknowledging that
206 nearshore PPCP concentrations were likely positively proportional to a developed location's
207 shoreline length, we scaled a developed site's population density by its shoreline length. This
208 scaling represents population density that directly interfaces with the lake, thereby capturing the
209 idea that sewage-associated pollutants, such as PPCPs (Karnjanapiboonwong et al. 2010) and
210 nutrients (de Vries 1972), contributed away from the shoreline can be removed via the soil
211 matrix en route to the lake.

212

213 Our calculation of IDW population was done in five steps. First, we traced polygons and
214 shorelines from satellite imagery for each developed site in Google Earth. Polygons were traced
215 for the entire area of visible development (Figure 2). Similarly, shoreline traces only reflected
216 shoreline length for which there was visible development (Figure 2). Second, polygon and line
217 geometries were downloaded from Google Earth as a .kml file. Third, the .kml file was imported
218 into the R statistical environment (R Core Team 2019) where, using the sf package (Pebesma,
219 2018), we calculated shoreline length, polygon area, and centroid location for each developed
220 site. Fourth, we joined point locations of each sampling site with the spatial polygons to calculate
221 the distance from each sampling location to each developed site's centroid. Fifth, we calculated
222 IDW population for each sampling location, using formula (1)

$$223 (1) I_j = \frac{P_{LI} * L_{LI}}{A_{LI}} + \frac{P_{BK} * L_{BK}}{A_{BK}} + \frac{P_{BGO} * L_{BGO}}{A_{BGO}}$$

224 where I is the IDW population at sampling location j , P is the population at each of the three
225 developed sites Listvyanka (LI), Bolshie Koty (BK), Bolshoe Goloustnoe (BGO), A is the area of
226 a developed site in km^2 , L is the shoreline length at a developed site in km , and D is the distance
227 from sampling site j to each developed site's centroid in km . This formulation implies that all
228 sampling locations are influenced by all three developed sites. Thus, the influence of an
229 individual developed site on each sampling location is positively influenced by the size and
230 spatial density of the population and its orientation toward the shoreline, and inversely
231 proportional to a sampling location's distance from each of the three developed sites.

232

233 *Water samples*

234 At both pelagic and littoral sites, water samples were collected for nutrient, chlorophyll,
235 microplastic, and PPCP analysis. Samples were collected by hand from 0.75 m depth for each
236 littoral site and with a bucket from aboard the Irkutsk State University “Kozhov” research vessel
237 for pelagic sites. Each water sample collection procedure is described below.

238

239 *Nutrients*

240 Water samples for nutrient analyses were collected in 150 mL glass jars that had been washed
241 with phosphate-free soap and rinsed three times with water from the sampling location. Samples
242 were collected in duplicates and immediately frozen at -20°C until processing at the A.P.
243 Vinogradov Institute of Geochemistry (Siberian Branch of the Russian Academy of Sciences,
244 Irkutsk). Samples were not filtered prior to freezing, meaning that nitrogen and ammonium
245 concentrations may potentially include intracellular nitrogen and overestimate nitrogenous forms
246 in the water column.

247

248 For each water sample, nitrate, ammonium, and total phosphorus concentrations were measured.
249 For ammonium (2016a) and nitrate (2017) concentrations, samples were analyzed with a
250 spectrophotometer following the addition of Nessler’s reagent and disulfuric acid respectively.
251 Total phosphorus concentration was measured with a spectrophotometer following the addition
252 of persulfate (2016b). Concentrations are reported in mg/L.

253

254 *Chlorophyll a*

255 Water samples were collected in 1.5 L plastic bottles from a depth of approximately 0.75 m.
256 Within 12 h of collection, three subsamples (up to 150 mL each) were filtered through 25-mm

257 diameter, 0.2 μm pore size nitrocellulose filters. Filters were then placed in a 35-mm petri dish
258 and frozen in the dark until processing.

259

260 Chlorophyll samples were processed in a manner similar to that of Parsons and Strickland (1963)
261 and Lorenzen (1967). Nitrocellulose filters were ground in 90% acetone, in which chlorophyll
262 extraction was allowed to proceed overnight. Samples were then centrifuged for 15-20 minutes.
263 After centrifugation, absorbance of the chlorophyll extract was measured in a spectrophotometer
264 at 630, 645, 665, and 750 nm. Concentrations were calculated using the formula: $C =$
265 $11.64(A_{665} - A_{750}) - 2.16(A_{645} - A_{750}) - 0.1(A_{630} - A_{750}) / (V_2/V_1)$; where A is the
266 absorbance value of a particular wavelength, V₁ is the volume of the filtered water, and V₂ is the
267 volume of extract. Concentrations are reported as mg/L.

268

269 PPCPs

270 Water samples for PPCP analysis were collected in 250 mL amber glass bottles that were rinsed
271 with either methanol or acetone and then three times with sample water prior to collections.
272 Following collection, samples were refrigerated and kept in the dark until solid phase extraction
273 (SPE).

274

275 Within 12 h of collection, samples were filtered directly from the amber glass bottle using an in-
276 line Teflon filter holder with glass microfiber GMF (1.0 μm pore size, WhatmanGrad 934-AH)
277 in tandem with a solid phase extraction (SPE) cartridge (200 mg HLB, Waters Corporation,
278 Milford, MA) connected to a 1-liter vacuum flask. Lab personnel wore gloves and face masks to
279 minimize contamination. Prior to filtration, SPE cartridges were primed with at least 5 mL of

280 either methanol or acetone and then washed with at least 5 mL of sample water. Rate of
281 extraction was maintained at approximately 1 drop per second. Extraction proceeded until water
282 could no longer pass through the SPE cartridge or until all collected water was filtered.
283 Cartridges were stored in Whirlpacks at -20°C until analysis for 18 PPCP residues using liquid
284 chromatography tandem mass spectrometry (LC-MS-MS) following methods of Lee et al. (2016)
285 and D'Alessio et al (2018). Concentrations are reported in µg/L.

286

287 *Microplastics*

288 At each location, samples were collected in triplicate using 1.5 L clear plastic bottles that were
289 washed thoroughly with sample water before each collection. Samples were collected by hand
290 for each littoral site and with a metal bucket from aboard the ship for pelagic sites.

291

292 For processing, each sample was vacuum filtered on to a 47-mm diameter GF/F filter. During
293 filtration, aluminum foil was used to cover the filtration funnel to prevent contamination from
294 airborne microplastic particles. After filtration, filters were dried under vacuum pressure and
295 then stored in 50-mm petri dishes. Following filtration of all three replicates, the filtrate was
296 collected and then re-filtered through a GF/F filter as a control for contamination from the plastic
297 vacuum funnel or potentially airborne microplastics.

298

299 Microplastic counting involved visual inspection of the entire GF/F in a similar manner to
300 methods described in Hanvey et al. (2017). Visual enumeration was conducted under a stereo
301 microscope with ~100x magnification, and microplastics were classified into one of three
302 categories: fibers, fragments, or beads. For all categories, plastics were defined as observed

303 objects with apparent artificial colors, so as to not enumerate plastics potentially contributed
304 from the sampling bottle itself. Fibers were defined as smooth, long plastics with consistent
305 diameters. Fragments were defined as plastics with irregularly sharp or jagged edges. Beads were
306 defined as spherical plastics. Although we did not measure microplastic size, this technique
307 likely allowed us to reliably quantify microplastics as small as ~300 µm (Hanvey et al. 2017).
308 During enumeration, GF/Fs remained covered in the petri dish to minimize potential for
309 contamination from the air. Following enumeration of both experimental and control samples,
310 fibers, fragments, and beads enumerated in the controls were subtracted from the experimental
311 microplastic densities for each plastic type and from each replicate. One location (BK-1) had two
312 control replicates, which were averaged for each plastic type and then subtracted from the
313 experimental samples. Results are reported as the average number of microplastics/L.

314

315 *Benthic biological samples*

316 At each littoral site, periphyton and macroinvertebrates were collected for relative abundance
317 estimates and food web analysis by wading and snorkeling.

318

319 *Benthic algal collection*

320 At each littoral site, we haphazardly selected three rocks representative of local substrate. A
321 plastic stencil was used to define a surface area of each rock from which we scraped a
322 standardized 14.5 cm² patch of periphyton. Samples were preserved with Lugol's solution and
323 stored in plastic scintillation vials. Additional periphyton was collected in composite from each
324 site for fatty acid and stable isotope analysis.

325

326 Periphyton taxonomic identification and enumeration was performed by subsampling 10 µL
327 aliquots from each preserved sample. For all 10 µL aliquots, cells, filaments, and colonies were
328 counted, for the entire subsample, until at least 300 cells were identified for a given sampling
329 replicate. If the first aliquot contained less than 300 cells, we counted additional subsamples until
330 we reached at least 300 cells in total. In instances when 300 cells were counted before finishing a
331 subsample, we still counted the entire aliquot. Taxa were classified into broad categories
332 consistent with Baikal algal taxonomy (Izboldina 2007), using coarse groupings to capture
333 general patterns in relative algal abundance. As a result, algal groups consisted of diatoms,
334 *Ulothrix*, *Spirogyra*, and the green algal Order Tetrasporales.

335

336 *Benthic invertebrate collection*

337 At each littoral site, three kick-net samples were collected for assessment of benthic community
338 composition and abundance. Using a D-net, we collected macroinvertebrates by flipping over 1-3
339 rocks, and then sweeping five times in a left-to-right motion across approximately 1 m. After the
340 series of sweeps, the catch was rinsed into a plastic bucket. For each replicate, bucket contents
341 were concentrated using a 64-µm mesh and placed in glass jars with 40% ethanol (vodka; the
342 only preservative available to us at the time) for preservation and refrigerated at 4°C aboard the
343 research vessel. The 40% ethanol preservative was replaced with ~80% ethanol upon return to
344 the lab within 24 to 48 hours, and samples were stored at ~4°C.

345

346 Separate collections were conducted for invertebrate fatty acid and stable isotope analyses.
347 Invertebrates were collected using a D-net in a similar fashion as the community enumeration.
348 Additional invertebrates were also collected by hand. Collected organisms were then live-sorted,

349 identified to species, and frozen at -20°C at the field station. The samples were later transferred
350 to the lab in the US via a Dewar flask with dry ice.

351

352 Invertebrate taxonomic identification and enumeration were performed under a stereo
353 microscope. All invertebrates were identified to species with the exception of juveniles
354 (Takhteev and Didorenko (2015) for amphipods; Sitnikova (2012) for mollusks; Table 2). All
355 samples contained oligochaetes and polychaetes, but due to poor preservation, these taxa were
356 not counted. Six samples of the 42 collected were not well-preserved and were excluded from
357 further analyses, in order to reduce errors in identification. KD-1 and LI-1 were the only sites
358 with 1 sample counted. BK-2 and KD-2 each had two samples counted.

359

360 *Food web characterization*

361 To characterize littoral food webs, we analyzed carbon and nitrogen stable isotopes as well as
362 fatty acid profiles for periphyton and macroinvertebrates. Prior to isotopic and fatty acid
363 analysis, periphyton and macroinvertebrate samples were lyophilized for ~24 hours,
364 homogenized to powder, and then weighed.

365

366 *Stable isotope analysis*

367 Measurements of $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ were performed on an elemental analyzer-isotope ratio mass
368 spectrometer (EA-IRMS; Finnigan DELTAplus XP, Thermo Scientific) at the Large Lakes
369 Observatory, University of Minnesota Duluth. The EA-IRMS was calibrated against certified
370 reference materials including L-glutamic acid (NIST SRM 8574), low organic soil and sorghum
371 flour (standards B-2153 and B-2159 from Elemental Micro-analysis Ltd., Okehampton, UK) and

372 in-house standards (acetanilide and caffeine). Replicate analyses of external standards showed a
373 mean standard deviation of 0.06 ‰ and 0.09 ‰, for $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$, respectively.

374

375 *Fatty acid analysis*

376 Following freeze-drying, samples were transferred to 10 mL glass centrifuge vials, and 2 mL of
377 100% chloroform was added to each under nitrogen gas. Samples remained in chloroform
378 overnight at -80°C. Fatty acid extractions generally involved three phases: (1) 100% chloroform
379 extraction, (2) chloroform-methanol extraction, and (3) fatty acid methylation. Fatty acid
380 extraction methods were adapted from Schram et al. (2018).

381

382 After overnight chloroform extraction, samples underwent a chloroform-methanol extraction
383 three times. To each sample, we added 1 mL cooled 100% methanol, 1 mL chloroform:methanol
384 solution (2:1), and 0.8 mL 0.9% NaCl solution. Samples were inverted three times and sonicated
385 on ice for 10 minutes. Next, samples were vortexed for 1 minute, and centrifuged for 5 minutes
386 (3,000 rpm) at 4°C. Using a double pipette technique, the lower organic layer was removed and
387 kept under nitrogen. After the third extraction, samples were evaporated under nitrogen flow, and
388 resuspended in 1.5 mL chloroform and stored at -20°C overnight.

389

390 Once resuspended in chloroform, 1 mL of chloroform extract was transferred to a glass
391 centrifuge tube with a glass syringe as well as an internal standard of 4 μL of 19-carbon fatty
392 acid. Samples were then evaporated under nitrogen, and then 1 mL of toluene and 2 mL of 1%
393 sulfuric acid-methanol was added. The vial was closed under nitrogen gas and then incubated in
394 50°C water bath for 16 hours. After incubation, samples were removed from the bath, allowed to
395 reach room temperature and stored on ice. Next, we performed a potassium carbonate-hexane

396 extraction twice. To each sample, we added 2 mL of 2% potassium bicarbonate and 5 mL of
397 100% hexane, inverting the capped vial so as to mix the solution. Samples were centrifuged for 3
398 minutes (1,500 rpm) at 4°C. The upper hexane layer was then removed and placed in a vial to
399 evaporate under nitrogen flow. Once almost evaporated, 1 mL of 100% hexane was added and
400 stored in a glass amber autosampler vial for GC/MS quantification. GC/MS quantification was
401 performed with a Shimadzu QP2020 GC/MS following Schram et al. (2018).

402

403 *Statistical analyses*

404 Total phosphorus, nitrate, ammonium, microplastic abundance and density, total PPCP
405 concentration, and $\delta^{15}\text{N}$ values in macroinvertebrate tissues were log-transformed and regressed
406 against log-transformed IDW population using a linear model. Analytically, log-transforming
407 made sites comparable, as values spanned three orders of magnitude. Physically, we assumed
408 that sewage indicators were likely subject to exponential processes (e.g., mixing, diffusion), and
409 log-transforming the data should linearize the relationships between predictor and response
410 variables. Residuals were assessed for normality and homogeneity of variance.

411

412 To assess if benthic community composition was associated with increasing sewage indicators,
413 periphyton and macroinvertebrate abundance data were each analyzed with a consistent
414 multivariate workflow. First, replicates were averaged, and taxonomic groups representing less
415 than 1% of the inter-site community were removed from analysis, in order to reduce the
416 influence of rare species on results. Second, community compositions for both periphyton and
417 macroinvertebrates were visualized using non-metric multidimensional scaling (NMDS) with a
418 Bray-Curtis similarity metric. Periphyton community compositions were calculated as relative

proportions, whereas invertebrate abundances were grouped at the genus-level and then square-root transformed to minimize influence of more abundant taxa. Visual inspection of the NMDS plot suggested that sites generally tended to separate by increasing PPCP concentrations and IDW population (see Table 2). To test whether sites' benthic communities significantly differed with increasing PPCP concentration and IDW population, we first used k-medoids, also known as Partitioning Around the Medoids (PAM; Kaufman and Rousseeuw 2005), clustering to identify an optimal number of groupings (Figure S1). For this process, we iterated through multiple numbers of clusters (i.e., 1 to 10) and calculated the within-group-sum-of-squares (wss) and average silhouette width. We identified the optimal number of groups when wss decreased most markedly and when silhouette width was greatest (i.e., the elbow method) (Johnson and Wichern 2007). To confirm the optimal number as determined by non-hierarchical PAM clustering, we also used Weighted Pair-Group Centroid Clustering (WPGMC) as a hierarchical approach (Sneath and Sokal 1973), which corrects for clusters that may not be strongly discriminated regardless of how many samples are assigned to a given cluster (Legendre and Legendre 2012). We then performed two permutational multivariate analyses of variance (PERMANOVA; Anderson 2001) with 999 permutations: the first where community compositions were responses to the groups identified through clustering and the second where community compositions were responses to the continuous IDW population. Unlike traditional multivariate analyses of variance (MANOVA), PERMANOVA does not require assumptions of multivariate normality (Anderson 2001). When significant differences were identified, post-hoc SIMPER analysis (Clarke 1993) was performed following the PERMANOVA to identify which taxonomic groups contributed to 85% of the cumulative variance that most influenced site separation.

442

443 To assess whether benthic food webs restructured with increasing sewage indicator
444 concentrations, fatty acid data were analyzed in a manner similar to periphyton and
445 macroinvertebrate abundance data. First, species' fatty acid profiles were visualized by
446 performing NMDS with Bray-Curtis similarity for all organisms' relative fatty acid abundance
447 (Figure S2). This technique broadly demonstrated that, as expected, interspecific variation in
448 fatty acid composition was greater than intraspecific variation. The same pattern was observed
449 for all fatty acids quantified as well as solely essential fatty acids (EFAs; Figure S2). Together,
450 these NMDS plots suggested that periphyton fatty acids at sites differentiated based on sewage
451 indicator concentrations, which was likely a reflection of differences in periphyton community
452 composition (Taipale et al. 2013). Among all taxa and sites, 18:3ω3, 18:1ω9, and 20:5ω3 had the
453 highest coefficients of variation, enabling comparisons between sites. These fatty acids tend to
454 be associated with filamentous green algae (i.e., 18:3ω3 and 18:1ω9) and diatoms (i.e., 20:5ω3).
455 To increase the robustness of our analysis, we expanded our approach to include major fatty
456 acids within each taxonomic group, including 18:2ω6 (abundant in green algae); 16:1ω7 and
457 14:0 (abundant in diatoms); and 16:0 (abundant in both green algae and diatoms) (Taipale et al.
458 2013). To evaluate how relative fatty acid abundance may relate to sewage pollution, we
459 assessed patterns among these seven fatty acids with both multivariate and univariate
460 approaches. Within a multivariate framework, we created two NMDS plots with Bray-Curtis
461 similarity, one just with primary producer (Figure S5) and the other with macroinvertebrate
462 (Figure S6) fatty acid profiles. Because multivariate patterns suggested fatty acid profiles may
463 relate to sewage pollution, we regressed a filamentous:diatom fatty acid ratio (Equation 2)
464 (2)
$$\frac{18:3\omega3\% + 18:1\omega9\% + 18:2\omega6\% + 16:0\%}{20:5\omega3\% + 16:1\omega7\% + 16:0\% + 14:0\%}$$

465 against log-transformed PPCP concentrations using a linear model. Additionally, we evaluated
466 how three essential fatty acids (18:3ω3, 18:2ω6, and 20:5ω3), lipids thought to accumulate in
467 biological systems, may differ in abundance across the sewage gradient. Therefore, we similarly
468 regressed the ratio of $\frac{18:3\omega3\% + 18:2\omega6\%}{20:5\omega3\%}$ against log-transformed PPCP concentrations using a
469 linear model.

