



1 Hyporheic Zone Respiration is Jointly Constrained by Organic

2 Carbon Concentration and Molecular Richness

3 James C. Stegen¹, Vanessa A. Garayburu-Caruso¹, Robert E. Danczak¹, Amy E. Goldman², Lupita
4 Renteria¹, Joshua M. Torgeson², and Jacqueline Hager²

5 ¹Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, U.S.A

6 ²Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, WA, U.S.A

7 *Correspondence to:* James C. Stegen (James.Stegen@pnnl.gov)

8 **Abstract.**

9 River corridors are fundamental components of the Earth system, and their biogeochemistry can be heavily influenced
10 by processes in subsurface zones immediately below the riverbed, referred to as the hyporheic zone. Within the
11 hyporheic zone, organic matter (OM) fuels microbial respiration, and OM chemistry heavily influences aerobic and
12 anaerobic biogeochemical processes. The link between OM chemistry and respiration has been hypothesized to be
13 mediated by OM molecular diversity, whereby respiration is predicted to decrease with increasing diversity. Here we
14 test the specific prediction that aerobic respiration rates will decrease with increases in the number of unique organic
15 molecules (i.e., OM molecular richness, as a measure of diversity). We use publicly available data across the United
16 States from crowdsourced samples taken by the Worldwide Hydrobiogeochemical Observation Network for Dynamic
17 River Systems (WHONDRS) consortium. Our continental-scale analyses rejected the hypothesis of a direct limitation
18 of respiration by OM molecular richness. In turn, we found that organic carbon (OC) concentration imposes a primary
19 constraint over hyporheic zone respiration, with additional potential influences of OM richness. We specifically
20 observed respiration rates to decrease nonlinearly with the ratio of OM richness to OC concentration. This relationship
21 took the form of a constraint space with respiration rates in most systems falling below the constraint boundary. A
22 similar, but slightly weaker, constraint boundary was observed when relating respiration rate to the inverse of OC
23 concentration. These results indicate that maximum respiration rates may be governed primarily by OC concentration,
24 with secondary influences from OM richness. Our results also show that other variables often suppress respiration
25 rates below the maximum associated with the richness-to-concentration ratio. An important focus of future research
26 efforts will identify factors that suppress hyporheic zone respiration below the constraint boundaries observed here.

27 **1 Introduction**

28 River corridors are key components of the Earth system that connect terrestrial landscapes to the ocean through the
29 transport and transformation of organic matter (OM) and nutrients (Harvey and Gooseff, 2015; Schläpfer and
30 Schneider, 2000; Schlesinger and Melack, 1981). In addition, river corridors have strong connections to the
31 atmosphere in terms of significant emissions of greenhouse gasses such as CO₂ (Raymond et al., 2013). Within river
32 corridors the hyporheic zone (Orghidan, 2010) can have a dominant influence over net metabolism and



33 biogeochemical transformations (Boulton et al., 1998; Naegeli and Uehlinger, 1997; Krause et al., 2011) to a degree
34 that it can act as the “river’s liver” to remove contaminants (Fischer et al., 2005). Recent work has found that
35 detailed properties of OM chemistry can significantly influence respiration rates in hyporheic zone sediments
36 (Stegen et al., 2018; Garayburu-Caruso et al., 2020a; Sengupta et al., 2020; Graham et al., 2018, 2017; Song et al.,
37 2020a). These observations demonstrate a need to deepen understanding of the relationships between hyporheic
38 zone biogeochemistry (e.g., respiration rates) and OM chemistry.

39

40 A conceptual hypothesis was recently developed that may provide new insight into the connections between OM
41 chemistry and biogeochemical rates. More specifically, Lehmann et al. (2020) hypothesize that OM can be protected
42 from degradation (in part) by high levels of molecular diversity. Biogeochemical rates that depend on OM oxidation
43 (e.g., aerobic respiration) may therefore be suppressed with increases in the number of unique organic molecules
44 (referred to here as OM molecular richness). The concept is that high levels of OM molecular richness lead to low
45 returns-on-investment, relative to the energy invested in building and maintaining the molecular machinery needed
46 to metabolize any given type of organic molecule. The consequence is low respiration rates. The underlying
47 mechanism has been proposed to help protect OM in some ecosystems such as deep sea (Arrieta et al., 2015) and
48 river corridor (Stegen et al., 2018) environments.

49

50 The hypothesis of lower biogeochemical rates with higher OM molecular richness has not been evaluated in
51 hyporheic zone sediments despite the established connection between OM chemistry and hyporheic zone respiration
52 rates. We posit that higher levels of hydrologic connectivity in hyporheic zones relative to unsaturated systems (e.g.,
53 soil) may diminish influences of spatial isolation such as an OM stabilization mechanism (Schmidt et al., 2011),
54 potentially leading to particularly strong relationships between respiration rates and OM chemistry. In turn, it is
55 plausible that the hyporheic zone is an ecosystem in which we may find support for the hypothesized negative
56 relationship between respiration rates and OM molecular richness. Here we test this hypothesis at the continental
57 scale using publicly available data from the Worldwide Hydrobiogeochemical Observation Network for Dynamic
58 River Systems (WHONDRS) consortium (Stegen and Goldman, 2018; Garayburu-Caruso et al., 2020b; Toyoda et
59 al., 2020; Goldman et al., 2020).

