

## MIT Open Access Articles

*Mineral protection regulates long-term  
global preservation of natural organic carbon*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Hemingway, Jordon D. et al. "Mineral protection regulates long-term global preservation of natural organic carbon." *Nature*, 570, 7760 (June 2019): 228–231 © 2019 The Author(s)

**As Published:** 10.1038/S41586-019-1280-6

**Publisher:** Springer Science and Business Media LLC

**Persistent URL:** <https://hdl.handle.net/1721.1/126782>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



1 **Mineral protection regulates long-term global preservation of natural organic**  
2 **carbon**

3 **Authors:** Jordon D. Hemingway<sup>1,\*</sup>, Daniel H. Rothman<sup>2</sup>, Katherine E. Grant<sup>3</sup>, Sarah Z.  
4 Rosengard<sup>4</sup>, Timothy I. Eglinton<sup>5</sup>, Louis A. Derry<sup>3</sup>, Valier V. Galy<sup>6</sup>

5 **Affiliations:**

6 <sup>1</sup>Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge,  
7 MA 02138, USA

8 <sup>2</sup>Lorenz Center, Department of Earth, Atmospheric, and Planetary Science, Massachusetts  
9 Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

10 <sup>3</sup>Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14853, USA

11 <sup>4</sup>Departments of Geography and Earth, Ocean, and Atmospheric Sciences, University of British  
12 Columbia, 2207 Main Mall, Vancouver, BC V6T 1Z4, Canada

13 <sup>5</sup>Geological Institute, Department of Earth Sciences, ETH Zürich, Sonneggstraße 5, Zürich 8092,  
14 Switzerland

15 <sup>6</sup>Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution,  
16 266 Woods Hole Road, Woods Hole, MA 02543, USA

17 \*Correspondence to: [jordon\\_hemingway@fas.harvard.edu](mailto:jordon_hemingway@fas.harvard.edu)

18 **The balance between photosynthetic organic carbon production and respiration controls**  
19 **atmospheric composition and climate<sup>1,2</sup>. The majority of organic carbon is respired back to**  
20 **carbon dioxide in the biosphere, but a small fraction escapes remineralization and is**  
21 **preserved over geologic timescales<sup>3</sup>. By removing reduced carbon from Earth’s surface,**  
22 **this sequestration process promotes atmospheric oxygen accumulation<sup>2</sup> and carbon dioxide**  
23 **removal<sup>1</sup>. Two major mechanisms have been proposed to explain organic carbon**  
24 **preservation: selective preservation of biochemically unreactive compounds<sup>4,5</sup> and**  
25 **protection resulting from interactions with a mineral matrix<sup>6,7</sup>. While both mechanisms can**  
26 **play a role across a range of environments and timescales, their global relative importance**  
27 **on 10<sup>3</sup>- to 10<sup>5</sup>-year timescales remains uncertain<sup>4</sup>. Here we present a global dataset of the**  
28 **distributions of organic carbon activation energy and corresponding radiocarbon ages in**  
29 **soils, sediments, and dissolved organic carbon; we find that activation energy distributions**  
30 **broaden over time in all mineral-containing samples. This result requires increasing bond-**  
31 **strength diversity, consistent with the formation of organo-mineral bonds<sup>8</sup> but inconsistent**  
32 **with selective preservation. Radiocarbon ages further reveal that high-energy, mineral-**  
33 **bound organic carbon persists for millennia relative to low-energy, unbound organic**  
34 **carbon. Our results provide globally coherent evidence for the proposed<sup>7</sup> importance of**  
35 **mineral protection in promoting organic carbon preservation. We suggest that similar**  
36 **studies of bond-strength diversity in ancient sediments may elucidate how and why organic**  
37 **carbon preservation—and thus atmospheric composition and climate—has varied over**  
38 **geologic time.**