470

471 All analyses were conducted in the R statistical environment (R Core Team 2019), using the
472 tidyverse (Wickham et al. 2019), factoextra (Kassambara and Mundt 2019), cluster (Maechler et
473 al. 2019), pvclust (Suzuki et al. 2019), ggrepel (Slowikowski 2019), viridis (Garnier 2018), fs
474 (Hester and Wickham 2019), spdplyr (Sumner 2019), janitor (Firke 2020), sf (Pebesma 2018),
475 ggpibr (Kassambara 2019), ggttext (Wilke 2020), OpenStreetMap (Fellows and Stotz 2019),
476 cowplot (Wilke 2019), and vegan (Oksanen et al. 2019) packages. All data, including .kml files
477 used to calculate IDW metric, are publicly available from the Environmental Data Initiative
478 repository (Meyer et al. 2020), and all R scripts are available from the GitHub repository of this
479 project's Open Science Framework account (Meyer et al. 2015).

480

481 Results

482 Water samples

483 Nearshore water nitrate ($R^2 = 0.01$, $p = 0.68$), ammonium ($R^2 = 0.17$, $p = 0.11$), total phosphorus
484 ($R^2 = 0.14$, $p = 0.14$), and chlorophyll a ($R^2 = 0.11$, $p = 0.20$) concentrations were not
485 significantly correlated with IDW population (Figure 3). Total PPCP ($R^2 = 0.26$, $p = 0.04$)
486 concentrations were significantly related with IDW population (Figure 3). In the littoral zone,
487 PPCPs detected included caffeine, 1,7-dimethylxanthine/paraxanthine (main human metabolite

488 of caffeine), cotinine (main human metabolite of nicotine), and acetaminophen/paracetamol
489 (Table 3). Other PPCPs, including carbamazepine, diphenhydramine, thiabendazole,
490 amphetamine, methamphetamine, MDA, MDMA, morphine, phenazone, sulfachloropyridazine,
491 sulfamethazine, sulfadimethoxine, sulfamethazole, trimethoprim, and cimetidine, were not
492 detected.

493

494 Microplastics were detected in samples from both littoral and pelagic sites. Bead microplastics
495 were only detected near Listvyanka. Fibers (mean = 0.85 microplastics/L, std dev = 1.21
496 microplastics/L) and fragments (mean = 0.83 microplastics/L, std dev = 1.35 microplastics/L)
497 were the most abundant types of microplastics across all sites, whereas beads were relatively rare
498 (mean = 0.08 microplastics/L, std dev = 0.31 microplastics/L). Total microplastic densities were
499 not significantly correlated with IDW population ($R^2 = 0.01$, $p = 0.65$; Figure 3), although more
500 types of microplastics were generally observed near areas with higher IDW population values,
501 such as Listvyanka.

502

503 *Benthic biological samples*

504 *Periphyton*

505 Major taxonomic groupings of periphyton consisted of diatoms, *Tetrasporales* spp., *Spirogyra*
506 spp., and *Ulothrix* spp. K-medoids (Figures S1a; S2a) and WPGMC (Figure S3a) cluster
507 analyses of periphyton abundance demonstrated two groupings capture most variance, and visual
508 inspection of relative periphyton community abundance NMDS suggested groupings were
509 related to IDW population values (Figure 4). PERMANOVA results demonstrated that
510 periphyton communities were significantly different based on IDW population groupings ($R^2 =$

511 0.52, $p = 0.001$) and the continuous IDW population ($R^2 = 0.43$, $p = 0.001$). Post-hoc SIMPER
512 results suggested that these differences were primarily associated with sites that had higher
513 *Ulothrix* spp. and *Spirogyra* spp. relative abundance. Additionally, sites with high IDW
514 populations had lower diatom relative abundance in comparison to sites with low and moderate
515 IDW populations.

516

517 *Macroinvertebrates*

518 Taxonomic groupings included five amphipod genera: *Eulimnogammarus*, *Poekilogammarus*,
519 *Cryptoropus*, *Brandtia* and *Pallasea*; six mollusk families: Planorbidae, Valvatidae, Baicaliidae,
520 Benedictidae, Acroloxidae, Maackia; flatworms; caddisflies; and leeches (summarized in Table
521 S1). K-mediod cluster analysis of macroinvertebrate community composition suggested 2 or 3
522 major groupings would capture most variance (Figure S1b; S2b), whereas WPGMC analyses
523 suggested 2 groupings would enable all sites except for one to be assigned a cluster (S3b).
524 Because both forms of hierarchical and non-hierarchical clustering suggested two groupings as
525 optimal, we proceeded using two groupings. Visual inspection of NMDS suggested clusters were
526 related to IDW population (Figure 5). PERMANOVA results supported the hypothesis that
527 macroinvertebrate communities significantly differed both among our IDW population groupings
528 ($R^2 = 0.19$, $p = 0.02$) and along our continuous gradient of increasing IDW population ($R^2 =$
529 0.19 , $p = 0.02$). Post-hoc SIMPER analyses suggested that *Poekilogammarus*,
530 *Eulimnogammarus*, Valvatidae, Caddisflies, *Brandtia*, Baicaliidae, Planorbidae, *Cryptoropus*,
531 and flatworms contributed the greatest differences between high and moderate/low IDW
532 population groupings (see Table 2).

533

534 *Food web characterization: stable isotopes and fatty acids*

535 Among periphyton and amphipod samples, $\delta^{13}\text{C}$ values ranged from -19.5 to -9.5 ‰ (Figure 6).

536 Among periphyton samples, $\delta^{15}\text{N}$ values ranged from 0.77 to 3.76 ‰, whereas amphipod $\delta^{15}\text{N}$

537 values ranged from 6.42 to 7.92 ‰.

538

539 For grazers, $\delta^{15}\text{N}$ significantly increased with IDW population ($p = 0.01$; Figure 3, Figure 6).

540 Periphyton $\delta^{15}\text{N}$ signatures did not significantly increase with IDW population ($p = 0.27$). In

541 contrast, $\delta^{13}\text{C}$ concentrations were not related with IDW population for either periphyton or

542 macroinvertebrates.

543

544 With respect to fatty acids, macroinvertebrates tended to be characterized by mono-unsaturated

545 fatty acids (MUFAAs) and long-chain (i.e. ≥ 20 -Carbons) polyunsaturated fatty acids (LCPUFAs),

546 whereas periphyton tended to be characterized by short-chain (i.e., 16- and 18-Carbons)

547 polyunsaturated fatty acids (SCPUFAs) (Table 3). When comparing proportions within taxa

548 across the sewage gradient, periphyton SCPUFA proportion tended to increase (Figure S4) and

549 periphyton SAFA proportions generally decreased. In contrast, benthic macroinvertebrate fatty

550 acid class proportions tended to remain consistent across the entire gradient (Figure S4).

551

552 For both periphyton and grazers, our analyses focused mainly on the fatty acids consistently

553 associated with filamentous green algae (i.e., 18:3 ω 3, 18:1 ω 9, 18:2 ω 6, and 16:0) as well as

554 diatoms (i.e., 20:5 ω 3, 16:1 ω 7, 14:0, and 16:0). For periphyton, the ratio of green

555 filamentous:diatom-associated fatty acids significantly increased with an increasing PPCP

556 concentration ($R^2 = 0.62$; $p = 0.04$, Figure 7; S5) but not with an increasing IDW population ($p =$

557 0.08). Amphipod fatty acid ratios were not significantly related with either increasing IDW
558 population or increasing PPCP concentrations (Figure 7; S6). When focusing solely on the
559 essential fatty acids 18:3ω3, 18:2ω6, and 20:5ω3, the same pattern was observed in both
560 periphyton ($R^2 = 0.73$; $p = 0.02$) and amphipods (Figure 7).

561

562 Discussion

563 Our combined results corroborate previous findings (e.g., Timoshkin et al., 2016; 2018) that
564 sewage pollution is entering Lake Baikal's nearshore area and likely is responsible for changes in
565 nearshore benthic communities. Unlike previous studies, we were able to incorporate highly
566 specific indicators of sewage pollution and food web structure to offer direct, quantitative
567 relationships between human development and ecological responses.

568

569 *Relating human settlements to sewage indicator concentrations*

570 In agreement with our expectations, some sewage pollution indicators in the nearshore of Lake
571 Baikal were associated with size of and distance from human settlements. Total PPCP,
572 macroinvertebrate $\delta^{15}\text{N}$, and, to some degree, total phosphorus concentrations increased with
573 IDW population. These sewage gradients created by highly localized settlements are noteworthy
574 considering that Baikal's shoreline, including our study area, is largely free of lakeside
575 development (Moore et al. 2009). Furthermore, the use of sewage-associated indicators, such as
576 PPCPs and $\delta^{15}\text{N}$, proved necessary for defining sewage gradients. The use of nutrients as
577 indicators alone would not reveal sewage pollution gradients, since nutrients were not strongly
578 correlated with IDW population and could come from diverse sources. For example, melting
579 permafrost in Lake Baikal's watershed (Anisimov and Reneva 2006) and the Selenga River basin

580 (Tornqvist et al. 2014) as well as climate-driven changes in mixing processes (Swann et al. 2020)
581 have the potential to contribute substantial nutrient loadings to the nearshore. While nutrients
582 also could be contributed by agriculture (Powers et al. 2016), atmospheric deposition (Galloway
583 et al. 2004; Monteith et al. 2007; Stoddard et al. 2016), and changing terrestrial plant
584 communities (Moran et al. 2012), these are not currently known to be major sources of elevated
585 nutrients in the Baikal watershed, relative to sewage (Timoshkin et al., 2016, Timoshkin et al.,
586 2018) and permafrost melt (Anisimov & Reneva, 2006).

587

588 This is the first known study to detect PPCPs in Lake Baikal, a voluminous lake in a largely
589 unpopulated watershed. We detected PPCPs nearshore but not at our three offshore sites,
590 suggesting that sewage inputs in Baikal become diluted as pollutants move out of the nearshore
591 area. More generally, these results are important for lake monitoring, as PPCPs are robust
592 indicators of sewage pollution. Beyond Lake Baikal, these data are important for understanding
593 PPCPs' prevalence in lakes, as lakes have remained less represented in the PPCP literature in
594 comparison to lotic and subsurface systems (Meyer et al. 2019). This literature imbalance creates
595 opportunities to assess how PPCPs, and sewage pollution more broadly, may lead to differing
596 ecological responses in lotic and lentic systems. As lakes tend to have longer hydraulic residence
597 times relative to rivers and streams, pollutants may be more prone to accumulate (Yang et al.
598 2018; Meyer et al. 2019). In the case of our data, comparing contemporaneous littoral and
599 pelagic PPCP concentrations revealed littoral-pelagic sewage gradients, as PPCPs were
600 degraded, metabolized or accumulated by biota, preserved within sediments, or diluted to
601 undetectable concentrations. In the context of the entire lake, analyses of sediments have shown
602 how PPCPs can remain within lake systems for decades, thereby enabling researchers to

603 reconstruct histories of wastewater pollution in a system (Czekalski et al. 2015; Yang et al.
604 2018).

605

606 Investigating PPCP concentrations across limnic environments could also establish how
607 ecological communities respond differently not only to sewage but also to the PPCPs themselves.

608 While we focus on PPCPs as indicators of sewage, previous studies have shown that PPCPs,

609 even at concentrations we observed in Lake Baikal, can elicit biological responses from

610 physiological (e.g., del Rey et al. 2011; Feijão et al. 2020) and behavioral (e.g., Brodin et al.

611 2013; Dziewczynski et al. 2016) levels to food webs (e.g., Lagesson et al. 2016; Richmond et

612 al. 2018) and ecosystems (e.g., Rosi-Marshall et al. 2013; Richmond et al. 2019; Robson et al.

613 2020). Although our study was not designed to evaluate the ecotoxicological effects of PPCPs

614 themselves, future studies could potentially address effects of PPCPs on nearshore Baikal biota

615 by using *in situ* sewage gradients as a guide.

616

617 In contrast to PPCP concentrations and $\delta^{15}\text{N}$ values, microplastics were not associated with IDW
618 population and may be poor proxies for sewage pollution in Lake Baikal. Additionally,

619 microplastics may originate from non-sewage sources, such as agriculture (Steinmetz et al. 2016)

620 and fish nets (Eerkes-Medrano et al. 2015). Because of their long degradation time (Brandon et

621 al. 2016), microplastics can indicate accumulated pollution, which likely enables wider

622 distribution from nearshore inputs to the offshore (Fischer et al. 2016; Hendrickson et al. 2018).

623 Unlike microplastic concentrations identified in Lake Hovsgol (Free et al. 2014), Lake Superior

624 (Hendrickson et al. 2018), or Lake Erie (Eriksen et al. 2013), microplastic concentrations in

625 Baikal, as quantified by our methods, may be poor proxies for capturing pollution from

626 seasonally varying human populations. It is worth noting that since the time of our field
627 sampling, evidence has accumulated that our methods likely dramatically underestimated
628 microplastic abundance (Wang and Wang 2018; Brandon et al. 2020), and there is potential for
629 the microplastics themselves to cause deleterious ecological responses. While we focus here on
630 microplastics as an indicator of sewage pollution, microplastics are increasingly shown to disrupt
631 food web dynamics by altering grazing patterns (Green 2016) and providing carbon substrate for
632 microbial growth (Romera-Castillo et al. 2018). Recent investigations of microplastics in Lake
633 Baikal near Bolshie Koty (BK) used analogous methods and measured similarly low
634 concentrations (Karnaukhov et al. 2020). When considering Lake Baikal's large volume,
635 Karnaukov et al. (2020) noted that the number of plastic pieces may well exceed those observed
636 in other lakes, such as Lake Hovsgol. Together these growing uncertainties suggest that
637 microplastic pollution in Baikal and elsewhere deserves increased attention.

638

639 *Relating sewage indicators with benthic algal communities*

640 Congruent with our hypotheses, increasing sewage indicators tended to be associated with higher
641 relative abundance of filamentous taxa in periphyton. Previous studies investigating Baikal's
642 periphyton composition noted that areas adjacent to human development often had increased
643 abundance of filamentous algae such as *Ulothrix* and *Spirogyra* (Timoshkin et al. 2016, 2018).
644 Lake Baikal's southwestern shore historically experiences short *Ulothrix* blooms in late August
645 (Kozhov 1963), potentially confounding sewage signals with an annually occurring
646 phenomenon. Our data are consistent with the results of Timoshkin et al. (2016) and show that
647 relative abundance of filamentous algae is greatest near areas of higher lakeside development.

648

649 While community composition shifted with increasing sewage indicator concentrations,
650 periphyton $\delta^{15}\text{N}$ values did not differ along our transect. Previous studies in marine (Gartner et
651 al. 2002; Savage and Elmgren 2004; Risk et al. 2009) and freshwater (Wayland and Hobson
652 2001; Camilleri and Ozersky 2019) systems have highlighted how sewage-associated $\delta^{15}\text{N}$ can
653 increase in algal samples and even throughout the food web. Like PPCPs in our study, $\delta^{15}\text{N}$
654 values are often most enriched near the source of sewage pollution and can decrease over several
655 kilometers (Savage and Elmgren 2004), with concentrations varying based on species-specific
656 uptake rates and advective, dispersive, and diffusive transport (Gartner et al. 2002). While
657 previous studies using $\delta^{15}\text{N}$ signatures in macroalgae and vascular macrophytes have
658 successfully tracked sewage gradients (Cole et al. 2004), periphyton $\delta^{15}\text{N}$ as a sewage indicator
659 potentially can be confounded by terrestrial $\delta^{15}\text{N}$ contributions such as through agricultural
660 runoff (Chang et al. 2012). In our study, periphyton $\delta^{15}\text{N}$ signatures may be explained by
661 periphyton's typically high cell turnover rates (e.g., days; Swamikannu and Hoagland 1989)
662 dampening isotopic patterns, $\delta^{15}\text{N}$ -accumulating algal taxa being grazed more readily by
663 macroinvertebrates (Rosenberger et al. 2008), or co-limitation dynamics between ammonium and
664 nitrate (York et al. 2007; Piñón-Gimate et al. 2009).

665

666 Fatty acid analyses suggested that changes in periphyton community composition altered the
667 nutritional quality of periphyton across the pollution gradient. Periphyton fatty acid profiles from
668 sites with higher sewage pollution had higher levels of 18:3 ω 3, 18:1 ω 9, 18:2 ω 6, and 16:0
669 relative to 20:5 ω 3, 16:1 ω 7, 16:0, and 14:0 fatty acids. This pattern likely reflects the higher
670 abundance of green algae relative to diatoms (Iverson et al. 2004; Osipova et al. 2009; Taipale et
671 al. 2013; Galloway and Winder 2015; Shishlyannikov et al. 2018), which we observed from our

672 periphyton community composition analysis (Figure 3). Together, our periphyton composition
673 and fatty acid results suggest that Baikal's nearshore periphyton communities near human
674 lakeside developments are more dominated by filamentous green algae, and therefore, have
675 lower nutritional content.