60 **2 Methods**

61 *Sample collection and data generation*

62 During the summer of 2019, the WHONDRS consortium carried out a multi-continent river corridor study to
63 evaluate interactions between metabolomes, microbial metabolism, biogeochemical function, and ecosystem
64 features. Garayburu-Caruso et al. 2020b describe details on metadata, sample collection, analysis, and processing of
65 ultrahigh resolution mass spectrometry data. Briefly, during late July and August 2019 sediment samples were
66 collected across multiple continents, but the current study focuses on samples collected in the contiguous United
67 States (ConUS) (Fig. 1). Shallow sediments (~1-5 cm depth) were collected at three separate depositional zones at
68 each site. The zones were ~ 10 m away from each other and were labeled as upstream, midstream, and downstream.



69 Samples were shipped to the Pacific Northwest National Laboratory (PNNL) campus in Richland, WA (USA) on ice
70 within 24 hours of collection.

71

72 In the laboratory, sediments were sieved with a 2 mm sieve, and subsampled into 50 mL conical tubes (Genesee
73 Scientific Olympus™ Plastics) to separate Field and Incubation aliquots. Note that in the methods provided by
74 Garayburu-Caruso et al. (2020b) there is an error in the description of the sediment preservation prior to mass
75 spectrometry analysis. Corrected preservation methods are described immediately below. Sediments from the Field
76 aliquot were flash frozen in liquid nitrogen immediately after sieving to maintain the sediment characteristics
77 observed in the field and stored at -80°C until analysis. The Incubation aliquots were not flash frozen immediately;
78 instead they were kept in the dark inside an environmental chamber at 21°C along with other sediments to be used
79 for respiration measurements (see below) so that the two sets of sediment samples experienced the same conditions
80 leading up to the use of the sediment for respiration estimation. The next morning, Incubation aliquots were
81 retrieved from the environmental chamber, flash frozen in liquid nitrogen, and stored at -80°C until analysis. In our
82 analyses we used the “Field” sediments to study water-extractable organic carbon concentration and OM chemistry
83 prior to the respiration incubation. We used the “Incubation” sediments as a check for changes or variation in
84 organic carbon concentration between Field sediments and those sediments that were actually incubated. As a
85 quality assurance procedure (detailed below), we removed samples with the largest changes in organic carbon
86 concentration between Field and Incubation sediments.

87

88 Field and Incubation sediments were extracted with milli-Q water, and the resulting supernatant from sediment
89 extractions was filtered through a 0.22 µm sterivex filter (EMD Millipore). Non-purgeable organic carbon (NPOC)
90 was determined on the supernatant by a Shimadzu combustion carbon analyzer TOC-L CSH/CSN E100V with ASI-
91 L autosampler. We only included data from sites that had similar NPOC concentrations between the paired Field and
92 Incubation samples. Our rationale for this approach is based on the assumption that if NPOC is highly variable
93 across replicate sub-samples (i.e., across paired Field and Incubation samples), the associated sediments used for
94 respiration measurements may have been highly heterogeneous despite our efforts to homogenize sediments prior to
95 analyses. In turn, we assume that high heterogeneity may lead to unreliable estimates of NPOC, respiration, and OM
96 molecular richness for a given site. Focusing analyses on the subset of sites that had relatively good correspondence
97 in NPOC between Field and Incubation samples is, therefore, a conservative approach aimed at working with only
98 the most reliable data.

99

100 To subset the data, we calculated the ratio between Field and Incubation NPOC concentrations within each site. If
101 the ratio was less than 1, it was inverted so that all ratios were greater than 1. We regressed log-transformed Field
102 NPOC vs. log-transformed Incubation NPOC, and calculated the R^2 of the associated regression. Log-transformation
103 was used due to the presence of skewed NPOC distributions. Subsequently, we removed samples in order of their
104 ratio, starting with the largest ratio (i.e., the largest proportional difference between Field and Incubation NPOC).
105 Higher R^2 values indicated a tighter relationship between Field and Incubation NPOC, and thus more reliable data.



106 We repeated these steps for all the samples in the Field-Incubation dataset ($n = 228$). We then plotted the R^2 vs. the
107 number of samples removed and selected a threshold for the number of samples to remove (Fig. S1). The resulting
108 curve showed that R^2 increased as a function of points removed until it leveled off. This nonlinear saturating
109 relationship was well-described by a Michaelis-Menten function (Michaelis and Menten, 1913; Johnson and Goody,
110 2011). In this function, the half saturation constant indicates the resource availability at which half of the maximum
111 intake is reached (Mulder and Hendriks, 2014). We used the half saturation constant, estimated from fitting the
112 function to the data in Fig. S1, in a conceptually analogous way. That is, the half saturation constant indicated the
113 number of samples that would need to be removed to gain half of the maximum potential increase in fit between
114 Field and Incubation NPOC. This resulted in removing 30 samples, leading to $R^2 = 0.74$ for the relationship between
115 Field and Incubation NPOC, which was half way between a minimum $R^2 = 0.47$ and maximum $R^2 = 1$. This
116 procedure was used to increase the reliability of the OM molecular richness estimates by removing samples that had
117 the greatest variability in NPOC, which could translate into variability in OM richness as there was a weak but
118 significant relationship between OM richness and NPOC ($R^2 = 0.20$, $p < 0.001$, Fig. S2).