39 Two classes of mechanisms—selectivity and protection—have been proposed to explain why  
40 some organic carbon (OC) escapes remineralization in soils and sediments<sup>4-7</sup>. Biochemical  
41 selectivity hypotheses state that intrinsically bioavailable compounds such as sugars and amino  
42 acids are rapidly respired, whereas “recalcitrant” (macro)molecules such as lignin are selectively  
43 preserved due to their low energy yield, large size, and/or a lack of enzymes that can decompose  
44 them<sup>4,5</sup>. Selective preservation has been extensively documented in dissolved OC (DOC)<sup>9</sup>,  
45 decaying woody tissue<sup>10</sup>, and sapropel sediments containing almost exclusively organic matter<sup>5</sup>.  
46 In contrast, protection hypotheses state that particles shield OC from respiration regardless of  
47 intrinsic recalcitrance, potentially due to occlusion within pore spaces that are inaccessible to  
48 microbes and their extracellular enzymes<sup>4,8,11-14</sup>. Specifically, protection often involves

49 physicochemical interactions with a mineral matrix. Mineral protection is evidenced by  
50 laboratory incubation experiments<sup>15</sup>, non-selective OC preservation in sinking marine particles<sup>16</sup>,  
51 and strong observed correlations between OC content, iron content<sup>17,18</sup>, and mineral surface  
52 area<sup>7,8,19</sup> in soils and sediments.

53 Although both selectivity and protection must play some role across a range of  
54 environments and timescales, directly measuring their relative importance using traditional  
55 geochemical techniques has proven difficult. To address this issue, first we develop a conceptual  
56 framework that predicts how chemical compositions, as recorded by OC bond strength, and  
57 preservation timescales evolve when OC is exposed to each mechanism. We then compare  
58 predictions to results from a global dataset.

59 **Figure 1** illustrates our model. Selective preservation states that OC diversity—that is,  
60 the distribution of carbon bond strengths contained within a given sample—decreases as certain  
61 chemical structures are preferentially respired. OC chemical composition becomes more  
62 homogeneous with time (**Fig. 1a**, top). In contrast, mineral protection allows for respiration of  
63 different chemical structures at similar rates. However, some OC molecules become occluded  
64 within aggregates<sup>8,14</sup> or bound to mineral surfaces during aging and persist for longer than they  
65 would in the absence of the mineral matrix (**Fig. 1a**, bottom). Because bound, stabilized  
66 compounds represent a new chemical and energetic state relative to their unbound  
67 counterparts<sup>20,21</sup>, we expect bond-strength diversity to increase with time when OC is protected.  
68 Whereas both mechanisms predict a similar decrease in the amount of OC remaining (**Fig. 1b**,  
69 top), they produce opposite trends for the temporal evolution of bond-strength diversity (**Fig. 1b**,  
70 bottom).

71 Information on bond-strength diversity is contained in a continuous function  $p(E)$  that  
72 describes the fraction of OC associated with each activation energy,  $E$ . Critically, in addition to  
73 the energy needed to break covalent bonds (a function of OC molecular structure),  $E$  includes the  
74 energy required to remove bound carbon from particles<sup>20-22</sup>. Here, we estimate  $p(E)$  by  
75 determining the bond-strength distribution that best predicts the measured OC decay profile  
76 when a sample is exposed to increasing temperature using ramped pyrolysis/oxidation (RPO)<sup>22</sup>.  
77  $E$  distributions are unique to each remineralization pathway<sup>22</sup>; we therefore treat  $p(E)$  calculated  
78 by RPO as a proxy for OC bond-strength diversity but emphasize that thermal  $E$  is not equivalent  
79 to the activation energy of any environmental enzymatic respiration reaction (**Methods**). Our

80 interpretation is informed by  $E$  distribution differences between samples subject to the same  
81 experimental conditions, not by absolute  $E$  values. In addition to providing an estimate of  $p(E)$ ,  
82 RPO quantifies the OC radiocarbon activity ( $^{14}\text{C}/^{12}\text{C}$ , expressed in  $^{14}\text{C}$  yr), a measure of the time  
83 that has passed since initial formation, over multiple decay windows. In doing so, RPO results  
84 allow us to determine how OC age varies as bonds become stronger.