676

677 Among the array of fatty acids synthesized in algal communities, essential fatty acids (EFAs) are
678 most likely to be taxonomically associated with, and influenced by, changing community
679 composition. EFAs are a subgroup of polyunsaturated fatty acids (PUFAs) that are prone to
680 accumulating in organisms (see Kelly & Scheibling, 2012). Among the eight common EFAs
681 (Taipale et al. 2013), 18:3 ω 3, 18:2 ω 6, and 20:5 ω 3 had the highest coefficients of variation
682 between sites. Because these three EFAs demonstrated the greatest variation between sites, our
683 analyses focused on how their relative abundances related to PPCP concentrations and IDW
684 populations. The fatty acids 18:3 ω 3 and 18:2 ω 6 have been previously associated with
685 filamentous algae, such as Baikalian *Ulothrix* (Osipova et al. 2009), whereas 20:5 ω 3 have
686 previously been associated with Baikalian diatoms (Shishlyannikov et al. 2018). Comparing the
687 ratio of filamentous green algae to diatoms could therefore function as proxy for each algal
688 taxon's relative abundance and potentially offer insights into feeding patterns for the grazers.

689

690 *Relating sewage indicators with macroinvertebrate feeding guilds*

691 In assessing benthic consumer communities' responses to changing periphyton, our data suggest
692 macroinvertebrate guilds reshape with increasing sewage pollution. Our results support the
693 general conclusion of Timoshkin et al. (2016) that Baikalian mollusk abundance tends to
694 decrease with increasing sewage pollution. Decreased mollusk abundance may have several

causes, including low tolerance for increased concentrations of PPCPs or other components of sewage (e.g., Hollingsworth et al. 2002, Timoshkin et al. 2016), inability to consume filamentous algae (Mazzella and Russo 1989), or filamentous algae not offering the proper nutrition (Lowe and Hunter 1988). In contrast to mollusks, amphipods were generally prevalent at all littoral sites, regardless of sewage indicator concentrations. *Brandtia* spp. was the only amphipod genus less abundant with sewage indicator signals. This genus tends to be associated with endemic sponges (Taakhteev & Didorenko, 2015), which may also be decreasing in abundance near areas of lakeside development (Timoshkin et al., 2016). *Eulimnogammarus* spp., one of the most speciose Baikal genera (Takhteev and Didorenko 2015), was prevalent at all sites, and $\delta^{15}\text{N}$ values in its tissue increased slightly but significantly with increasing IDW population. Unlike periphyton, amphipods' increasing $\delta^{15}\text{N}$ values may relate to amphipods having longer cellular turnover rates (e.g., weeks; McIntyre and Flecker 2006) relative to periphyton. Consequently, amphipods' enhanced $\delta^{15}\text{N}$ values suggest that sewage-derived nutrients are being incorporated into the food web. While we did not test amphipod tissues for other sewage indicators such as PPCPs and microplastics, the potential for PPCPs to bioaccumulate and biomagnify in food webs has been recently demonstrated, with ecological ramifications remaining uncertain (Lagesson et al., 2016; Richmond et al., 2018). These combined results suggest that mollusk abundance and amphipod $\delta^{15}\text{N}$ values may be longer-term indicators of sewage pollution in Baikal.

In contrast to variation in $\delta^{15}\text{N}$ values, amphipod fatty acid profiles did not differ markedly between sites (Figure 7). Amphipods from all collected sites expressed consistent 20:5 ω 3 signatures relative to 18:3 ω 3 and 18:2 ω 6. Consumers usually accumulate fatty acids from their food source. Yoshii's (1999) study as well as our own stable isotope data suggest that Baikal's

718 benthic, littoral amphipods are likely a combination of grazers and omnivores. Because fatty acid
719 profiles in amphipods largely reflected fatty acid signatures in periphyton, our data suggest that
720 amphipods likely continue grazing on periphyton, despite the food resource changing in
721 community composition and nutritional content. As a consequence, amphipods may be
722 compensating for the shifting nutritional quality of periphyton through at least two potential
723 mechanisms. First, amphipods may selectively consume diatoms as opposed to filamentous
724 algae, meaning diatom relative abundance could decrease both from increased grazing and lesser
725 efficiency at taking up nutrients relative to filamentous taxa. Second, amphipods themselves
726 (e.g., Desvilettes et al. 1997; Castell et al. 2004) or heterotrophic symbionts (Klein Breteler et al.
727 1999; Veloza et al. 2006; Hiltunen et al. 2017) may upgrade fatty acids by investing energy to
728 convert C18 fatty acids to C20 fatty acids. Regardless of the exact mechanism, our data suggest
729 that food web interactions would change with increasing sewage pollution and may imply a net
730 energetic cost through amphipods' differential grazing patterns.

731

732 *Conclusions*

733 Over the past decade, Lake Baikal has shown signs of nearshore eutrophication, despite the
734 pelagic zone remaining ultra-oligotrophic. While Baikal receives nutrients from multiple sources,
735 sewage-specific indicators used in this study implicate wastewater pollution as one of the
736 sources. Our results corroborate work by Timoshkin et al. (2016, 2018), demonstrating how
737 patchy hot spots of lakeside development at Baikal can create gradients in sewage concentrations
738 and ecological responses. Unlike previous studies, our approach pairs community abundance
739 data (i.e., periphyton and macroinvertebrate counts) and nuanced dietary tracers (i.e., fatty acids)
740 to assess benthic community and food web consequences of sewage pollution. While sewage

741 pollution may lead to changing resources for macroinvertebrate grazers, Baikal's amphipods
742 appear to be compensating either (1) by selectively grazing on diatoms or (2) by consuming less
743 desirable food and upgrading fatty acids. In both cases, our results suggest shifting community
744 interactions and may imply a net energetic cost for amphipods, as they expend energy either by
745 foraging selectively for diatoms or by catabolizing certain essential fatty acids.

746

747 *Future trajectories: a call for increased nearshore monitoring*

748 Our results underscore the importance of nearshore monitoring in detecting sewage pollution in
749 large lakes. Lake Baikal is considered ultra-oligotrophic based on pelagic sampling (Yoshida et
750 al. 2003; O'Donnell et al. 2017), but nearshore hot spots of eutrophication are developing
751 throughout the lake (Timoshkin et al. 2016, 2018). While pelagic samples are representative of
752 the lake's overall status, nearshore sampling aids managers in identifying pollution loading
753 before the entire system is affected (Jacoby et al. 1991; Lambert et al. 2008; Hampton et al.
754 2011). Beyond Baikal, several large, deep, oligotrophic lakes have likewise experienced
755 localized sewage pollution with nearshore biological responses, despite pelagic measurements
756 suggesting oligotrophic status (e.g., Jacoby et al. 1991, Rosenberger et al. 2008; Hampton et al.,
757 2011). Once eutrophication of the open water has occurred, mitigation can involve complex
758 socio-economic factors (Carpenter et al. 1999), require system-specific information (Jeppesen et
759 al. 2005), and necessitate long-term strategies (Tong et al. 2020). Because nutrients may enter
760 systems from numerous sources, incorporating sewage specific indicators, such as PPCPs, may
761 be necessary. PPCP sampling has the potential to not only identify sewage-associated nutrient
762 pollution but also assess heterogeneities in sewage loading along a shoreline. When PPCP data
763 are paired with co-located benthic community composition and food web data, managers can

764 take system-specific actions to mitigate ecological consequences before sewage concentrations
765 are detected throughout the lake. Across larger spatial and temporal scales, these paired PPCP-
766 biological samples have potential to offer a synoptic view of the impacts of sewage pollution,
767 enabling regional and local monitoring to coordinate mitigation strategies
768

For Review Only

769

770

Works Cited

- 771 Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral
772 Ecology* **26**: 32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x
- 773 Andersson, E., and A.-K. Brunberg. 2006. Inorganic nutrient acquisition in a shallow clearwater lake –
774 dominance of benthic microbiota. *Aquatic Sciences* **68**: 172–180. doi:10.1007/s00027-006-0805-
775 x
- 776 Anisimov, O., and S. Reneva. 2006. Permafrost and Changing Climate: The Russian Perspective. *Ambio*
777 **35**: 169–175.
- 778 Barnes, D. K. A., F. Galgani, R. C. Thompson, and M. Barlaz. 2009. Accumulation and fragmentation of
779 plastic debris in global environments. *Philos Trans R Soc Lond B Biol Sci* **364**: 1985–1998.
780 doi:10.1098/rstb.2008.0205
- 781 Bendz, D., N. A. Paxéus, T. R. Ginn, and F. J. Loge. 2005. Occurrence and fate of pharmaceutically
782 active compounds in the environment, a case study: Höje River in Sweden. *Journal of Hazardous
783 Materials* **122**: 195–204. doi:10.1016/j.jhazmat.2005.03.012
- 784 Brandon, J. A., A. Freibott, and L. M. Sala. 2020. Patterns of suspended and sulp-ingested microplastic
785 debris in the North Pacific investigated with epifluorescence microscopy. *Limnology and
786 Oceanography Letters* **5**: 46–53. doi:10.1002/lol2.10127
- 787 Brandon, J., M. Goldstein, and M. D. Ohman. 2016. Long-term aging and degradation of microplastic
788 particles: Comparing in situ oceanic and experimental weathering patterns. *Marine Pollution
789 Bulletin* **110**: 299–308. doi:10.1016/j.marpolbul.2016.06.048
- 790 Brodin, T., J. Fick, M. Jonsson, and J. Klaminder. 2013. Dilute Concentrations of a Psychiatric Drug
791 Alter Behavior of Fish from Natural Populations. *Science* **339**: 814–815.
792 doi:10.1126/science.1226850

- 793 Camilleri, A. C., and T. Ozersky. 2019. Large variation in periphyton $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in the upper
794 Great Lakes: Correlates and implications. *Journal of Great Lakes Research* **45**: 986–990.
795 doi:10.1016/j.jglr.2019.06.003
- 796 Carpenter, S. R., D. Ludwig, and W. A. Brock. 1999. Management of Eutrophication for Lakes Subject to
797 Potentially Irreversible Change. *Ecological Applications* **9**: 751–771. doi:10.2307/2641327
- 798 Castell, J. D., E. J. Kennedy, S. M. C. Robinson, G. J. Parsons, T. J. Blair, and E. Gonzalez-Duran. 2004.
799 Effect of dietary lipids on fatty acid composition and metabolism in juvenile green sea urchins
800 (*Strongylocentrotus droebachiensis*). *Aquaculture* **242**: 417–435.
801 doi:10.1016/j.aquaculture.2003.11.003
- 802 Chang, H.-Y., S.-H. Wu, K.-T. Shao, and others. 2012. Longitudinal variation in food sources and their
803 use by aquatic fauna along a subtropical river in Taiwan. *Freshwater Biology* **57**: 1839–1853.
804 doi:10.1111/j.1365-2427.2012.02843.x
- 805 Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. *Australian
806 Journal of Ecology* **18**: 117–143. doi:<https://doi.org/10.1111/j.1442-9993.1993.tb00438.x>
- 807 Cole, M. L., I. Valiela, K. D. Kroeger, and others. 2004. Assessment of a $\delta^{15}\text{N}$ isotopic method to
808 indicate anthropogenic eutrophication in aquatic ecosystems. *J. Environ. Qual.* **33**: 124–132.
809 doi:10.2134/jeq2004.1240
- 810 Costanzo, S. D., M. J. O'Donohue, W. C. Dennison, N. R. Loneragan, and M. Thomas. 2001. A New
811 Approach for Detecting and Mapping Sewage Impacts. *Marine Pollution Bulletin* **42**: 149–156.
812 doi:10.1016/S0025-326X(00)00125-9
- 813 Craine, J. M., A. J. Elmore, L. Wang, and others. 2018. Isotopic evidence for oligotrophication of
814 terrestrial ecosystems. *Nature Ecology & Evolution* **2**: 1735–1744. doi:10.1038/s41559-018-
815 0694-0
- 816 Czekalski, N., R. Sigdel, J. Birtel, B. Matthews, and H. Bürgmann. 2015. Does human activity impact the
817 natural antibiotic resistance background? Abundance of antibiotic resistance genes in 21 Swiss
818 lakes. *Environment International* **81**: 45–55. doi:10.1016/j.envint.2015.04.005

- 819 D'Alessio, M., S. Onanong, D. D. Snow, and C. Ray. 2018. Occurrence and removal of pharmaceutical
820 compounds and steroids at four wastewater treatment plants in Hawai'i and their environmental
821 fate. *Science of The Total Environment* **631–632**: 1360–1370.
822 doi:10.1016/j.scitotenv.2018.03.100
- 823 Dalsgaard, J., M. St. John, G. Kattner, D. Müller-Navarra, and W. Hagen. 2003. Fatty acid trophic
824 markers in the pelagic marine environment, p. 225–340. *In Advances in Marine Biology*.
825 Elsevier.
- 826 Desvilettes, Ch., G. Bourdier, and J. Ch. Breton. 1997. On the occurrence of a possible bioconversion of
827 linolenic acid into docosahexaenoic acid by the copepod *Eucyclops serrulatus* fed on microalgae.
828 *J Plankton Res* **19**: 273–278. doi:10.1093/plankt/19.2.273
- 829 Dziewczynski, T. L., B. A. Campbell, and J. L. Kane. 2016. Dose-dependent fluoxetine effects on
830 boldness in male Siamese fighting fish. *Journal of Experimental Biology* **219**: 797–804.
831 doi:10.1242/jeb.132761
- 832 Eerkes-Medrano, D., R. C. Thompson, and D. C. Aldridge. 2015. Microplastics in freshwater systems: A
833 review of the emerging threats, identification of knowledge gaps and prioritisation of research
834 needs. *Water Research* **75**: 63–82. doi:10.1016/j.watres.2015.02.012
- 835 Eriksen, M., S. Mason, S. Wilson, C. Box, A. Zellers, W. Edwards, H. Farley, and S. Amato. 2013.
836 Microplastic pollution in the surface waters of the Laurentian Great Lakes. *Marine Pollution
837 Bulletin* **77**: 177–182. doi:10.1016/j.marpolbul.2013.10.007
- 838 Feijão, E., R. Cruz de Carvalho, I. A. Duarte, and others. 2020. Fluoxetine Arrests Growth of the Model
839 Diatom *Phaeodactylum tricornutum* by Increasing Oxidative Stress and Altering Energetic and
840 Lipid Metabolism. *Front Microbiol* **11**. doi:10.3389/fmicb.2020.01803
- 841 Fellows, I., and using the Jm. library by J. P. Stotz. 2019. OpenStreetMap: Access to Open Street Map
842 Raster Images,.
- 843 Firke, S. 2020. janitor: Simple Tools for Examining and Cleaning Dirty Data.,

- 844 Fischer, E. K., L. Paglia longa, E. Czech, and M. Tamminga. 2016. Microplastic pollution in lakes and
845 lake shoreline sediments – A case study on Lake Bolsena and Lake Chiusi (central Italy).
846 Environmental Pollution **213**: 648–657. doi:10.1016/j.envpol.2016.03.012
- 847 Focazio, M. J., D. W. Kolpin, K. K. Barnes, E. T. Furlong, M. T. Meyer, S. D. Zaugg, L. B. Barber, and
848 M. E. Thurman. 2008. A national reconnaissance for pharmaceuticals and other organic
849 wastewater contaminants in the United States - II) Untreated drinking water sources. SCIENCE
850 OF THE TOTAL ENVIRONMENT **402**: 201–216. doi:10.1016/j.scitotenv.2008.02.021
- 851 Free, C. M., O. P. Jensen, S. A. Mason, M. Eriksen, N. J. Williamson, and B. Boldgiv. 2014. High-levels
852 of microplastic pollution in a large, remote, mountain lake. Marine Pollution Bulletin **85**: 156–
853 163. doi:10.1016/j.marpolbul.2014.06.001
- 854 Galloway, A. W. E., and M. Winder. 2015. Partitioning the Relative Importance of Phylogeny and
855 Environmental Conditions on Phytoplankton Fatty Acids. PLOS ONE **10**: e0130053.
856 doi:10.1371/journal.pone.0130053
- 857 Galloway, J. N., F. J. Dentener, D. G. Capone, and others. 2004. Nitrogen Cycles: Past, Present, and
858 Future. Biogeochemistry **70**: 153–226. doi:10.1007/s10533-004-0370-0
- 859 Garnier, S. 2018. viridis: Default Color Maps from “matplotlib.”
- 860 Gartner, A., P. Lavery, and A. J. Smit. 2002. Use of delta N-15 signatures of different functional forms of
861 macroalgae and filter-feeders to reveal temporal and spatial patterns in sewage dispersal. Mar.
862 Ecol.-Prog. Ser. **235**: 63–73. doi:10.3354/meps235063
- 863 Green, D. S. 2016. Effects of microplastics on European flat oysters, *Ostrea edulis* and their associated
864 benthic communities. Environmental Pollution **216**: 95–103. doi:10.1016/j.envpol.2016.05.043
- 865 Guzzo, M. M., G. D. Haffner, S. Sorge, S. A. Rush, and A. T. Fisk. 2011. Spatial and temporal
866 variabilities of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ within lower trophic levels of a large lake: implications for
867 estimating trophic relationships of consumers. Hydrobiologia **675**: 41–53. doi:10.1007/s10750-
868 011-0794-1