119

120 *Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS)*

121 We used ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) to generate
122 mass spectra of sediment OM pools. Field sediment extracts were normalized to 1.5 mg C L^{-1} , acidified to pH 2 and
123 extracted with solid phase extraction (SPE) PPL cartridges following procedures described by (Dittmar et al., 2008).
124 Note that all samples were normalized to a consistent NPOC concentration prior to SPE and the same sample
125 volume was extracted with the same cartridges and resin mass. Since concentrations were normalized prior SPE, we
126 did not measure extraction efficiency post-extraction. While extraction efficiency will vary across samples, our
127 approach assumes that variation in extraction efficiency is not systematically linked to respiration rate to such a
128 degree that the number of detected peaks becomes correlated with respiration. Although we cannot definitively
129 determine that this assumption is upheld in this dataset, it seems extremely unlikely that the number of observed
130 peaks would become systematically and spuriously linked to respiration due to variation in extraction efficiency.

131

132 FTICR-MS analyses were carried out at the Environmental Molecular Science Laboratory (EMSL) in Richland, WA
133 using a 12 Tesla (12T) Bruker Solarix FTICR mass spectrometer (Bruker, Solarix, Billerica, MA, USA) in negative
134 mode. The method used to assign molecular formulas to FTICR-MS spectra is described in Garayburu-Caruso et al.
135 (Garayburu-Caruso et al., 2020b). Briefly, Formularity (Tolić et al., 2017) was used to align mass lists generated
136 using Bruker DataAnalysis V4.2. Resulting reports were processed using ftmsRanalysis (Bramer et al., 2020). It is
137 important to note that FTICR-MS is a non-targeted approach to reliably identify molecular formulas of organic
138 molecules with masses, but it is not quantitative and does not provide information about the structure of the
139 molecular formulas identified. Our analyses on the Field FTICR-MS data only included samples that passed through
140 the subsetting process described above based on Field and Incubation NPOC. We calculated OM richness as the
141 total number of unique peaks present in one sample.

142



143

144

145 *Incubations and respiration rates*

146 Respiration rates were determined following methods described by Garayburu-Caruso et al. (2020a). Sieved
147 sediments were subsampled into 40 mL clear glass vials (I-Chem amber VOA glass vials) with a 0.5 cm diameter
148 factory calibrated oxygen sensor dot (Fibox 3; PreSens GmbH, Regensburg, Germany). Vials with sediments and
149 unfiltered water from each site were kept in the dark inside the environmental chamber at a 21°C until next day
150 incubations. Reactors consisted of 10 mL of sieved sediments and ~30-35 mL of aerated unfiltered water with no
151 headspace, shaken at 250 rpm for 2 hours. Dissolved oxygen (DO) was measured noninvasively every 15 min for
152 the first hour and every 30 min during the second hour using an oxygen optical meter (Fibox 3; PreSens GmbH,
153 Germany) to read the oxygen sensor dots on the vials. Respiration rates were calculated as the slope of the linear
154 regression between DO concentration and incubation time for each reactor and further normalized per gram of
155 sediment in each reactor. Normalized and not-normalized rates are reported in this manuscript.

156

157 *Statistical analysis*

158 All statistical analyses were completed using R (version 3.6.3)(R Core Team, 2021) with $p < 0.05$ as the significance
159 threshold. We used ordinary least squares regressions (function “lm”) to evaluate relationships between respiration
160 rates and OM richness or NPOC. While not initially expected, we observed an apparent non-linear constraint-based
161 relationship between respiration rate and the inverse of NPOC. To evaluate the statistical significance of the
162 constraint boundary, we subdivided the 1/NPOC data into 10 even bins and found the maximum respiration rate in
163 each of those bins. We then fit a negative exponential function to the relationship between maximum respiration rate
164 and 1/NPOC. To evaluate the potential contribution of OM richness, we used the same approach to regress
165 respiration rate against the ratio of OM richness to NPOC concentration. Base functions in R and ggplot2
166 (Wickham, 2016) were used for these analyses and associated plotting.

167

168 Scripts necessary to reproduce the primary results of this manuscript are available at

169 https://github.com/WHONDRS-Hub/Respiration_and_OM_Richness/. Goldman et al. (2020) provides the raw,
170 unprocessed FTICR-MS data and respiration rate data. FTICR-MS data used in this manuscript were processed
171 following instructions provided in the Goldman et al. (2020) data package.

172

173 **3 Results and Discussion**

174 Both respiration rates and OM molecular richness varied significantly across samples, providing a useful dataset to
175 study the hypothesized negative relationship between these two variables. More specifically, the distribution of
176 aerobic respiration rates revealed a broad range of rates that were highly skewed for rates that were either not
177 normalized (Fig. 2A) or were normalized (Fig. 2B) per gram of sediment. These skewed distributions indicate the
178 potential for biogeochemical “hot spots” (McClain et al. 2003) or “control points” (Bernhardt et al. 2017) at the
179 continental scale. The distribution for OM molecular richness (i.e., number of identified organic molecules)



180 appeared to be multimodal, but dominated by one primary peak (Fig. 2C). OM molecular richness ranged from
181 ~2000-5000 peaks, and we took advantage of this variation to evaluate the relationship between OM richness and
182 respiration rates.

183

184 We did not observe a clear negative relationship between sediment aerobic respiration rates and OM molecular
185 richness, which rejected the hypothesis that higher OM richness will suppress respiration (Fig. 3A,C). The data in
186 Fig.3 suggest there may instead be a peak in maximum respiration rate near intermediate levels of OM molecular
187 richness. There may, therefore, be an optimal level of OM molecular richness that enables high respiration rates, but
188 does not guarantee elevated rates, leading to a unimodal constraint space. Regressions based on maximum
189 respiration rates across the OM richness axis were not significant, however (Fig. 3). These results further reject the
190 hypothesis of any direct relationship between respiration rate and OM richness.