85 To assess global trends across all environments, this study compiles RPO results from  
86 riverine and marine DOC, soils, riverine particulate OC (POC), and marine sediments (**Extended**  
87 **Data Fig. 1**). These samples were primarily obtained from aerobic environments and represent  
88 oxic respiration, although some were exposed to periodic anoxia during their diagenetic history  
89 (*e.g.*, Arabian Sea sediments). This global dataset includes samples analyzed using slightly  
90 different protocols (*e.g.*, pyrolysis vs. oxidation), but an uncertainty assessment indicates that  
91 methodological variability is small and that complex kinetic behavior during sample heating is  
92 unlikely to explain observed trends (**Methods**)<sup>22,23</sup>. Furthermore, because our focus is the  
93 persistence of biospherically produced OC, we excluded samples containing carbonates,  
94 anthropogenic OC, and/or rock-derived OC, resulting in a final dataset of 62 samples. This  
95 screening procedure removes permafrost-dominated soils since these samples exhibit  
96 anomalously old  $^{14}\text{C}$  ages due to protracted storage under frozen conditions. It is possible that  
97 some retained samples contain minor rock-derived OC inputs. However, such contributions  
98 never exceed approximately 5% of total OC and cannot explain resulting  $E$  and  $^{14}\text{C}$  trends  
99 (**Methods**).

100 **Figure 2** shows the  $E$  distributions for each sample, grouped by environment. We  
101 observe significant OC bond-strength diversity differences between DOC and all mineral-  
102 containing samples. Here, we quantify OC bond-strength diversity in terms of  $\sigma_E$ , the standard  
103 deviation of the distribution  $p(E)$ . DOC exhibits the narrowest distribution of any sample type. In  
104 contrast,  $p(E)$  is substantially broader in all mineral-containing sample types, indicating greater  
105 OC bond-strength diversity. We additionally quantify the mean  $E$  value,  $\mu_E$ , for each sample type  
106 as a measure of the average OC bond strength. Across sample types,  $\mu_E$  increases in the order:  
107 soils, river POC, DOC, and marine sediments. Finally, intra-sample  $^{14}\text{C}$  trends for mineral-  
108 containing samples show a globally coherent increase in  $^{14}\text{C}$  reservoir age with increasing  $E$ .  
109 High- $E$  OC is consistently thousands of  $^{14}\text{C}$  yr older than low- $E$  OC, especially within riverine  
110 POC and marine sediments, indicating protracted preservation of high- $E$  material in these

111 environments. In contrast, DOC samples show either the opposite trend or no relationship  
112 between  $E$  and  $^{14}\text{C}$  age.

113 To determine whether these observations are diagnostic of a particular preservation  
114 mechanism, we compare the observed temporal evolution of  $\sigma_E$  to that predicted by selective  
115 preservation and by mineral protection. Because the timescale of exposure to respiration cannot  
116 be measured directly, we approximate it using OC content ( $\%OC$ )<sup>24</sup>. This approximation  
117 assumes that loss by respiration is the dominant control on  $\%OC$ ; it ignores secondary effects  
118 such as dilution. This is supported at the global scale by the observation that intra-sample  $^{14}\text{C}$  age  
119 variances increase with decreasing  $\%OC$  for all mineral-containing samples following a single  
120 power-law relationship (**Extended Data Fig. 2**). This relationship persists across  $\sim 3$  orders of  
121 magnitude in  $\%OC$  despite large differences in sample latitude, physical environment, grain size,  
122 *etc.*, suggesting that samples with lower  $\%OC$  have been exposed to aerobic respiration for  
123 longer timescales (*c.f.*, **Fig. 1b**, top)<sup>6</sup>. Furthermore, soils generally contain higher  $\%OC$  and  
124 lower  $^{14}\text{C}$  age variances than do river POC and marine sediments, consistent with the idea that  
125 sediments represent longer temporal integration. Thus, while we do not expect this relationship  
126 to strictly hold at the individual soil/sediment profile level,  $\%OC$  remains a robust respiration  
127 timescale proxy for global theory-data comparisons.