- 869 Hadwen, W. L., and S. E. Bunn. 2005. Food web responses to low-level nutrient and ^{15}N -tracer
870 additions in the littoral zone of an oligotrophic dune lake. *Limnology and Oceanography* **50**:
871 1096.
- 872 Hampton, S. E., S. C. Fradkin, P. R. Leavitt, and E. E. Rosenberger. 2011. Disproportionate importance
873 of nearshore habitat for the food web of a deep oligotrophic lake. *Marine and Freshwater
874 Research* **62**: 350. doi:10.1071/MF10229
- 875 Hampton, S. E., S. McGowan, T. Ozersky, and others. 2018. Recent ecological change in ancient lakes.
876 *Limnology and Oceanography* **63**: 2277–2304. doi:10.1002/lo.10938
- 877 Hanvey, J. S., P. J. Lewis, J. L. Lavers, N. D. Crosbie, K. Pozo, and B. O. Clarke. 2017. A review of
878 analytical techniques for quantifying microplastics in sediments. *Anal. Methods* **9**: 1369–1383.
879 doi:10.1039/C6AY02707E
- 880 Hendrickson, E., E. C. Minor, and K. Schreiner. 2018. Microplastic Abundance and Composition in
881 Western Lake Superior As Determined via Microscopy, Pyr-GC/MS, and FTIR. *Environ. Sci.
882 Technol.* **52**: 1787–1796. doi:10.1021/acs.est.7b05829
- 883 Hester, J., and H. Wickham. 2019. fs: Cross-Platform File System Operations Based on “libuv.”
- 884 Hiltunen, M., M. Honkanen, S. Taipale, U. Strandberg, and P. Kankaala. 2017. Trophic upgrading via the
885 microbial food web may link terrestrial dissolved organic matter to Daphnia. *J Plankton Res* **39**:
886 861–869. doi:10.1093/plankt/fbx050
- 887 Hollingsworth, R. G., J. W. Armstrong, and E. Campbell. 2002. Caffeine as a repellent for slugs and
888 snails. *Nature* **417**: 915–916. doi:10.1038/417915a
- 889 Interfax-Tourism. 2018. Байкал с января по август 2018 года посетили 1,2 миллиона туристов (1.2
890 million tourists vistied Baikal from January through August 2018). Interfax-Tourism, October 25
- 891 Iverson, S. J., C. Field, W. D. Bowen, and W. Blanchard. 2004. Quantitative Fatty Acid Signature
892 Analysis: A New Method of Estimating Predator Diets. *Ecological Monographs* **74**: 211–235.
893 doi:10.1890/02-4105

- 894 Izhboldina, L. A. 2007. Guide and Key to Benthic and Periphyton Algae of Lake Baikal (meio- and
895 macrophytes) with Brief Notes on Their Ecology, Nauka-Centre.
- 896 Jacoby, J. M., D. D. Bouchard, and C. R. Patmont. 1991. Response of Periphyton to Nutrient Enrichment
897 in Lake Chelan, WA. *Lake and Reservoir Management* 7: 33–43.
898 doi:10.1080/07438149109354252
- 899 Jeppesen, E., M. Søndergaard, J. P. Jensen, and others. 2005. Lake responses to reduced nutrient loading
900 – an analysis of contemporary long-term data from 35 case studies. *Freshwater Biology* 50:
901 1747–1771. doi:10.1111/j.1365-2427.2005.01415.x
- 902 Johnson, R. A., and D. V. Wichern. 2007. *Applied Multivariate Statistical Analysis*, 6th ed. Prentice Hall.
- 903 Karnaukhov, D., S. Biritskaya, E. Dolinskaya, M. Teplykh, N. Silenko, Y. Ermolaeva, and E. Silow.
904 2020. POLLUTION BY MACRO- AND MICROPLASTIC OF LARGE LACUSTRINE
905 ECOSYSTEMS IN EASTERN ASIA. *Pollution Research* 2: 353–355.
- 906 Karnjanapiboonwong, A., A. N. Morse, J. D. Maul, and T. A. Anderson. 2010. Sorption of estrogens,
907 triclosan, and caffeine in a sandy loam and a silt loam soil. *Journal of Soils and Sediments* 10:
908 1300–1307. doi:10.1007/s11368-010-0223-5
- 909 Kassambara, A. 2019. ggpibr: “ggplot2” Based Publication Ready Plots.,
910 Kassambara, A., and F. Mundt. 2019. factoextra: Extract and Visualize the Results of Multivariate Data
911 Analyses.,
- 912 Kaufman, L., and P. J. Rousseeuw. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*,
913 1st Edition. Wiley-Interscience.
- 914 Kelly, J. R., and R. E. Scheibling. 2012. Fatty acids as dietary tracers in benthic food webs. *Marine
915 Ecology Progress Series* 446: 1–22. doi:10.3354/meps09559
- 916 Klein Breteler, W. C. M., N. Schogt, M. Baas, S. Schouten, and G. W. Kraay. 1999. Trophic upgrading of
917 food quality by protozoans enhancing copepod growth: role of essential lipids. *Marine Biology*
918 135: 191–198. doi:10.1007/s002270050616

- 919 Klein, S., E. Worch, and T. P. Knepper. 2015. Occurrence and Spatial Distribution of Microplastics in
920 River Shore Sediments of the Rhine-Main Area in Germany. *Environ. Sci. Technol.* **49**: 6070–
921 6076. doi:10.1021/acs.est.5b00492
- 922 Kolpin, D. W., E. T. Furlong, M. T. Meyer, E. M. Thurman, S. D. Zaugg, L. B. Barber, and H. T. Buxton.
923 2002. Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams,
924 1999–2000: A National Reconnaissance. *Environmental Science & Technology* **36**: 1202–1211.
925 doi:10.1021/es011055j
- 926 Kozhov, M. M. 1963. *Lake Baikal and its Life*, Springer Science & Business Media.
- 927 Kozhova, O. M., and L. R. Izmest'eva. 1998. *Lake Baikal: Evolution and Biodiversity*, Backhuys
928 Publishers.
- 929 Kravtsova, L. S., L. A. Izhboldina, I. V. Khanaev, and others. 2014. Nearshore benthic blooms of
930 filamentous green algae in Lake Baikal. *Journal of Great Lakes Research* **40**: 441–448.
931 doi:10.1016/j.jglr.2014.02.019
- 932 Lagesson, A., J. Fahlman, T. Brodin, J. Fick, M. Jonsson, P. Byström, and J. Klaminder. 2016.
933 Bioaccumulation of five pharmaceuticals at multiple trophic levels in an aquatic food web -
934 Insights from a field experiment. *Science of The Total Environment* **568**: 208–215.
935 doi:10.1016/j.scitotenv.2016.05.206
- 936 Lambert, D., A. Cattaneo, and R. Carignan. 2008. Periphyton as an early indicator of perturbation in
937 recreational lakes. *Can. J. Fish. Aquat. Sci.* **65**: 258–265. doi:10.1139/f07-168
- 938 Legendre, P., and L. Legendre. 2012. *Numerical Ecology*, 3rd ed. Elsevier.
- 939 Li, J., C. Green, A. Reynolds, H. Shi, and J. M. Rotchell. 2018. Microplastics in mussels sampled from
940 coastal waters and supermarkets in the United Kingdom. *Environmental Pollution* **241**: 35–44.
941 doi:10.1016/j.envpol.2018.05.038
- 942 Lorenzen, C. J. 1967. Determination of Chlorophyll and Pheo-Pigments: Spectrophotometric Equations1.
943 Limnology and Oceanography **12**: 343–346. doi:<https://doi.org/10.4319/lo.1967.12.2.0343>

- 944 Lowe, R. L., and R. D. Hunter. 1988. Effect of Grazing by *Physa integra* on Periphyton Community
945 Structure. *Journal of the North American Benthological Society* **7**: 29–36. doi:10.2307/1467828
- 946 Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2019. *cluster: Cluster Analysis Basics*
947 and Extensions.,
- 948 Mazzella, L., and G. F. Russo. 1989. Grazing effect of two *Gibbula* species (Mollusca,
949 Archaeogastropoda) on the epiphytic community of *Posidonia oceanica* leaves.
- 950 McIntyre, P. B., and A. S. Flecker. 2006. Rapid turnover of tissue nitrogen of primary consumers in
951 tropical freshwaters. *Oecologia* **148**: 12–21. doi:10.1007/s00442-005-0354-3
- 952 Meyer, M. F., T. Ozersky, K. H. Woo, and others. 2020. A unified dataset of co-located sewage pollution,
953 periphyton, and benthic macroinvertebrate community and food web structure from Lake Baikal
954 (Siberia).doi:10.6073/PASTA/76F43144015EC795679BAC508EFA044B
- 955 Meyer, M. F., S. M. Powers, and S. E. Hampton. 2019. An Evidence Synthesis of Pharmaceuticals and
956 Personal Care Products (PPCPs) in the Environment: Imbalances among Compounds, Sewage
957 Treatment Techniques, and Ecosystem Types. *Environ. Sci. Technol.* **53**: 12961–12973.
958 doi:10.1021/acs.est.9b02966
- 959 Meyer, M., T. Ozersky, K. Woo, A. W. E. Galloway, M. R. Brousil, and S. Hampton. 2015. Baikal Food
960 Webs.doi:10.17605/OSF.IO/9TA8Z
- 961 Monteith, D. T., J. L. Stoddard, C. D. Evans, and others. 2007. Dissolved organic carbon trends resulting
962 from changes in atmospheric deposition chemistry. *Nature* **450**: 537–540.
963 doi:10.1038/nature06316
- 964 Moore, J. W., D. E. Schindler, M. D. Scheuerell, D. Smith, and J. Frogge. 2003. Lake eutrophication at
965 the urban fringe, Seattle region, USA. *AMBIO: A Journal of the Human Environment* **32**: 13–18.
- 966 Moore, M. V., S. E. Hampton, L. R. Izmost'eva, E. A. Silow, E. V. Peshkova, and B. K. Pavlov. 2009.
967 Climate Change and the World's "Sacred Sea"-Lake Baikal, Siberia. *Bioscience* **59**: 405–417.
968 doi:10.1525/bio.2009.59.5.8

- 969 Moran, P. W., S. E. Cox, S. S. Embrey, R. L. Huffman, T. D. Olsen, and S. C. Fradkin. 2012. Sources and
970 Sinks of Nitrogen and Phosphorus in a Deep, Oligotrophic Lake, Lake Crescent, Olympic
971 National Park, Washington. US Geological Survey.
- 972 O'Donnell, D. R., P. Wilburn, E. A. Silow, L. Y. Yampolsky, and E. Litchman. 2017. Nitrogen and
973 phosphorus colimitation of phytoplankton in Lake Baikal: Insights from a spatial survey and
974 nutrient enrichment experiments. Limnology and Oceanography **62**: 1383–1392.
975 doi:10.1002/lno.10505
- 976 Oksanen, J., F. G. Blanchet, M. Friendly, and others. 2019. vegan: Community Ecology Package.,
977 Osipova, S., L. Dudareva, N. Bondarenko, A. Nasarova, N. Sokolova, L. Obolkina, O. Glyzina, and O.
978 Timoshkin. 2009. Temporal variation in fatty acid composition of *Ulothrix zonata* (Chlorophyta)
979 from ice and benthic communities of Lake Baikal. Phycologia **48**: 130–135.
- 980 Ozersky, T., E. A. Volkova, N. A. Bondarenko, O. A. Timoshkin, V. V. Malnik, V. M. Domysheva, and
981 S. E. Hampton. 2018. Nutrient limitation of benthic algae in Lake Baikal, Russia. Freshwater
982 Science **37**: 472–482. doi:10.1086/699408
- 983 Parsons, T. R., and J. D. H. Strickland. 1963. Discussion of spectrophotometric determination of marine-
984 plant pigments, with revised equations for ascertaining chlorophylls and carotenoids. Journal of
985 Marine Research.
- 986 Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal
987 **10**: 439–446. doi:10.32614/RJ-2018-009
- 988 Piñón-Gimate, A., M. F. Soto-Jiménez, M. J. Ochoa-Izaguirre, E. García-Pagés, and F. Páez-Osuna. 2009.
989 Macroalgae blooms and $\delta^{15}\text{N}$ in subtropical coastal lagoons from the Southeastern Gulf of
990 California: Discrimination among agricultural, shrimp farm and sewage effluents. Marine
991 Pollution Bulletin **58**: 1144–1151. doi:10.1016/j.marpolbul.2009.04.004
- 992 Powers, S. M., T. W. Bruulsema, T. P. Burt, and others. 2016. Long-term accumulation and transport of
993 anthropogenic phosphorus in three river basins. Nature Geoscience **9**: 353–356.
994 doi:10.1038/ngeo2693

- 995 R Core Team. 2019. R: A Language and Environment for Statistical Computing.,
- 996 del Rey, Z. R., E. F. Granek, and B. A. Buckley. 2011. Expression of HSP70 in *Mytilus californianus*
997 following exposure to caffeine. *Ecotoxicology* **20**: 855–861. doi:10.1007/s10646-011-0649-6
- 998 Richmond, E. K., E. J. Rosi, A. J. Reisinger, B. R. Hanrahan, R. M. Thompson, and M. R. Grace. 2019.
999 Influences of the antidepressant fluoxetine on stream ecosystem function and aquatic insect
1000 emergence at environmentally realistic concentrations. *Journal of Freshwater Ecology* **34**: 513–
1001 531. doi:10.1080/02705060.2019.1629546
- 1002 Richmond, E. K., E. J. Rosi, D. M. Walters, J. Fick, S. K. Hamilton, T. Brodin, A. Sundelin, and M. R.
1003 Grace. 2018. A diverse suite of pharmaceuticals contaminates stream and riparian food webs.
1004 *Nature Communications* **9**: 4491. doi:10.1038/s41467-018-06822-w
- 1005 Risk, M. J., B. E. Lapointe, O. A. Sherwood, and B. J. Bedford. 2009. The use of δ15N in assessing
1006 sewage stress on coral reefs. *Marine Pollution Bulletin* **58**: 793–802.
1007 doi:10.1016/j.marpolbul.2009.02.008
- 1008 Robson, S. V., E. J. Rosi, E. K. Richmond, and M. R. Grace. 2020. Environmental concentrations of
1009 pharmaceuticals alter metabolism, denitrification, and diatom assemblages in artificial streams.
1010 *Freshwater Science* **39**: 256–267. doi:10.1086/708893
- 1011 Romera-Castillo, C., M. Pinto, T. M. Langer, X. A. Álvarez-Salgado, and G. J. Herndl. 2018. Dissolved
1012 organic carbon leaching from plastics stimulates microbial activity in the ocean. *Nat Commun* **9**:
1013 1–7. doi:10.1038/s41467-018-03798-5
- 1014 Rosenberger, E. E., S. E. Hampton, S. C. Fradkin, and B. P. Kennedy. 2008. Effects of shoreline
1015 development on the nearshore environment in large deep oligotrophic lakes. *Freshwater Biology*
1016 **53**: 1673–1691. doi:10.1111/j.1365-2427.2008.01990.x
- 1017 Rosi-Marshall, E. J., D. W. Kincaid, H. A. Bechtold, T. V. Royer, M. Rojas, and J. J. Kelly. 2013.
1018 Pharmaceuticals suppress algal growth and microbial respiration and alter bacterial communities
1019 in stream biofilms. *Ecological Applications* **23**: 583–593. doi:10.1890/12-0491.1

- 1020 Rosi-Marshall, E. J., and T. V. Royer. 2012. Pharmaceutical Compounds and Ecosystem Function: An
1021 Emerging Research Challenge for Aquatic Ecologists. *Ecosystems* **15**: 867–880.
1022 doi:10.1007/s10021-012-9553-z
- 1023 Sargent, J. R., and S. Falk-Petersen. 1988. The lipid biochemistry of calanoid copepods. *Hydrobiologia*
1024 **167–168**: 101–114. doi:10.1007/BF00026297
- 1025 Savage, C., and R. Elmgren. 2004. MACROALGAL (*FUCUS VESICULOSUS*) $\delta^{15}\text{N}$ VALUES TRACE
1026 DECREASE IN SEWAGE INFLUENCE. *Ecological Applications* **14**: 517–526. doi:10.1890/02-
1027 5396
- 1028 Schram, J. B., J. N. Kobelt, M. N. Dethier, and A. W. E. Galloway. 2018. Trophic Transfer of Macroalgal
1029 Fatty Acids in Two Urchin Species: Digestion, Egestion, and Tissue Building. *Front. Ecol. Evol.*
1030 **6**. doi:10.3389/fevo.2018.00083
- 1031 Shishlyannikov, S. M., A. A. Nikonova, Y. S. Bukin, and A. G. Gorshkov. 2018. Fatty acid trophic
1032 markers in Lake Baikal phytoplankton: A comparison of endemic and cosmopolitan diatom-
1033 dominated phytoplankton assemblages. *Ecological Indicators* **85**: 878–886.
1034 doi:10.1016/j.ecolind.2017.11.052
- 1035 Sitnikova, T. Ya. 2012. Определитель брюхоногих моллюсков бухты Большие Коты (юго-западное
1036 побережье озера Байкал) [Key of the Gastropod Molluscs in the Bay of Bolshie Koty (South-
1037 West shoreline of Lake Baikal)], Irkutsk State University.
- 1038 Slowikowski, K. 2019. *ggrepel*: Automatically Position Non-Overlapping Text Labels with “*ggplot2*,”
- 1039 Smith, V. H., G. D. Tilman, and J. C. Nekola. 1999. Eutrophication: impacts of excess nutrient inputs on
1040 freshwater, marine, and terrestrial ecosystems. *Environmental Pollution* **100**: 179–196.
1041 doi:10.1016/S0269-7491(99)00091-3
- 1042 Sneath, P. H. A., and R. R. Sokal. 1973. *Numerical Taxonomy: The Principles and Practice of Numerical
1043 Classification*, W. H. Freeman.