191

192 While we did not observe a direct link between respiration and OM richness, extending the Lehmann et al. (2020)
193 hypothesis revealed a potential influence of OM richness, after controlling for water soluble organic carbon (OC)
194 concentration, measured as NPOC. That is, we posit that any connection between OM richness and respiration is
195 likely modified by the amount of OC. The magnitude of OM richness relative to the concentration of OC could,
196 therefore, provide a stronger constraint over respiration than OM richness alone. High ratios indicate high levels of
197 OM richness relative to the amount of OC, while low ratios indicate low levels of OM richness relative to the
198 amount of OC. In the context of the Lehmann et al. (2020) hypothesis, respiration would therefore be expected to
199 decrease with increasing richness-to-concentration ratios.

200

201 Consistent with this extended hypothesis, we find that maximum respiration rates decreased with increasing
202 richness-to-concentration ratios (Fig. 4). This suggests that at the continental scale OM molecular richness may
203 indirectly influence aerobic respiration rates. However, the influence of OM richness is likely to be relatively minor.
204 That is, maximum respiration rate was also well-explained simply to the inverse of OC concentration (Fig. 5). Note
205 that the relationship between respiration rates and OC concentrations was relatively weak (Fig. S3). The statistical
206 models using the richness-to-concentration ratio are technically better models as they have higher R^2 values than
207 models using only the inverse of OC concentration (cf. Figs. 4, 5). We also note that both types of models are
208 univariate, so there is no penalty for multiple explanatory variables. The bulk of variation in maximum respiration
209 rates (~90%) is, however, explained simply by the inverse of OC concentration.

210

211 We infer that OC concentration could impose a primary constraint over maximum respiration rates, with OM
212 richness potentially contributing additional constraints. As such, any influences of OM richness over respiration
213 rates in hyporheic zone sediments are likely modulated by OC concentration. This is conceptually consistent with
214 observations in marine (Arrieta et al., 2015) and river corridor systems (Stegen et al., 2018). That is, when OM
215 molecular richness is high relative to OC concentration, the probability of a microbe repeatedly encountering the
216 same type of molecule is minimized. In that case, the costs of maintaining metabolic machinery to metabolize any



217 specific type of molecule may outweigh the energy gains (Arrieta et al., 2015). When costs outweigh gains,
218 respiration is expected to be minimized, which is consistent with the respiration constraint boundary to be lowest at
219 the highest richness-to-concentration ratios. In addition, the constraint boundary was non-linear, which is most likely
220 due to the fact that respiration rates cannot be below zero such that increasingly large richness-to-concentration
221 ratios cannot further suppress respiration.

222
223 The combined influences of OM richness and OC concentration are realized as a non-linear constraint space with the
224 vast majority of measured respiration rates falling well below the constraint boundary. This indicates that in most
225 cases, additional controls over respiration drive respiration below its potential maximum. Discerning these
226 additional controls is an important avenue of future work. For example, it would be useful to evaluate the degree to
227 which microbial community diversity, composition, biomass, and/or functional potential are related to deviations
228 from the constraint boundary. In addition, the FTICR-MS data used here provides presence-absence of organic
229 molecules, but not relative abundances of organic molecules. Accounting for among-molecule variation in
230 concentrations could provide insights into factors driving respiration below the constraint boundary. There are many
231 potential influences spanning physical (e.g., sediment texture), chemical (e.g., mineralogy), and biological (e.g.,
232 fungal-to-bacterial ratios) that require investigation in context of the constraint discovered here.

233
234 The outcomes of our study are useful for guiding models towards and away from features and processes that need to
235 be represented to enable predictions of river corridor biogeochemical function. For large scale models, it appears
236 there is a constraint envelope for hyporheic zone respiration rates that is related primarily to OC concentration and
237 even more strongly to OM richness relative to OC concentration. The richness-to-concentration ratio offers a simple
238 way to represent variation in maximum respiration rates across river corridors. Furthermore, for sites without
239 estimates of OM richness, our results suggest that variation in maximum respiration rate could be reasonably
240 estimated via OC concentration alone. Additionally, the constraint spaces observed here could be included in models
241 more mechanistically whereby they would emerge from the representation of how microbial metabolism is
242 influenced by both OM molecular richness and OC concentration. The constraint spaces could, alternatively, be
243 included more phenomenologically in models through probabilistic sampling respiration rates within the constraint
244 space. There may also be additional aspects of OM molecular richness and chemical diversity more broadly (e.g.,
245 thermodynamic properties, elemental ratios) that influence respiration and other biogeochemical rates in hyporheic
246 zone sediments (e.g., Garayburu-Caruso et al., 2020a; Song et al., 2020b). There is a need to examine such
247 possibilities at both local and global scales.

248
249 **4 Code availability:** Scripts necessary to reproduce the primary results of this manuscript are available at
250 https://github.com/WHONDRS-Hub/Respiration_and_OM_Richness/.