128 We test the hypothesis that selective preservation is the dominant control by quantifying  
129 how  $p(E)$  evolves following this mechanism. We treat each OC component as decaying  
130 according to a rate constant defined by its bond strength. Because high- $E$  OC is generally older  
131 than low- $E$  OC in our dataset, we assume a decreasing rate constant with increasing  $E$ . High- $E$   
132 material is therefore selectively preserved, as expected if stronger bonds are inherently more  
133 resistant to degradation<sup>4</sup>. The distribution  $p(E)$  can then be obtained at each point in time by  
134 solving the first-order differential equation that describes the evolution of  $p(E)$  as  $\%OC$   
135 decreases (**Methods**). Using a Gaussian distribution for the initial condition, we obtain the  
136 results shown in **Fig. 3a**. In contrast, **Fig. 3b** shows the observed average  $p(E)$  for all mineral-  
137 containing samples binned by  $\%OC$ . Although selective preservation is consistent with the  
138 observed narrow DOC distribution<sup>9</sup>, it fails to reproduce observed  $p(E)$  trends for all mineral-  
139 containing samples, which broaden with decreasing  $\%OC$  and suggest increasing bond-strength  
140 diversity.

141 To better quantify bond-strength evolution through time, we bin all mineral-containing  
142 samples and plot  $\mu_E$  and  $\sigma_E$  as a function of %OC (**Fig. 4**). This approach approximates a time-  
143 series. To estimate the autotrophic biomass starting composition, we include three samples  
144 representing mineral-free, vascular-plant OC: woody tissue, cellulose, and a Hawaiian soil  
145 organic (O) horizon.

146 **Figure 4b** reveals increasing  $\sigma_E$  with increasing time of exposure to respiration,  
147 consistent with low  $\sigma_E$  values for autotrophic biomass and with previously reported  $p(E)$   
148 evolution from a Hawaiian soil profile<sup>25</sup>. In contrast, selective preservation predicts that  $\sigma_E$  must  
149 decrease with time. Conceivably, selective preservation combined with resupply of respired,  
150 low- $E$  material could lead to a nearly constant  $\sigma_E$  over time, but it cannot explain the observation  
151 that  $p(E)$  broadens beyond its initial distribution. Specifically, resupply fails to predict the  
152 observed increase in  $p(E)$  at the high- $E$  end of the distribution as OC ages. Rather, mineral  
153 surfaces likely broaden  $p(E)$  relative to autotrophic biomass by protecting otherwise bioavailable  
154 OC and by strengthening OC bonding environments<sup>20,21</sup>. Secondary condensation<sup>4,26</sup> reactions  
155 have additionally been hypothesized to occur in the environment; these reactions could similarly  
156 form new bonds and thus influence  $E$  distributions. However, condensation is likely not the  
157 major driver of observed global  $E$  trends since (i) this would require DOC to behave similarly,  
158 which is not observed, and (ii) this would not necessarily lead to the formation of stronger bonds  
159 and thus higher- $E$  OC, which is observed. Still, condensation may play a minor role in OC  
160 preservation.

161 We propose that some OC fraction becomes strongly bound to minerals during aging in  
162 the environment, potentially due to ligand exchange or direct chelation with iron oxides<sup>18</sup>, and  
163 leads to the production of high- $E$  material that persists for millennia. This is supported by  
164 observed intra-sample <sup>14</sup>C trends, which show a global increase in <sup>14</sup>C reservoir age with  
165 increasing  $E$ . High- $E$  OC is consistently thousands of <sup>14</sup>C yr older than low- $E$  material, especially  
166 for riverine POC and marine sediments.