- 1044 Steinmetz, Z., C. Wollmann, M. Schaefer, and others. 2016. Plastic mulching in agriculture. Trading
1045 short-term agronomic benefits for long-term soil degradation? *Science of The Total Environment*
1046 **550**: 690–705. doi:10.1016/j.scitotenv.2016.01.153
- 1047 Stoddard, J. L., J. Van Sickle, A. T. Herlihy, J. Brahney, S. Paulsen, D. V. Peck, R. Mitchell, and A. I.
1048 Pollard. 2016. Continental-Scale Increase in Lake and Stream Phosphorus: Are Oligotrophic
1049 Systems Disappearing in the United States? *Environ. Sci. Technol.* **50**: 3409–3415.
1050 doi:10.1021/acs.est.5b05950
- 1051 Sumner, M. D. 2019. *spdpqr: Data Manipulation Verbs for the Spatial Classes*,
1052 Suzuki, R., Y. Terada, and H. Shimodaira. 2019. *pvclust: Hierarchical Clustering with P-Values via*
1053 *Multiscale Bootstrap Resampling*,
1054 Swamikannu, X., and K. D. Hoagland. 1989. Effects of Snail Grazing on the Diversity and Structure of a
1055 Periphyton Community in a Eutrophic Pond. *Can. J. Fish. Aquat. Sci.* **46**: 1698–1704.
1056 doi:10.1139/f89-215
- 1057 Swann, G. E. A., V. N. Panizzo, S. Piccolroaz, and others. 2020. Changing nutrient cycling in Lake
1058 Baikal, the world's oldest lake. *PNAS* **117**: 27211–27217. doi:10.1073/pnas.2013181117
- 1059 Taipale, S., U. Strandberg, E. Peltomaa, A. W. E. Galloway, A. Ojala, and M. T. Brett. 2013. Fatty acid
1060 composition as biomarkers of freshwater microalgae: analysis of 37 strains of microalgae in 22
1061 genera and in seven classes. *Aquatic Microbial Ecology* **71**: 165–178. doi:10.3354/ame01671
- 1062 Takhteev, V. V., and D. I. Didorenko. 2015. Fauna and ecology of amphipods of Lake Baikal: A Training
1063 manual, V.B. Sochava Institute of Geography SB RAS.
- 1064 Timoshkin, O. A., M. V. Moore, N. N. Kulikova, and others. 2018. Groundwater contamination by
1065 sewage causes benthic algal outbreaks in the littoral zone of Lake Baikal (East Siberia). *Journal*
1066 *of Great Lakes Research*. doi:10.1016/j.jglr.2018.01.008
- 1067 Timoshkin, O. A., D. P. Samsonov, M. Yamamoto, and others. 2016. Rapid ecological change in the
1068 coastal zone of Lake Baikal (East Siberia): Is the site of the world's greatest freshwater

- 1069 biodiversity in danger? *Journal of Great Lakes Research* **42**: 487–497.
- 1070 doi:10.1016/j.jglr.2016.02.011
- 1071 Tong, Y., M. Wang, J. Peñuelas, and others. 2020. Improvement in municipal wastewater treatment alters
1072 lake nitrogen to phosphorus ratios in populated regions. *Proc Natl Acad Sci USA* **117**: 11566–
1073 11572. doi:10.1073/pnas.1920759117
- 1074 Tornqvist, R., J. Jarsjo, J. Pietron, A. Bring, P. Rogberg, S. M. Asokan, and G. Destouni. 2014. Evolution
1075 of the hydro-climate system in the Lake Baikal basin. *Journal of Hydrology* **519**: 1953–1962.
1076 doi:10.1016/j.jhydrol.2014.09.074
- 1077 Turetsky, M. R., R. K. Wieder, C. J. Williams, and D. H. Vitt. 2000. Organic matter accumulation, peat
1078 chemistry, and permafrost melting in peatlands of boreal Alberta. *Écoscience* **7**: 115–122.
1079 doi:10.1080/11956860.2000.11682608
- 1080 Veloza, A. J., F.-L. E. Chu, and K. W. Tang. 2006. Trophic modification of essential fatty acids by
1081 heterotrophic protists and its effects on the fatty acid composition of the copepod *Acartia tonsa*.
1082 *Marine Biology* **148**: 779–788. doi:10.1007/s00227-005-0123-1
- 1083 Volkova, E. A., N. A. Bondarenko, and O. A. Timoshkin. 2018. Morphotaxonomy, distribution and
1084 abundance of *Spirogyra* (Zygnematophyceae, Charophyta) in Lake Baikal, East Siberia.
1085 *Phycologia* **57**: 298–308. doi:10.2216/17-69.1
- 1086 de Vries, J. 1972. Soil Filtration of Wastewater Effluent and the Mechanism of Pore Clogging. *Journal*
1087 (Water Pollution Control Federation) **44**: 565–573.
- 1088 Wang, W., and J. Wang. 2018. Investigation of microplastics in aquatic environments: An overview of
1089 the methods used, from field sampling to laboratory analysis. *TrAC Trends in Analytical*
1090 *Chemistry* **108**: 195–202. doi:10.1016/j.trac.2018.08.026
- 1091 Wayland, M., and K. A. Hobson. 2001. Stable carbon, nitrogen, and sulfur isotope ratios in riparian food
1092 webs on rivers receiving sewage and pulp-mill effluents. *Can. J. Zool.* **79**: 5–15. doi:10.1139/z00-
1093 169

- 1094 Wickham, H., M. Averick, J. Bryan, and others. 2019. Welcome to the tidyverse. Journal of Open Source
1095 Software 4: 1686. doi:10.21105/joss.01686
- 1096 Wilke, C. O. 2019. cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2.”
- 1097 Wilke, C. O. 2020. ggtext: Improved Text Rendering Support for “ggplot2.”
- 1098 Yang, Y., W. Song, H. Lin, W. Wang, L. Du, and W. Xing. 2018. Antibiotics and antibiotic resistance
1099 genes in global lakes: A review and meta-analysis. Environment International 116: 60–73.
1100 doi:10.1016/j.envint.2018.04.011
- 1101 Yang, Y.-Y., G. S. Toor, P. C. Wilson, and C. F. Williams. 2016. Septic systems as hot-spots of
1102 pollutants in the environment: Fate and mass balance of micropollutants in septic drainfields.
1103 Science of The Total Environment 566–567: 1535–1544. doi:10.1016/j.scitotenv.2016.06.043
- 1104 York, J. K., G. Tomasky, I. Valiela, and D. J. Repeta. 2007. Stable isotopic detection of ammonium and
1105 nitrate assimilation by phytoplankton in the Waquoit Bay estuarine system. Limnology and
1106 Oceanography 52: 144–155. doi:10.4319/lo.2007.52.1.0144
- 1107 Yoshida, T., T. Sekino, M. Genkai-Kato, and others. 2003. Seasonal dynamics of primary production in
1108 the pelagic zone of southern Lake Baikal. Limnology 4: 53–62. doi:10.1007/s10201-002-0089-3
- 1109 Yoshii, K. 1999. Stable isotope analyses of benthic organisms in Lake Baikal. Hydrobiologia 411: 145–
1110 159.
- 1111 2016a. Methods for determination of nitrogen-containing matters (with corrections) (Методы
1112 определения азотсодержащих веществ (с Поправками)).
- 1113 2016b. Methods for determination of phosphorus-containing matters (with corrections) (Методы
1114 определения фосфорсодержащих веществ).
- 1115 2017. Nitrate concentration in waters: Photometric methods with Giress reagent following stabilization in
1116 a cadmium reducer (Массовая концентрация нитратного азота в водах: Методика измерений
1117 фотометрическим методом с реагентом Грисса после восстановления в кадмием
1118 редукторе).
- 1119

1120 Acknowledgments

1121 We would like to thank the faculty, students, staff, and mariners of the Irkutsk State University's
1122 Biological Research Institute Biostation for their expert field, taxonomic, and laboratory support;
1123 Marianne Moore and Bart De Stasio for helpful advice; the researchers and students of the
1124 Siberian Branch of the Russian Academy of Sciences Limnological Institute for expert
1125 taxonomic and logistical assistance; Oleg A. Timoshkin, Tatiana Ya. Sitnikova, Irina V.
1126 Mekhanikova, Nina A. Bonderenko, Ekaterina Volkova, Yulia Zvereva, Vadim V. Takhteev,
1127 Stephanie G. Labou, Stephen L. Katz, Brian P. Lanouette, John R. Loffredo, Alli N. Cramer,
1128 Alexander K. Fremier, Erica J. Crespi, Stephen M. Powers, Daniel L. Preston, Gavin L.
1129 Simpson, and James J. Elser for offering insights throughout the development of this project.
1130 Funding was provided by the National Science Foundation (NSF-DEB-1136637) to S.E.H., a
1131 Fulbright Fellowship to M.F.M., a NSF Graduate Research Fellowship to M.F.M. (NSF-DGE-
1132 1347973), and the Russian Ministry of Science and Education (N FZZE-2020-0026; FZZE-2020-
1133 0023). This work serves as one chapter of M.F.M.'s doctoral dissertation in Environmental and
1134 Natural Resource Sciences at Washington State University.
1135

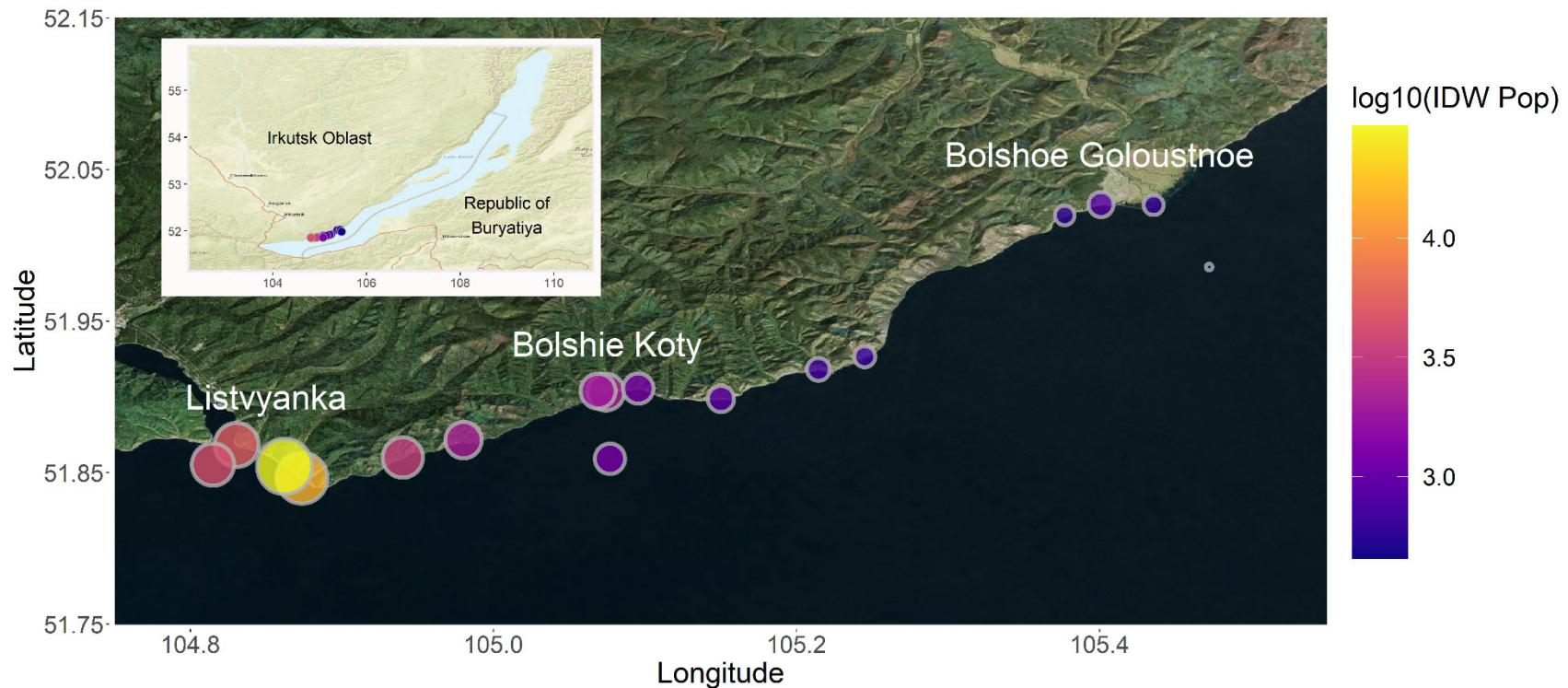
1136

Table 1: Location, depth, temperature and population information for each of the 17 sampling stations. “OS” refers to pelagic locations (i.e., “Offshore”), whereas other site abbreviations refer to littoral sampling locations.

Site	Latitude	Longitude	Depth (m)	Distance to shore (m)	Air Temperature (°C)	Surface Temperature (°C)	Adjacent Population
BK-1	51.90316	105.07404	0.7	10	18	14	80
BK-2	51.90365	105.069	0.9	17.5	19	13	80
BK-3	51.90536	105.0957	0.8	10	18	14	80
BGO-1	52.02693	105.40102	0.9	18	20	13	0
BGO-2	52.0197	105.37707	1.1	14	19	14	600
BGO-3	52.02649	105.43577	0.7	21	18	16	600
OS-1	51.98559	105.47237	900	NA	15	NA	NA
KD-1	51.92646	105.24504	0.8	20.75	23	NA	0
KD-2	51.91807	105.21456	0.9	14.5	23	16	0
MS-1	51.89863	105.15017	0.6	10.5	21	17	0
SM-1	51.87152	104.98006	0.9	11.5	21	15	0
LI-1	51.86825	104.83042	0.6	8.9	19	14	2000

LI-2	51.84626	104.87356	0.8	9.4	21	15	2000
LI-3	51.85407	104.86216	0.7	9.25	19.5	15	2000
EM-1	51.86005	104.93999	0.7	15.5	24.5	14	0
OS-2	51.8553	104.8148	1300	NA	21	NA	NA
OS-3	51.859108	105.0769	1400	5000	NA	14.5	NA

1137



1138

1139 Figure 1: Map of all sampling locations with sites sized and colored by log-transformed IDW population. IDW population was log-
1140 transformed so as to make IDW populations across three orders of magnitude more comparable. The entire transect included three
1141 developed sites (i.e., Listvyanka, Bolshie Koty, Bolshoe Goloustnoe). Three offshore samples were also collected to compare pelagic
1142 sewage signals to those in the littoral. Sampling locations west of Listvyanka are located farther from Listvyanka's centroid, and
1143 therefore have lower IDW population values than sites located closer to the centroid. This map was created using the R statistical

1144 environment (R Core Team 2019) and the tidyverse (Wickham et al. 2019), OpenStreetMap (Fellows and Stotz 2019), ggpubr
1145 (Kassambara 2019), cowplot (Wilke 2019), and ggrepel (Slowikowski 2019) packages.
1146

For Review Only

1147

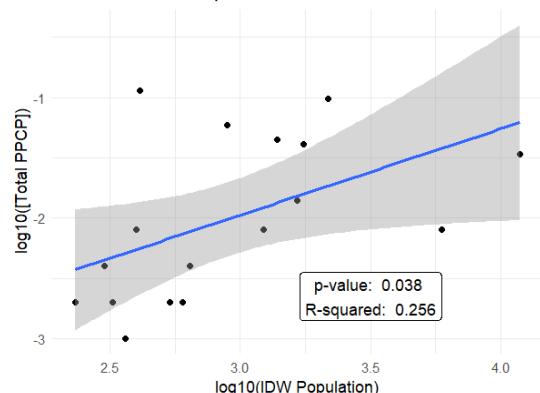
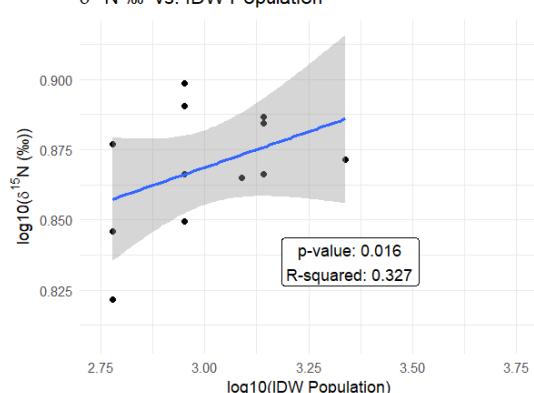
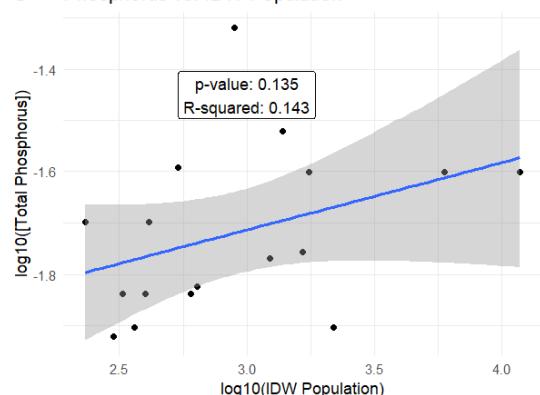
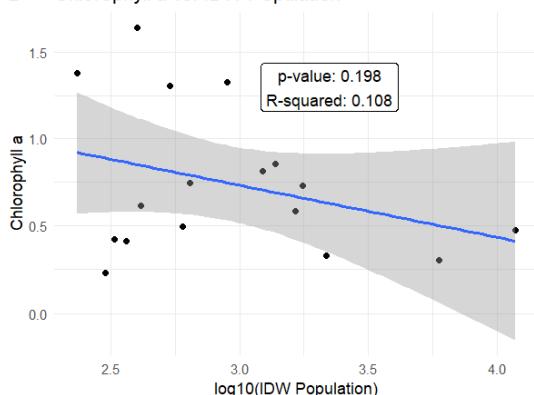
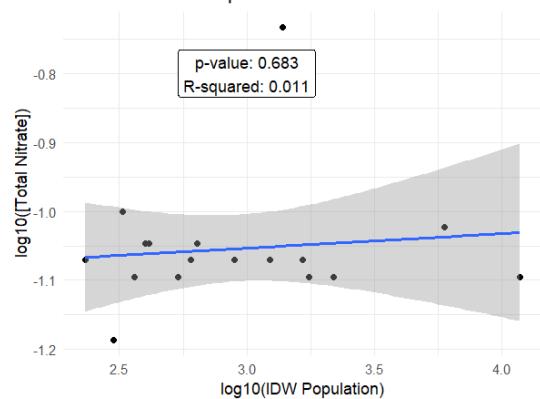
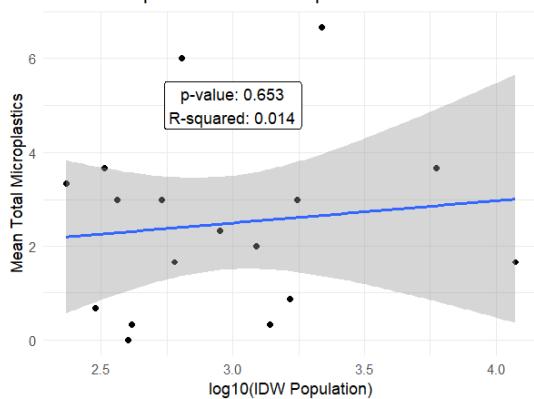
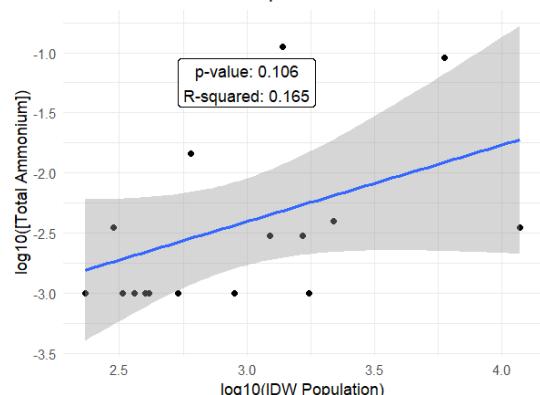
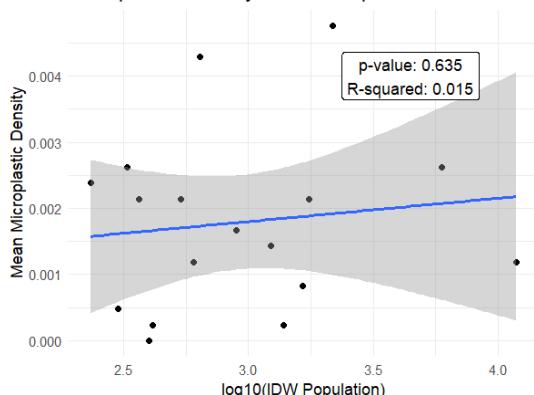


1148

1149 Figure 2: Photographs and Google Earth imagery of each developed area. Photographs were
1150 taken by Kara H. Woo and Michael F. Meyer.