251
252 **5 Data availability:** Data were published previously in Goldman et al. (2020).
253



254 **6 Author contributions:** JCS and VAG-C conceptualized the study, VAG-C performed analyses with feedback
255 from JCS, RED processed some of the data, AEG managed the sampling campaign, and LR, JMT, and JW
256 processed samples. JCS and VAG-C drafted the initial manuscript and all authors contributed to revisions.
257

258 **7 Competing interests:** The authors declare that they have no conflict of interest.
259

260 **8 Acknowledgements**

261 This study used data from the Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems
262 (WHONDRS) under the River Corridor Science Focus Area (SFA) at the Pacific Northwest National Laboratory
263 (PNNL) that was generated at the U.S. Department of Energy (DOE) Environmental Molecular Science Laboratory
264 User Facility. PNNL is operated by Battelle Memorial Institute for the U.S. DOE under Contract No. DE-AC05-
265 76RL01830. The SFA is supported by the U.S. DOE, Office of Biological and Environmental Research (BER),
266 Environmental System Science (ESS) Program.
267

268 **9 References**

- 269 Arrieta, J. M., Mayol, E., Hansman, R. L., Herndl, G. J., Dittmar, T., and Duarte, C. M.: Dilution limits dissolved
270 organic carbon utilization in the deep ocean, *Science*, 348, 331–333, <https://doi.org/10.1126/science.1258955>,
271 2015.
- 272 Boulton, A. J., Findlay, S., Marmonier, P., Stanley, E. H., and Valett, H. M.: The functional significance of the
273 hyporheic zone in streams and rivers, *Annual Review of Ecology and Systematics*, 29, 59–81,
274 <https://doi.org/10.1146/annurev.ecolsys.29.1.59>, 1998.
- 275 Bramer, L. M., White, A. M., Stratton, K. G., Thompson, A. M., Claborne, D., Hofmockel, K., and McCue, L. A.:
276 ftmsRanalysis: An R package for exploratory data analysis and interactive visualization of FT-MS data, *PLOS*
277 *Computational Biology*, 16, e1007654, <https://doi.org/10.1371/journal.pcbi.1007654>, 2020.
- 278 Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G.: A simple and efficient method for the solid-phase extraction of
279 dissolved organic matter (SPE-DOM) from seawater, *Limnology and Oceanography: Methods*, 6, 230–235,
280 <https://doi.org/10.4319/lom.2008.6.230>, 2008.
- 281 Fischer, H., Kloep, F., Wilzcek, S., and Pusch, M. T.: A river's liver—microbial processes within the hyporheic zone
282 of a large lowland river, *Biogeochemistry*, 76, 349–371, 2005.
- 283 Garayburu-Caruso, V. A., Stegen, J. C., Song, H.-S., Renteria, L., Wells, J., Garcia, W., Resch, C. T., Goldman, A.
284 E., Chu, R. K., Toyoda, J., and Graham, E. B.: Carbon Limitation Leads to Thermodynamic Regulation of
285 Aerobic Metabolism, *Environmental Science & Technology Letters*, 7, 517–524,
286 <https://doi.org/10.1021/acs.estlett.0c00258>, 2020a.
- 287 Garayburu-Caruso, V. A., Danczak, R. E., Stegen, J. C., Renteria, L., McCall, M., Goldman, A. E., Chu, R. K.,
288 Toyoda, J., Resch, C. T., Torgeson, J. M., Wells, J., Fansler, S., Kumar, S., and Graham, E. B.: Using
289 community science to reveal the global chemogeography of river metabolomes, *Metabolites*, 10, 518,
290 <https://doi.org/10.3390/metabo10120518>, 2020b.
- 291 Goldman, A. E., Chu, R. K., Danczak, R. E., Daly, R. A., Fansler, S., Garayburu-Caruso, V. A., Graham, E. B.,
292 McCall, M. L., Ren, H., and Renteria, L.: WHONDRS Summer 2019 Sampling Campaign: Global River



- 293 Corridor Sediment FTICR-MS, NPOC, and Aerobic Respiration, Environmental System Science Data
294 Infrastructure for a Virtual Ecosystem ..., 2020.
- 295 Graham, E. B., Tfaily, M. M., Crump, A. R., Goldman, A. E., Bramer, L. M., Arntzen, E., Romero, E., Resch, C. T.,
296 Kennedy, D. W., and Stegen, J. C.: Carbon Inputs From Riparian Vegetation Limit Oxidation of Physically
297 Bound Organic Carbon Via Biochemical and Thermodynamic Processes, *Journal of Geophysical Research:*
298 *Biogeosciences*, 122, 3188–3205, <https://doi.org/10.1002/2017JG003967>, 2017.
- 299 Graham, E. B., Crump, A. R., Kennedy, D. W., Arntzen, E., Fansler, S., Purvine, S. O., Nicora, C. D., Nelson, W.,
300 Tfaily, M. M., and Stegen, J. C.: Multi 'omics comparison reveals metabolome biochemistry, not microbiome
301 composition or gene expression, corresponds to elevated biogeochemical function in the hyporheic zone,
302 *Science of The Total Environment*, 642, 742–753, <https://doi.org/10.1016/j.scitotenv.2018.05.256>, 2018.
- 303 Harvey, J. and Gooseff, M.: River corridor science: Hydrologic exchange and ecological consequences from
304 bedforms to basins, *Water Resources Research*, 51, 6893–6922, <https://doi.org/10.1002/2015WR017617>, 2015.
- 305 Johnson, K. A. and Goody, R. S.: The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten
306 Paper, *Biochemistry*, 50, 8264–8269, <https://doi.org/10.1021/bi201284u>, 2011.
- 307 Krause, S., Hannah, D. M., Fleckenstein, J. H., Heppell, C. M., Kaeser, D., Pickup, R., Pinay, G., Robertson, A. L.,
308 and Wood, P. J.: Inter-disciplinary perspectives on processes in the hyporheic zone, *Ecology*, 4, 481–499,
309 <https://doi.org/10.1002/eco.176>, 2011.
- 310 Lehmann, J., Hansel, C. M., Kaiser, C., Kleber, M., Maher, K., Manzoni, S., Nunan, N., Reichstein, M., Schimel, J.
311 P., Torn, M. S., Wieder, W. R., and Kögel-Knabner, I.: Persistence of soil organic carbon caused by functional
312 complexity, *Nature Geoscience*, 13, 529–534, <https://doi.org/10.1038/s41561-020-0612-3>, 2020.
- 313 Michaelis, L. and Menten, M. L.: Die kinetik der invertinwirkung, *Biochem. z.*, 49, 352, 1913.
- 314 Mulder, C. and Hendriks, A. J.: Half-saturation constants in functional responses, *Global Ecology and Conservation*,
315 2, 161–169, 2014.
- 316 Naegeli, M. W. and Uehlinger, U.: Contribution of the Hyporheic Zone to Ecosystem Metabolism in a Prealpine
317 Gravel-Bed-River, *Journal of the North American Benthological Society*, 16, 794–804,
318 <https://doi.org/10.2307/1468172>, 1997.
- 319 Orghidan, T.: A new habitat of subsurface waters: the hyporheic biotope, *Fundamental and Applied Limnology*,
320 291–302, <https://doi.org/10.1127/1863-9135/2010/0176-0291>, 2010.
- 321 R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing,
322 Vienna, Austria, 2021.
- 323 Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R.,
324 Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., and Guth, P.: Global carbon
325 dioxide emissions from inland waters, *Nature*, 503, 355–359, <https://doi.org/10.1038/nature12760>, 2013.
- 326 Schlesinger, W. H. and Melack, J. M.: Transport of organic carbon in the world's rivers, *Tellus*, 33, 172–187,
327 <https://doi.org/10.3402/tellusa.v33i2.10706>, 1981.
- 328 Schlünz, B. and Schneider, R. R.: Transport of terrestrial organic carbon to the oceans by rivers: re-estimating flux-
329 and burial rates, *International Journal of Earth Sciences*, 88, 599–606, <https://doi.org/10.1007/s005310050290>,



330 2000.

331 Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-
332 Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.:
333 Persistence of soil organic matter as an ecosystem property, *Nature*, 478, 49–56,
334 <https://doi.org/10.1038/nature10386>, 2011.

335 Sengupta, A., Fansler, S. J., Chu, R. K., Danczak, R. E., Garayburu-Caruso, V. A., Renteria, L., Song, H.-S.,
336 Toyoda, J., Wells, J., and Stegen, J. C.: Disturbance Triggers Non-Linear Microbe-Environment Feedbacks,
337 *bioRxiv*, 2020.09.30.314328, <https://doi.org/10.1101/2020.09.30.314328>, 2020.

338 Song, H.-S., Stegen, J. C., Graham, E. B., Lee, J.-Y., Garayburu-Caruso, V. A., Nelson, W. C., Chen, X., Moulton,
339 J. D., and Scheibe, T. D.: Representing Organic Matter Thermodynamics in Biogeochemical Reactions via
340 Substrate-Explicit Modeling, *bioRxiv*, 2020.02.27.968669, <https://doi.org/10.1101/2020.02.27.968669>, 2020a.

341 Song, H.-S., Stegen, J. C., Graham, E. B., Lee, J.-Y., Garayburu-Caruso, V. A., Nelson, W. C., Chen, X., Moulton,
342 J. D., and Scheibe, T. D.: Representing Organic Matter Thermodynamics in Biogeochemical Reactions via
343 Substrate-Explicit Modeling, *Frontiers in Microbiology*, 11, <https://doi.org/10.3389/fmicb.2020.531756>, 2020b.

344 Stegen, J. C. and Goldman, A. E.: WHONDRS: a Community Resource for Studying Dynamic River Corridors,
345 *mSystems*, 3, <https://doi.org/10.1128/mSystems.00151-18>, 2018.

346 Stegen, J. C., Johnson, T., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E., Nelson, W. C., Arntzen, E. V.,
347 Chrisler, W. B., Chu, R. K., Fansler, S. J., Graham, E. B., Kennedy, D. W., Resch, C. T., Tfaily, M., and
348 Zachara, J.: Influences of organic carbon speciation on hyporheic corridor biogeochemistry and microbial
349 ecology, *Nature Communications*, 9, 585, <https://doi.org/10.1038/s41467-018-02922-9>, 2018.

350 Tolić, N., Liu, Y., Liyu, A., Shen, Y., Tfaily, M. M., Kujawinski, E. B., Longnecker, K., Kuo, L.-J., Robinson, E.
351 W., Paša-Tolić, L., and Hess, N. J.: Formularity: Software for Automated Formula Assignment of Natural and
352 Other Organic Matter from Ultrahigh-Resolution Mass Spectra, *Analytical Chemistry*, 89, 12659–12665,
353 <https://doi.org/10.1021/acs.analchem.7b03318>, 2017.

354 Toyoda, J. G., Goldman, A. E., Chu, R. K., and Danczak, R. E.: WHONDRS Summer 2019 Sampling Campaign:
355 Global River Corridor Surface Water FTICR-MS and Stable Isotopes [Dataset], ESS-DIVE,
356 <https://doi.org/10.15485/1603775>, 2020.