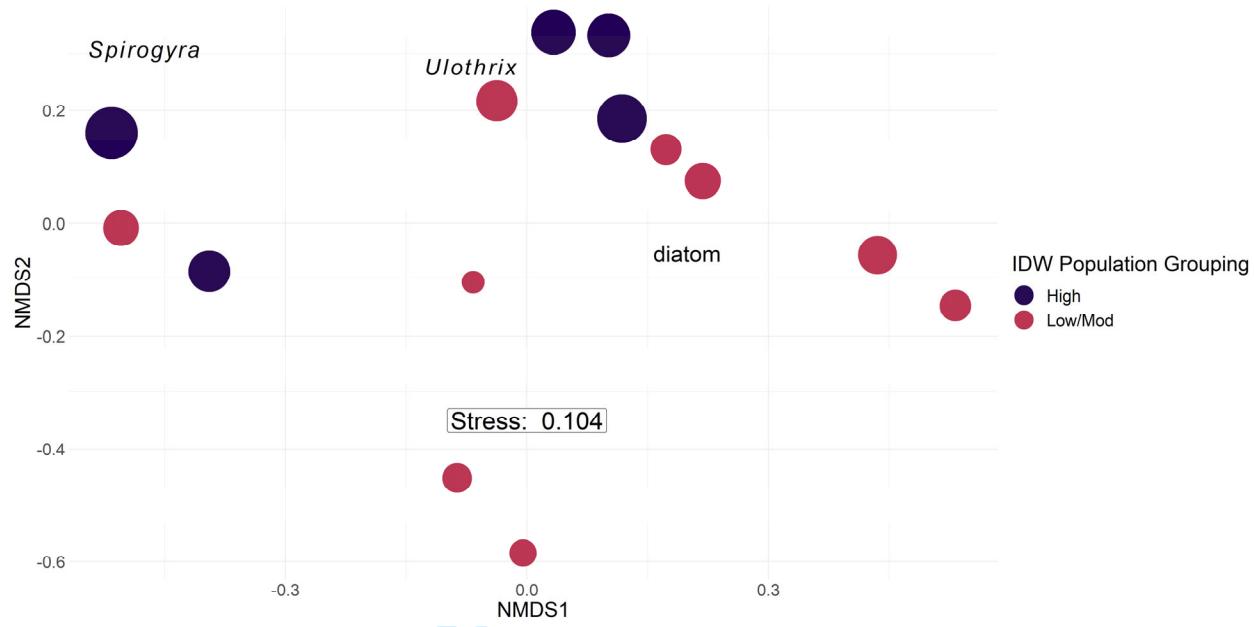
Table 2: Average sewage indicator concentrations and densities per sampling location. Caffeine, acetaminophen/paracetamol, paraxanthine, and cotinine detection limits are estimated to be 0.001 µg/L based on a 500 mL sample volume.

Site	NH ₄ ⁺ (mg/L)	NO ₃ ⁻ (mg/L)	Total Phosphorus (mg/L)	Caffeine (µg/L)	Acetaminophen (µg/L)	Paraxanthine (µg/L)	Cotinine (µg/L)	Fragment density (MPs/L)	Fiber density (MPs/L)	Bead density (MPs/L)	IDW population	Categorical IDW population
BK-1	0.003	0.085	0.054	0.011	0.001	0.002	0	0	0.000833	0	2304.039	High
BK-2	0.003	0.085	0.052	0.007	0.001	0	0	0.000952	0.000476	0	1891.558	Mod/Low
BK-3	0.068	0.09	0.045	0.003	0.001	0	0	0.003095	0.00119	0	1231.234	Mod/Low
BGO-1	0.0145	0.085	0.044	0	0.002	0	0	0.00119	0	0	838.5385	Mod/Low
BGO-2	0.001	0.08	0.0385	0	0.001	0	0	0.000238	0.001905	0	611.91	Mod/Low
BGO-3	0.001	0.09	0.044	0.005	0.003	0	0	0	0	0	624.455	Mod/Low
OS-1	0.001	0.085	0.061	0	0.001	0	0.001	0.002381	0	0	455.7733	Mod/Low
KD-1	0.0035	0.065	0.0375	0.003	0.001	0	0	0	0.000476	0	662.4151	Mod/Low
KD-2	0.001	0.1	0.0445	0.001	0.001	0	0	0.000714	0.001905	0	720.5484	Mod/Low
MS-1	0.001	0.09	0.061	0.064	0.035	0.015	0	0	0.000238	0	903.6733	Mod/Low
SM-1	0.001	0.085	0.1475	0.042	0.012	0.005	0	0	0.001667	0	2146.218	Mod/Low
LI-1	0.004	0.08	0.0385	0.05	0.04	0.006	0.002	0.00381	0.000238	0.000714	5403.209	High
LI-2	0.091	0.095	0.0775	0.001	0.007	0	0	0.001429	0.00119	0	14792.51	High
LI-3	0.0035	0.08	0.077	0.027	0.002	0.002	0.003	0.000476	0	0.000714	29511.73	High
EM-1	0.1125	0.185	0.092	0.029	0.014	0.002	0	0	0.000238	0	3389.949	High
OS-2	0.001	0.08	0.078	0.033	0.001	0.004	0.003	0.000238	0.001905	0	4340	High
OS-3	0.001	0.08	0.0795	0.001	0.001	0	0	0	0.002143	0	1221.424	Mod/Low

A PPCP vs. IDW Population**B** $\delta^{15}\text{N} \text{‰}$ vs. IDW Population**C** Phosphorus vs. IDW Population**D** Chlorophyll a vs. IDW Population**E** Nitrate vs. IDW Population**F** Total Microplastics vs. IDW Population**G** Ammonium vs. IDW Population**H** Microplastics Density vs. IDW Population

1153 Figure 3: Linear models of total PPCP concentrations (A), macroinvertebrate $\delta^{15}\text{N}$ (B), total
1154 phosphorus (C), chlorophyll a (D), nitrate (E), total microplastics (F), ammonium (G), and
1155 microplastic density (H) regressed against log-transformed inverse distance weighted (IDW)
1156 population. Total PPCP concentrations (A) and macroinvertebrate $\delta^{15}\text{N}$ (B) produced significant
1157 models. Total phosphorus (C), chlorophyll a (D), nitrate (E), total microplastics (F), ammonium
1158 (G), and microplastic density (H) did not produce significant models.

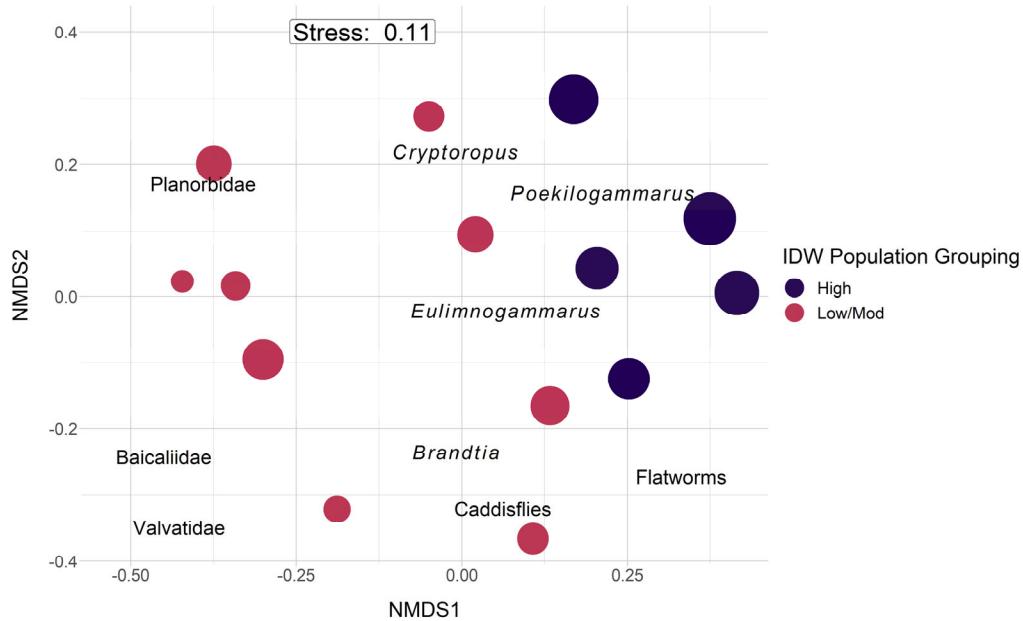
For Review Only

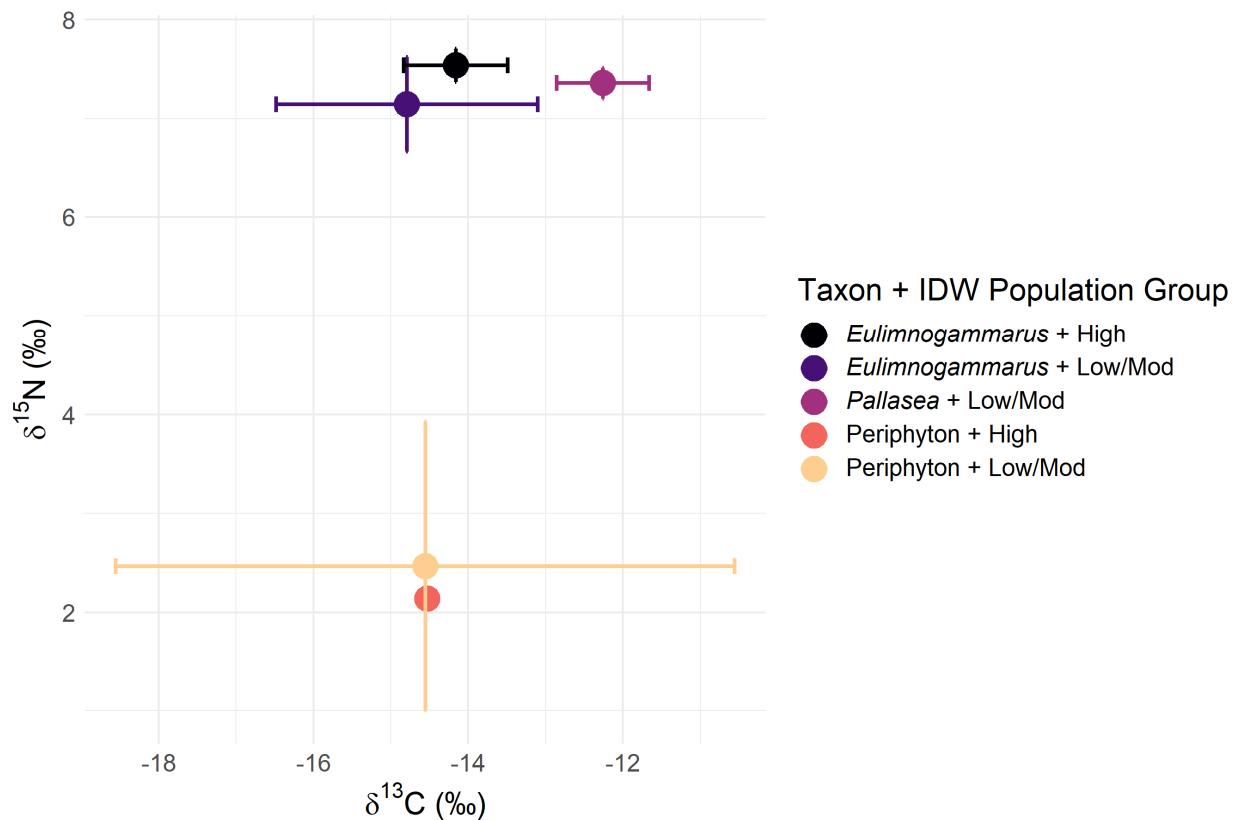


1159

1160 Figure 4: Periphyton abundance NMDS with Bray-Curtis dissimilarity. Points are sized by log10
1161 IDW population and colored by grouped IDW population values. Taxonomic labels represent
1162 species scores, which are weighted averages of species contributions from site scores. For
1163 periphyton, PERMANOVA ($p = 0.001$) and post-hoc SIMPER results suggested sites with a
1164 higher IDW population value tended to be more associated with filamentous algal groupings and
1165 separate from sites with moderate and low IDW population values, which were more associated
1166 with diatom abundance.

1167





1177

1178

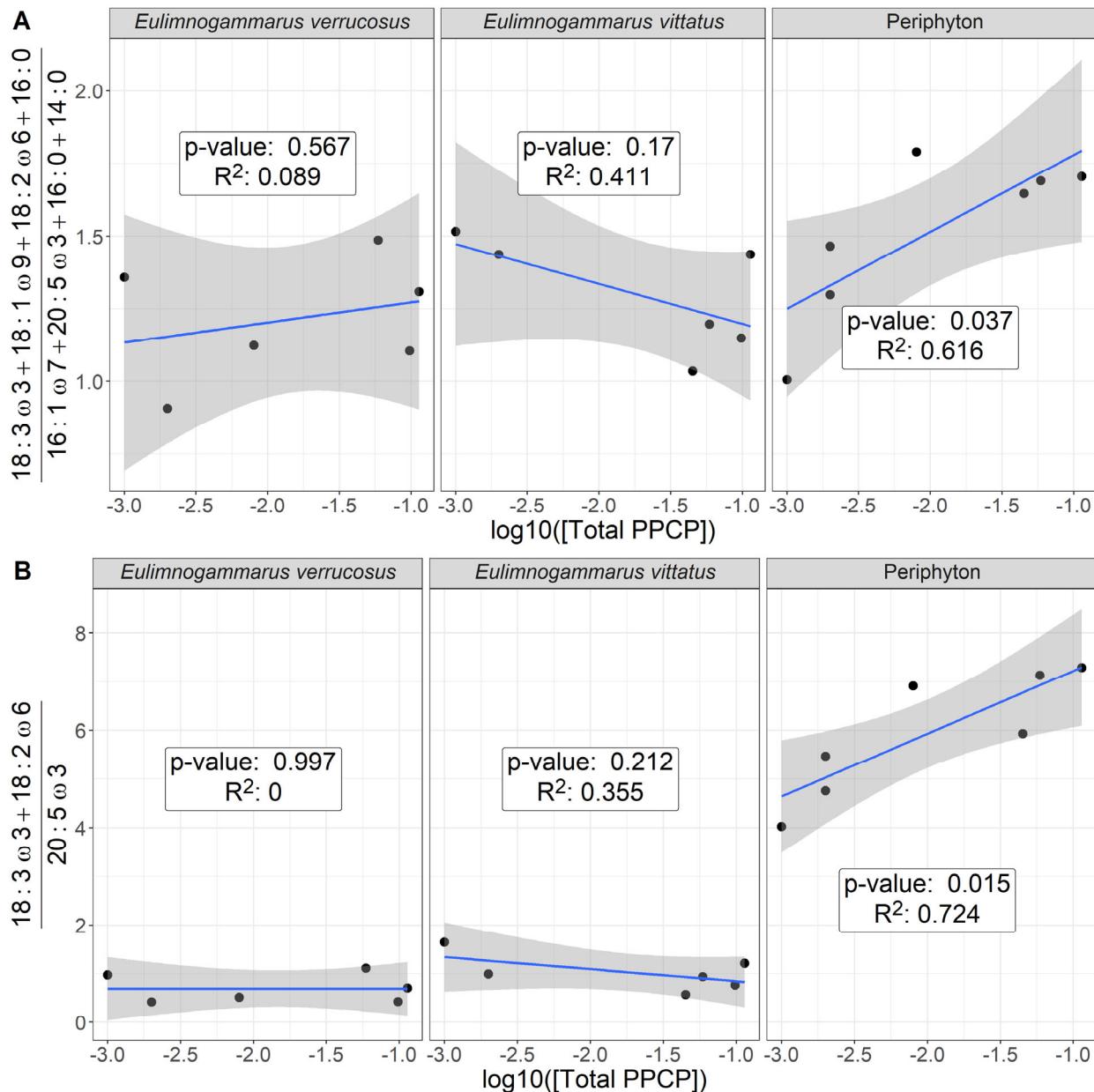
1179 Figure 6: Biplot of mean and standard deviation $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ stable isotope values for littoral
 1180 amphipods and periphyton, grouped by categorical IDW population (Table 3). In general,
 1181 periphyton did not differ in $\delta^{13}\text{C}$ or $\delta^{15}\text{N}$ signatures with increasing IDW population, whereas
 1182 *Eulimnogammarus* amphipods increased in $\delta^{15}\text{N}$ signatures with increasing IDW population.
 1183 *Pallasea* signatures differed from *Eulimnogammarus* most likely because *Pallasea* tends to
 1184 remain in the nearshore area, whereas *Eulimnogammarus* will regularly migrate to deeper zones
 1185 (Takhteev & Didorenko, 2015).

1186

Table 3: Mean inter-site fatty acid proportion of each taxon and fatty acid grouping (as defined in table S2).

Taxon	Number of sites	Branched	LCPUFA	MUFA	SAFA	SCPUFA
<i>Draparnaldia</i> spp.	4	0.000	0.012	0.088	0.189	0.710
<i>Eulimnogammarus cyaneus</i>	2	0.002	0.259	0.309	0.248	0.182
<i>Eulimnogammarus verrucosus</i>	6	0.000	0.188	0.385	0.240	0.187
<i>Eulimnogammarus vittatus</i>	6	0.001	0.171	0.371	0.241	0.216
<i>Pallasea cancellus</i>	3	0.001	0.282	0.359	0.187	0.171
Periphyton	7	0.000	0.073	0.092	0.284	0.550
Snail	3	0.002	0.470	0.123	0.194	0.211

1188



1189

1190 Figure 7: Ratio of filamentous:diatom-associated fatty acids (A) and essential fatty acids (B)
 1191 across our PPCP gradient. Our first analysis (A) focused solely on green filamentous algal fatty
 1192 acids (i.e., 18:3ω3, 18:1ω9, 18:2ω6, and 16:0 relative to diatom fatty acids (i.e., 20:5ω3, 16:1ω7,
 1193 16:0, 14:0) in relation to increasing PPCP concentrations. This first analysis suggested
 1194 periphyton reflected an increasing green, filamentous signature relative to diatoms, which

1195 corroborates analyses showing community compositional shifts (Figure 4). While periphyton
1196 fatty acids changed significantly across our sewage gradient, macroinvertebrate signatures
1197 remained consistent. Our second analysis (B) focused solely on the essential fatty acids, which
1198 further highlights the trends observed in periphyton and macroinvertebrate grazers.

For Review Only

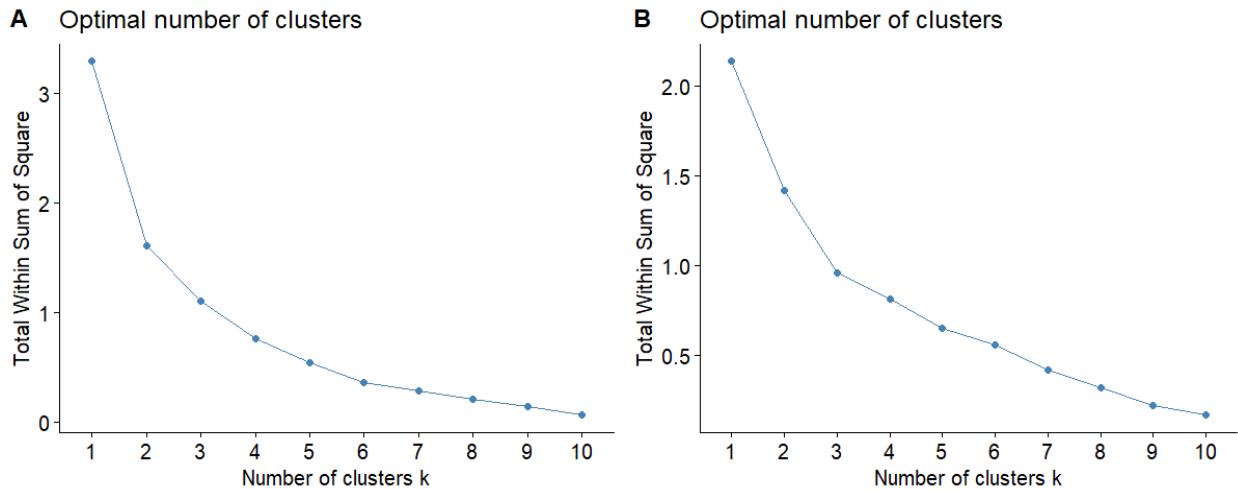
Table S1: Macroinvertebrate taxonomic groupings for abundance estimates. Amphipod taxa were defined as in Takhteev & Didorenko, 2015; mollusk taxa were defined as in Sitnikova, 2012.

Amphipoda	Mollusca	Other
<i>Brandtia latissima intermida</i> (Dorogostaiskii 1930)	Acroloxidae	Asellidae
<i>Brandtia latissima lata</i> (Dybowsky 1874)	Baicaliidae	Caddisflies
<i>Brandtia latissima latior</i> (Dybowsky 1874)	Benedictidate	Hirudinea
<i>Brandtia latissima latissima</i> (Gerstfeldt 1858)	Maackia	Planaria
<i>Brandtia parasitica parasitica</i> (Dybowsky 1874)	Planorbidae	
<i>Cryptoropus inflatus</i> (Dybowsky 1874)	Valvatidae	
<i>Cryptoropus pachytus</i> (Dybowsky 1874)		
<i>Cryptoropus rugosus</i> (Dybowsky 1874)		
<i>Eulimnogammarus capreolus</i> (Dybowsky 1874)		
<i>Eulimnogammarus cruentes</i> (Dorogostaiskii 1930)		
<i>Eulimnogammarus cyaneus</i> (Dybowsky 1874)		
<i>Eulimnogammarus grandimanus</i> (Bazikalova 1945)		
<i>Eulimnogammarus maacki</i> (Gerstfeldt 1858)		
<i>Eulimnogammarus maritui</i> (Bazikalova 1945)		
<i>Eulimnogammarus verucossus</i> (Gerstfeldt 1858)		
<i>Eulimnogammarus viridis viridis</i> (Dybowsky 1874)		
<i>Eulimnogammarus vittatus</i> (Dybowsky 1874)		
<i>Pallasea brandtia brandita</i> (Dybowsky 1874)		
<i>Pallasea brandtii tenera</i> (Sovinskii 1930)		

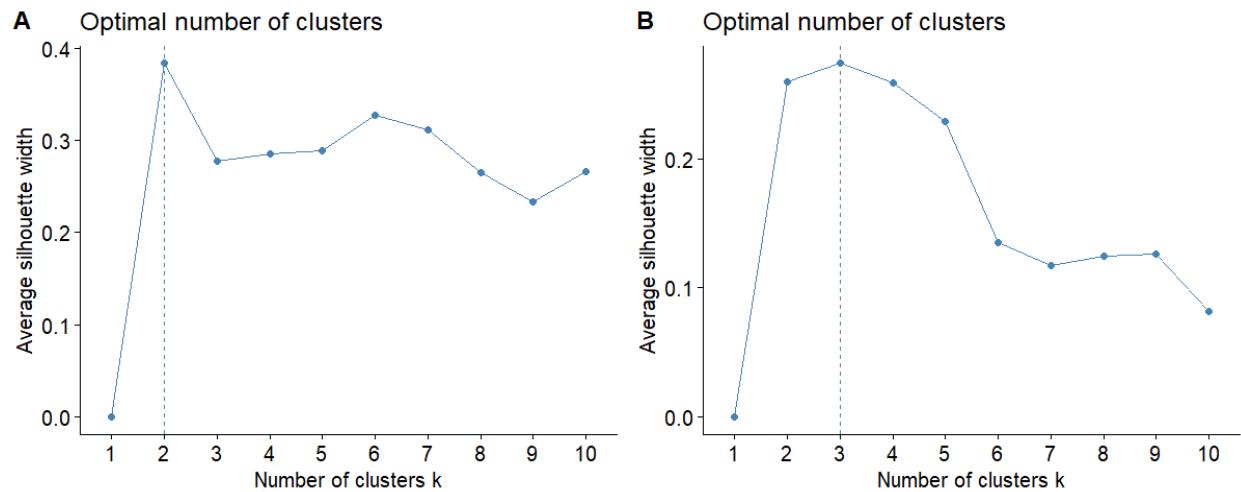
<i>Pallasea cancelloides</i> (Gerstfeldt 1858)		
<i>Pallasea cancellus</i> (Pallas 1776)		
<i>Pallasea viridis</i> (Garjajev 1901)		
<i>Poekilogammarus crassimus</i> (Sovinskii 1915)		
<i>Poekilogammarus ephippiatus</i> (Dybowsky 1874)		
<i>Poekilogammarus megonychus perpolitus</i> (Takhteev 2002)		
<i>Poekilogammarus pictus</i> (Dybowsky 1874)		

1199

For Review Only



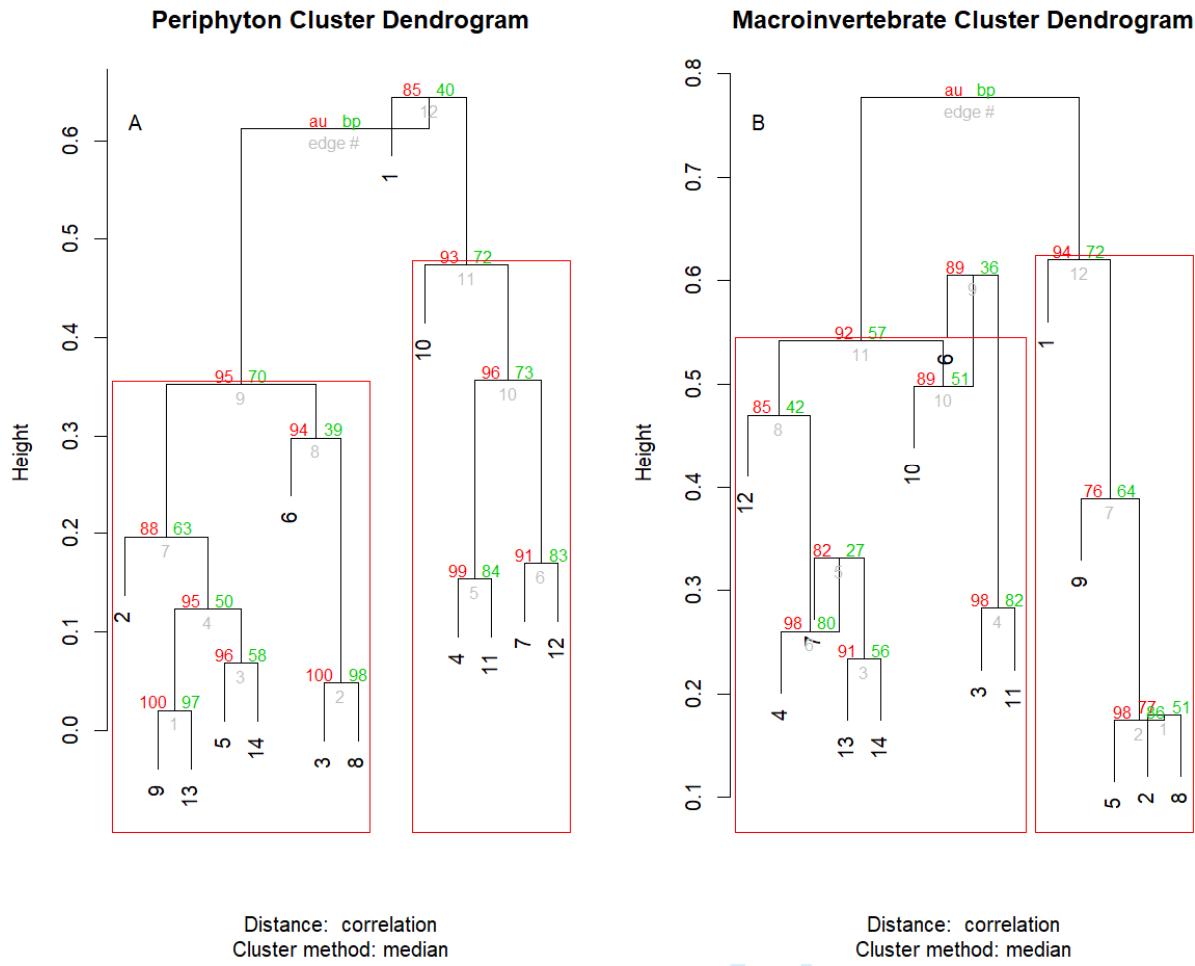
1200
1201
1202 Figure S1: With-group-sum-of-squares (wss) for increasing number of k-mediod clusters for
1203 periphyton (A) and invertebrate (B) community data. In the case of periphyton data, wss
1204 decreases most markedly with three clusters, whereas invertebrate community abundance is best
1205 described by potential two or three clusters.
1206



1207

1208 Figure S2: Average silhouette width for increasing number of k-mediod clusters for periphyton
1209 (A) and invertebrate (B) community data. In the case of periphyton data, average silhouette
1210 width decreases most markedly with three clusters, whereas invertebrate community abundance
1211 is best described by two or three clusters as the average silhouette width for both two and three
1212 clusters was highest before beginning to decrease.

1213

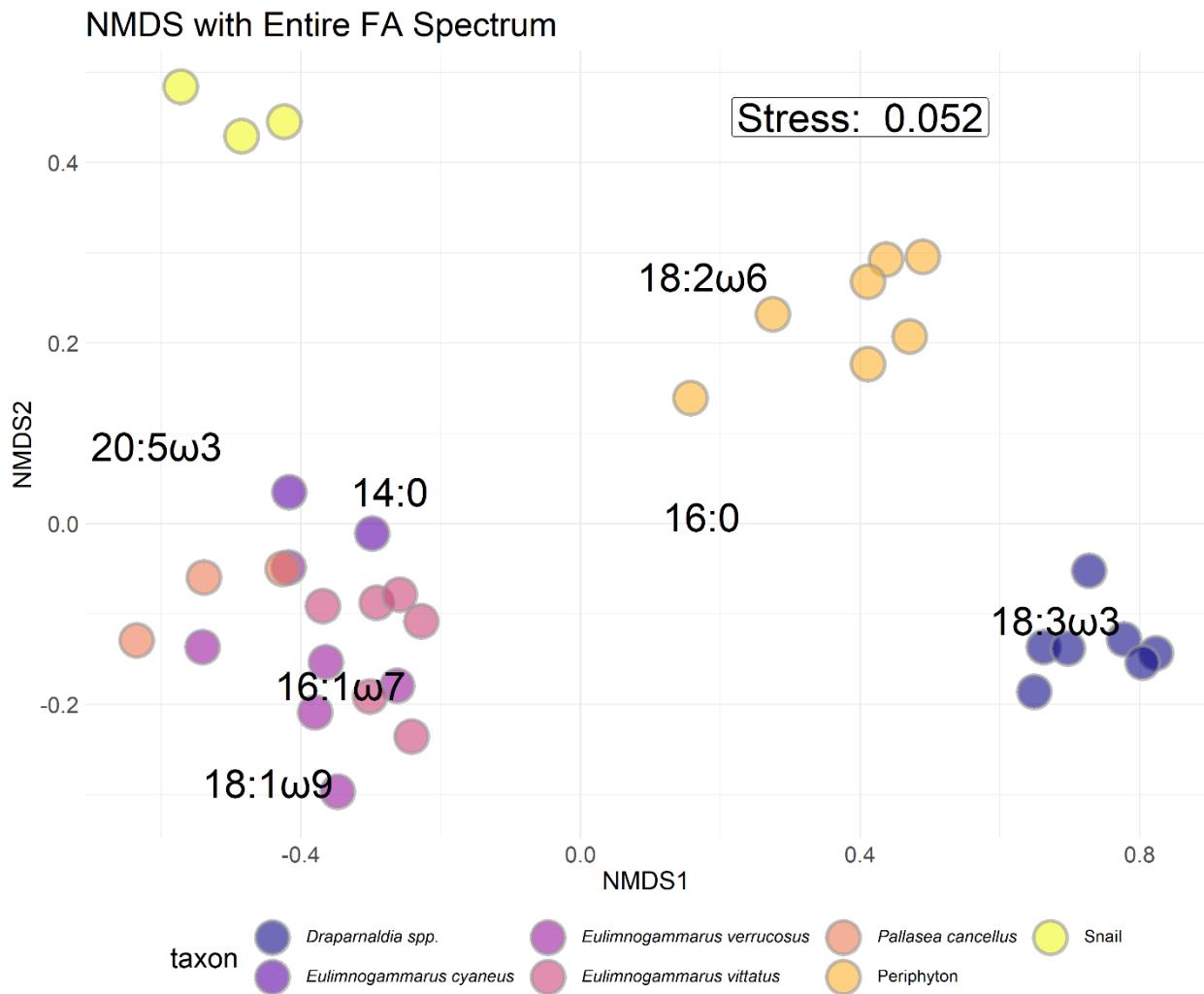


1214

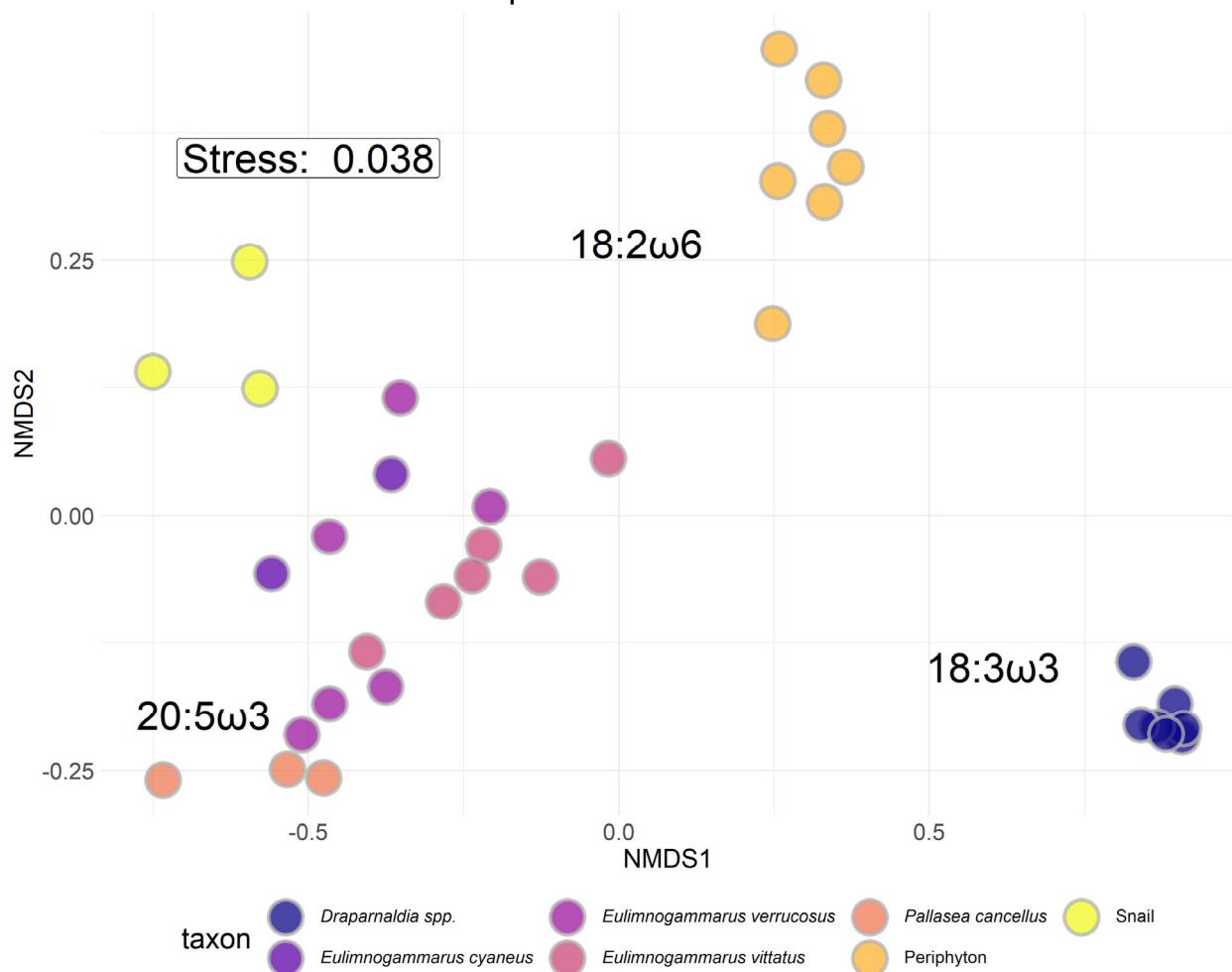
1215 Figure S3: Weighted Pair-Group Centroid Clustering (WPGMC) for periphyton (A) and
1216 macroinvertebrate (B) community compositions. Approximately unbiased (au) p-values are
1217 computed by multiscale bootstrap resampling, and displayed in red on the left side of each node.
1218 Bootstrapped probabilities (bp) are displayed in green on the right side of each node. Unlike k-
1219 medioids, WPGMC uses a hierarchical approach to assign clusters, which are bootstrapped in
1220 order to generate a probability of group membership. This technique suggested that both
1221 periphyton and macroinvertebrates could be grouped in two clusters. Grouping
1222 macroinvertebrates into three clusters was possible; however, three clusters resulted in 8 of the