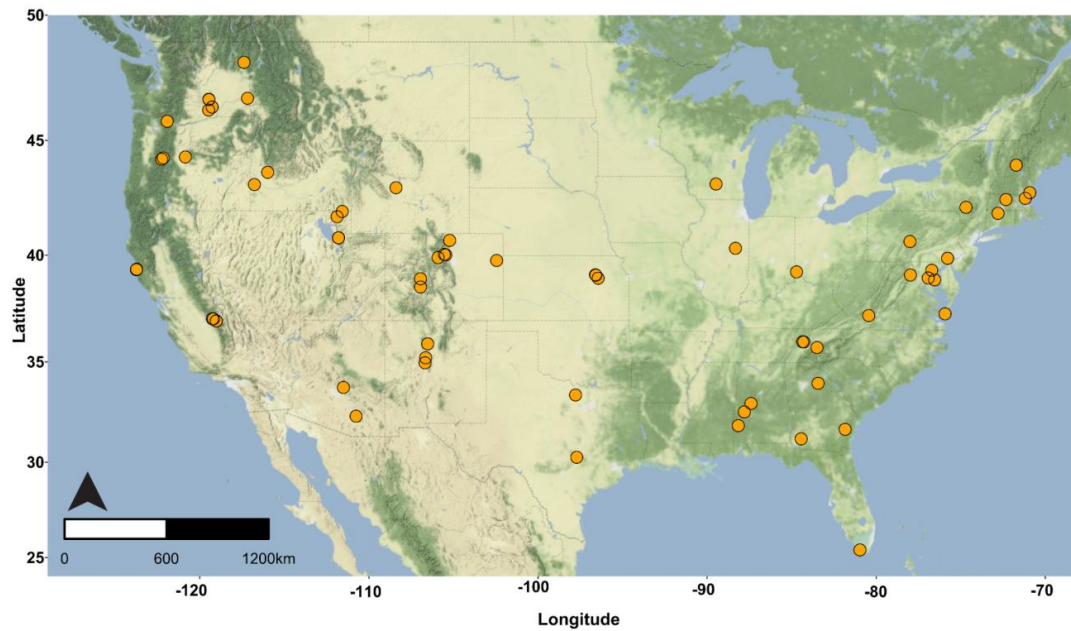
357 Wickham, H.: *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York, 2016.

358

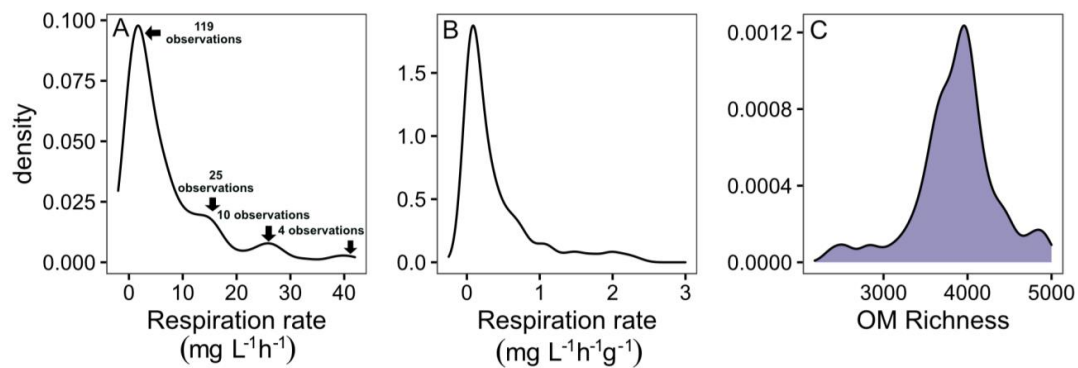
359



360 **Figures**
361

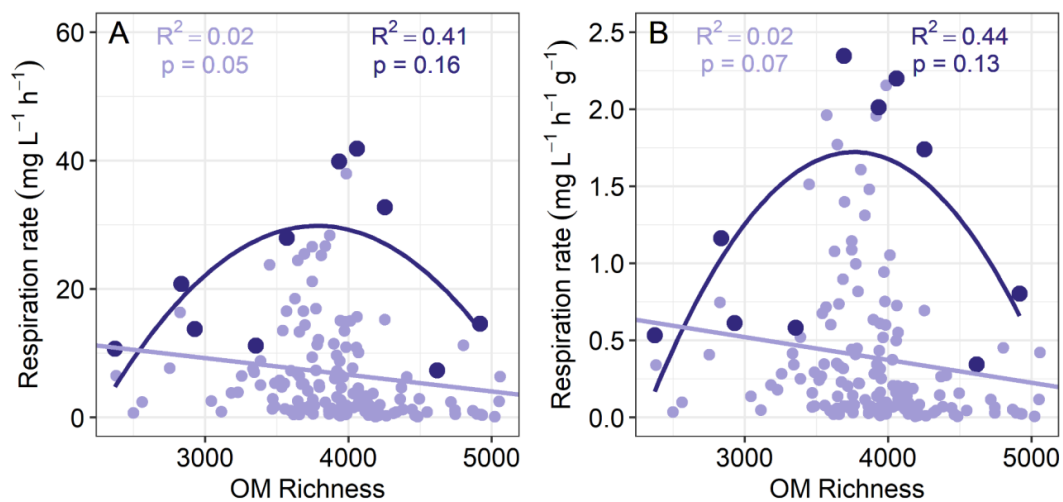


362
363 **Figure 1.** Spatial distribution of sampling locations. At each location three sediment samples were collected from locations
364 distributed along an upstream-downstream gradient within a single stream reach. The map was generated by Sophia McKeever
365 using QGIS, and the base map is copyrighted: ©OpenStreetMap contributors 2022. The map is distributed under the
366 Open Data Commons Open Database 504 License (ODbL) v1.0.
367



368
369
370
371
372
373

Figure 2. Density plots of aerobic respiration measured as oxygen consumption that was either not normalized relative to sediment mass (A) or normalized by sediment mass (B). Panel (C) is a density plot for the number of unique peaks identified in sediment samples, which we refer to as OM richness.