- 1223 14 sampling locations being assigned to a group. In contrast, two groups enabled 13 of the 14
1224 sampling locations to be assigned to a cluster.

For Review Only



NMDS with Essential FA Spectrum



1235

1236 Figure S3: NMDS with Bray-Curtis dissimilarity of proportional biologically essential fatty acid
 1237 compositions for each macroinvertebrate and primary producer collected. *Eulimnogammarus* and
 1238 *Pallasea* are endemic amphipod genera. *Draparnaldia* spp. are endemic filamentous algae that
 1239 are large and form very dense mats easily collected where it occurs. *Draparnaldia* spp. occurred
 1240 in large, visible colonies, allowing us to sample and analyze just the *Draparnaldia* spp. fatty
 1241 acids. Because *Draparnaldia* spp. fatty acids were dominated by 18:3 ω 3 more so than
 1242 periphyton, they formed their own cluster. Snails were not identified to species prior to fatty acid
 1243 analysis. Interspecific variation in fatty acid composition tended to be larger than intraspecific

1244 variation, implying that fatty acid signatures were largely species-specific and not
1245 environmentally driven.

1246

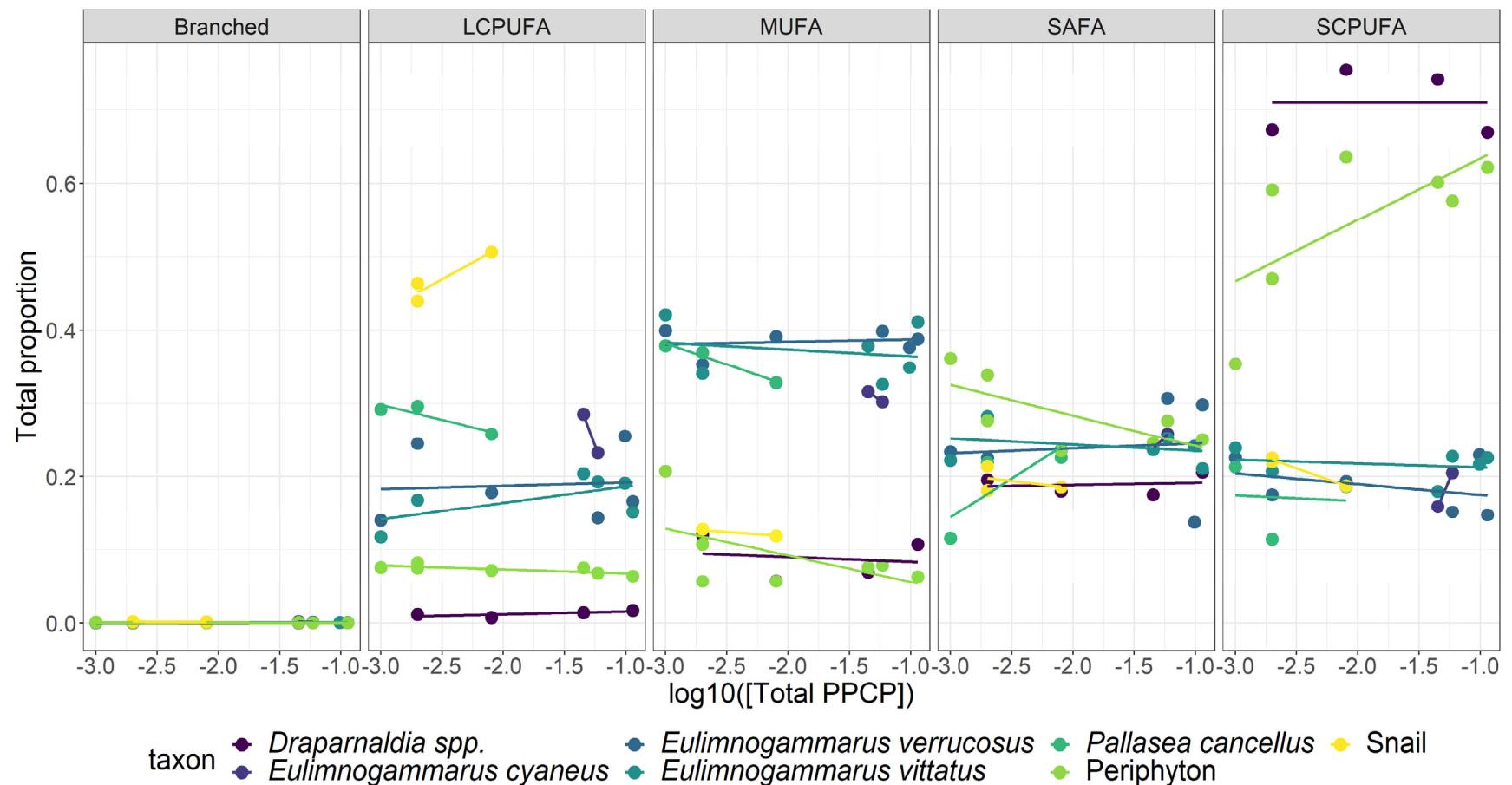
For Review Only

Table S2: Fatty acid groupings used in this analysis

Fatty Acid Group	Fatty acids considered
Branched	a-15:0, i-15:0, a-17:0, i-17:0
SAFA	12:0, 14:0, 15:0, 16:0, 17:0, 18:0, 20:0, 22:0, 24:0
MUFA	14:1 ω 5, 15:1 ω 7, 17:1n7, 16:1 ω 5, 16:1 ω 6, 16:1 ω 7, 16:1 ω 8, 16:1 ω 9, 18:1 ω 7, 18:1 ω 9, 20:1 ω 7, 20:1 ω 9, 22:1 ω 7, 22:1 ω 9
SCPUFA	16:2 ω 4, 16:2 ω 6, 16:2 ω 7, 16:3 ω 3, 16:3 ω 4, 16:3 ω 6, 16:4 ω 1, 16:4 ω 3, 18:2 ω 6, 18:2 ω 6t, 18:3 ω 3, 18:3 ω 6, 18:4 ω 3, 18:4 ω 4, 18:5 ω 3
LCPUFA	20:2 ω 5(11), 20:2 ω 5(13), 20:2 ω 6, 20:3 ω 3, 20:3 ω 6, 20:4 ω 3, 20:4 ω 6, 20:5 ω 3, 22:2 ω 6, 22:3 ω 3, 22:4 ω 3, 22:4 ω 6, 22:5 ω 3, 22:5 ω 6, 22:6 ω 3

1247

1248



1249

1250 Figure S4: Proportions of major fatty acid groups (as defined in Table S2) across the sewage gradient. Primary producers
 1251 (*Draparnaldia* spp. and periphyton) were largely characterized by SCPUFAs, amphipods were largely associated with high MUFA
 1252 abundance, and snails were generally characterized with high LCPUFA abundance. Across the sewage gradient, periphyton SCPUFA

1253 tended to increase, which lead to more targeted analyses on which specific fatty acids were increasing. In contrast to periphyton, all

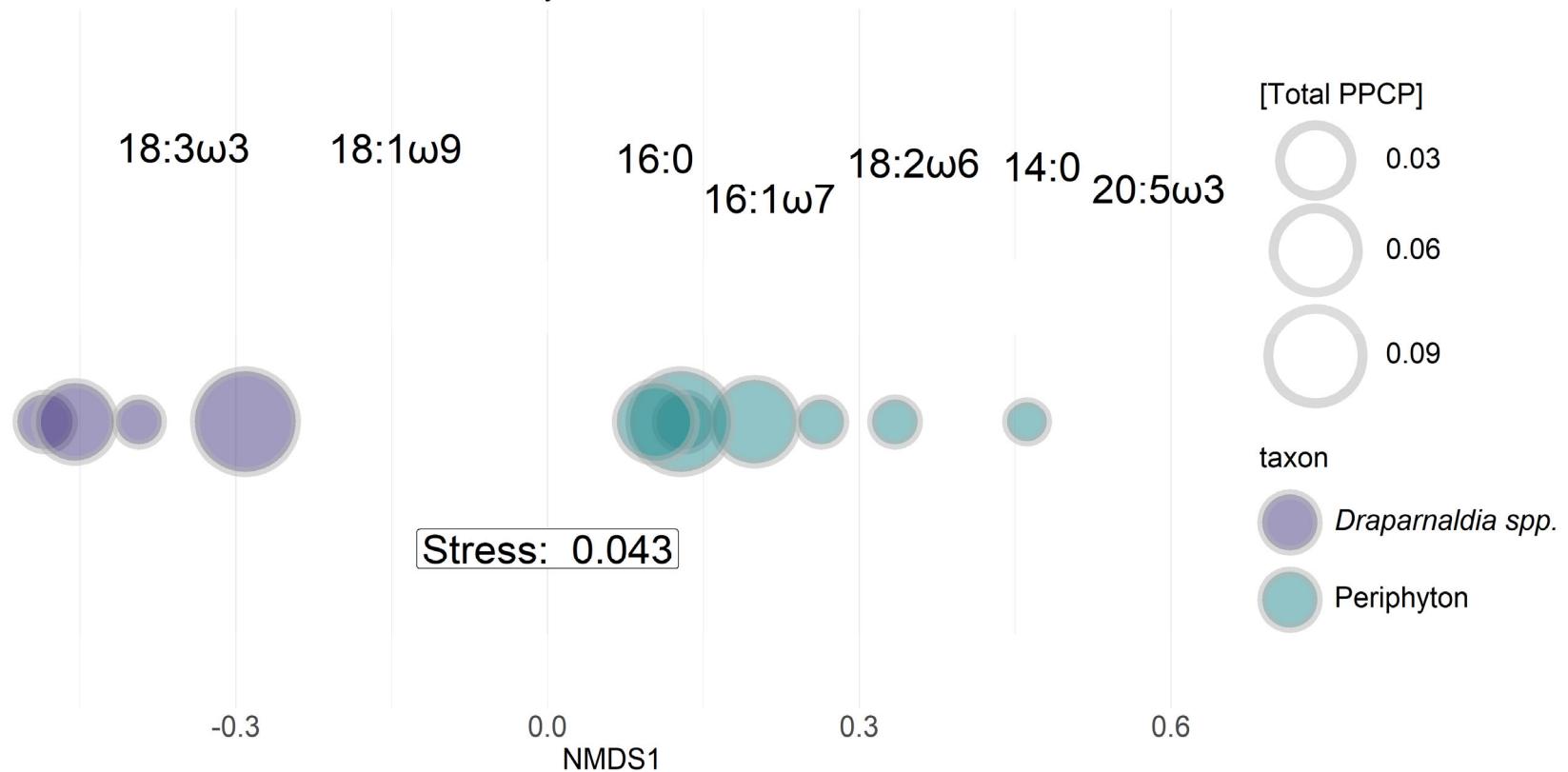
1254 other taxa remained consistent with respect to fatty acid proportions across the sewage gradient.

1255

1256

For Review Only

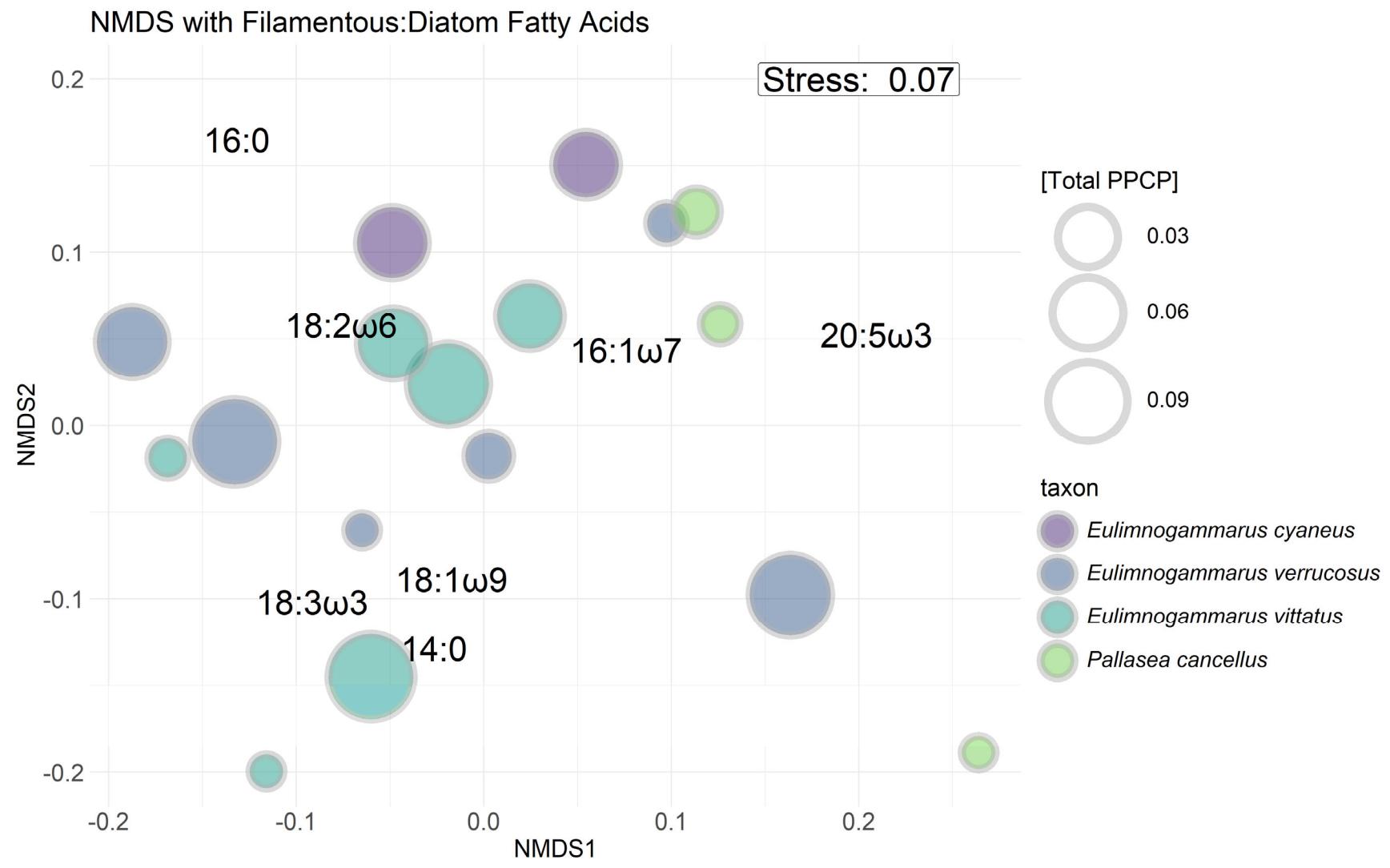
NMDS with Filamentous:Diatom Fatty Acids



1257

1258 Figure S5: One-dimensional NMDS with Bray-Curtis similarity of seven targeted fatty acids of interest for primary producers. Fatty
1259 acid scores are placed above shapes. Shapes are sized by total PPCP concentration. Periphyton (blue-green) tend to increase in size
1260 from right-to-left, suggesting that periphyton tend to include more 18:3 ω 3 and 18:1 ω 9 (indicators of green algal taxa) with an
1261 increasing sewage signal. In contrast, *Draparnaldia* spp. (purple) fatty acids tend to remain consistent across the sewage gradient.

1262



1264 Figure S6: NMDS with Bray-Curtis similarity of seven targeted fatty acids of interest for primary producers. Points are sized by total
1265 PPCP concentration. Visually, there appears to be no distinct separation among or within taxa unlike was observed with periphyton
1266 (Figure S5).

For Review Only

1267

For Review Only