374

375

376

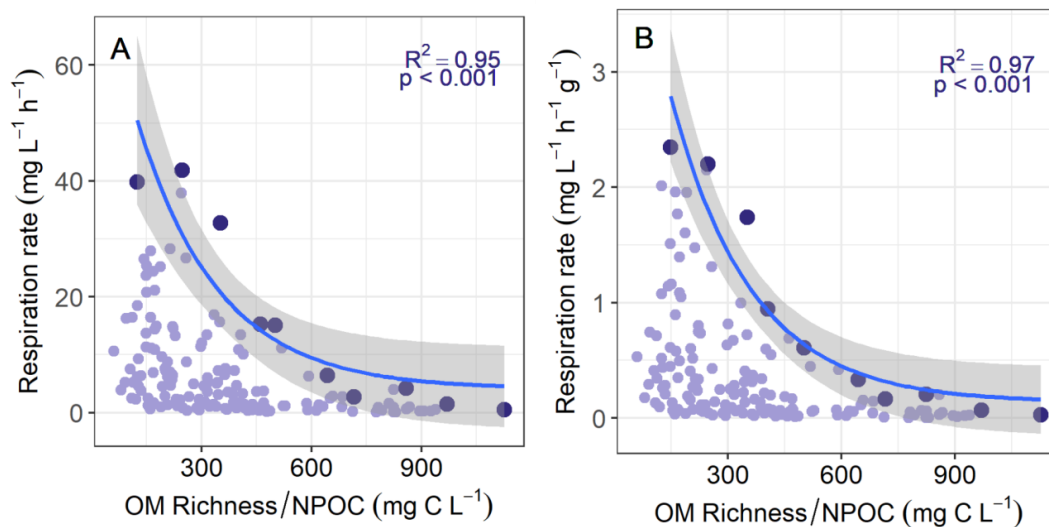
377

378

379

380

Figure 3. Sediment aerobic respiration as a function of OM richness. Respiration was measured as oxygen consumption and was either not normalized (A) or normalized by sediment mass (B). Quadratic regression models based on maximum respiration rates are shown in dark purple while linear regression models based on all respiration values are shown in light purple. Maximum respiration rates were found by subdividing each horizontal axis into 10 even bins. In all cases the models provided poor fits to the data.



381

382

383

384

385

386

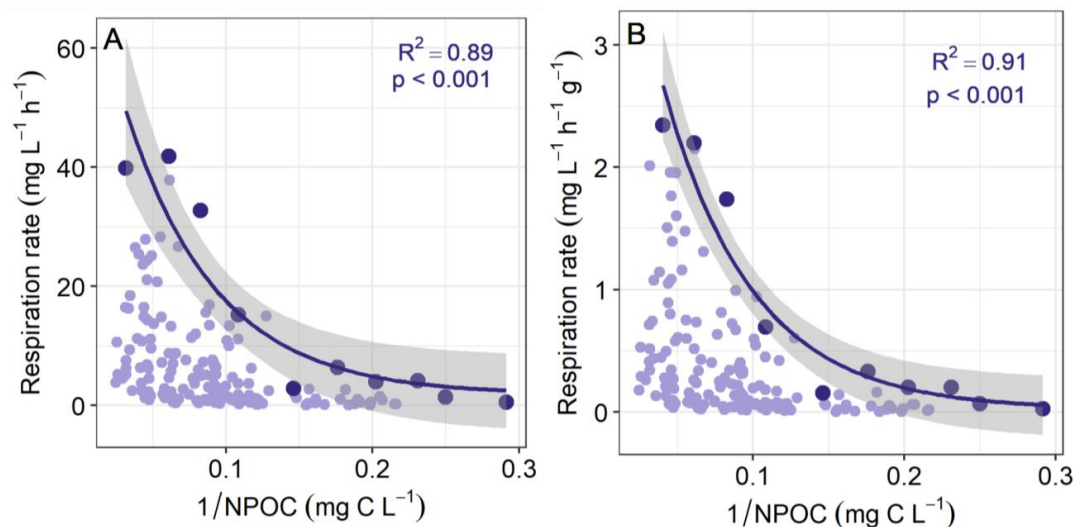
387

388

Figure 4. Maximum sediment respiration rate decreased with increasing values of ratio of OM molecular richness to non-purgeable organic carbon concentration (NPOC). Panels A and B are for respiration that was either not normalized or normalized by sediment mass, respectively. Maximum respiration rates (shown in the darker colors) were found by subdividing each horizontal axis into 10 even bins. All other respiration rates and the corresponding richness-to-concentration ratios are shown in lighter colors. Solid lines represent negative exponential models fit to the maximum respiration rates, with shaded areas indicating 95% confidence intervals. Statistics for each model are provided on each panel.



389



390

391

392

393

394

395

396

Figure 5. Maximum sediment respiration decreased with the inverse of non-purgeable organic carbon concentration (NPOC). Panels A and B are for respiration that was either not normalized or normalized by sediment mass, respectively. Maximum respiration rates (shown in the darker colors) were found by subdividing each horizontal axis into 10 even bins. All other respiration rates and the corresponding 1/NPOC values are shown in lighter colors. Solid lines represent negative exponential models fit to the maximum respiration rates, with shaded areas indicating 95% confidence intervals. Statistics for each model are provided on each panel.