167 All mineral-containing samples additionally include lower  $E$  OC than that found in DOC  
168 and autotrophic biomass, leading to only a weak increase in  $\mu_E$  with decreasing %OC (**Fig. 4a**)  
169 despite significant ingrowth of high- $E$  material. This requires production and preservation of  
170 low- $E$  compounds during heterotrophic alteration in soils and sediments<sup>25-28</sup>. Microbial  
171 heterotrophs are known to produce an abundance of small molecules containing

172 carboxyl/carbonyl (C=O) and hydroxyl (C-OH) carbon atoms (*e.g.*, amino acids, sugars)<sup>26,28</sup>, for  
173 example during extracellular hydrolysis of polymers<sup>27</sup>. These compounds can then interact with  
174 particle surfaces *via* cation bridging and/or hydrophobic interactions and become protected. In  
175 contrast, the lack of particle interactions further explains the absence of low-*E* material and  
176 narrow <sup>14</sup>C gradients for DOC; small, bioavailable molecules in mineral-free environments are  
177 free to be selectively respired over short timescales. For example, laboratory incubation  
178 experiments show increased heterotrophic respiration rates once weakly bound OC is desorbed  
179 from particles<sup>29</sup>. Furthermore, Barber *et al.*<sup>17</sup> have shown that C=O and C-OH carbon atoms,  
180 which exhibit low inherent *E* values<sup>21</sup>, are stabilized by binding to iron minerals.

181 Combined, these observations imply weak mineral association (possibly due to occlusion  
182 within aggregates<sup>8,14</sup>) as a mechanism to preserve low-*E* OC, consistent with its absence from  
183 DOC and vascular-plant biomass. However, unlike high-*E* material, which ages for millennia,  
184 low-*E* OC is continuously adsorbed/desorbed and resupplied by alteration of recently produced,  
185 <sup>14</sup>C-enriched biomass. This leads to the observed increase in <sup>14</sup>C reservoir age with increasing *E*.

186 Mineral protection carries implications for the fate of OC in soils and sediments and for  
187 atmospheric composition throughout Earth history. Whereas soils, riverine POC, and marine  
188 sediments appear to follow a globally coherent  $\sigma_E$  vs. %OC trend today (**Extended Data Fig. 2**),  
189 these trends could have differed in the geologic past due to changes in oxidant availability,  
190 mineral surface area, surface charge, and/or OC composition (*e.g.*, availability of organic  
191 ligands). We propose that such changes are reflected in the distribution  $p(E)$  of geologically  
192 preserved OC. For example, production of the secondary clay mineral smectite during  
193 weathering on land has been shown to promote protracted OC storage<sup>30</sup>. Furthermore, by  
194 enhancing the deposition of OC-rich sediments in marine continental margins, smectite  
195 production has been proposed as a carbon sequestration mechanism during warm periods such as  
196 Cretaceous ocean anoxic events<sup>31</sup>. Assuming the mechanism proposed here equally applies to the  
197 geologic past, then our framework provides a means to directly test this hypothesis using the  
198 evolution of  $p(E)$  through time. In general, we predict that strong interactions with mineral  
199 surfaces expand OC bond-strength diversity and increase preservation efficiency, thus decreasing  
200 atmospheric CO<sub>2</sub> levels and promoting oxygen accumulation.

201

202 **References:**



- 203 1. Berner, R. A. Atmospheric carbon dioxide levels over Phanerozoic time. *Science* **249**, 1382–  
204 1386 (1990).
- 205 2. Berner, R. A. & Canfield, D. E. A new model for atmospheric oxygen over Phanerozoic  
206 time. *Am. J. Sci.* **289**, 333–361 (1989).
- 207 3. Hayes, J. M. & Waldbauer, J. R. The carbon cycle and associated redox processes through  
208 time. *Phil. Trans. R. Soc. B* **361**, 931–950 (2006).
- 209 4. Burdige, D. J. Preservation of organic matter in marine sediments: controls, mechanisms,  
210 and an imbalance in sediment organic carbon budgets? *Chem. Rev.* **107**, 467–485 (2007).
- 211 5. Hatcher, P. G., Spiker, E. C., Szeverenyi, N. M. & Maciel, G. E. Selective preservation and  
212 origin of petroleum-forming aquatic kerogen. *Nature* **305**, 498–501 (1983).
- 213 6. Hedges, J. I. & Keil, R. G. Sedimentary organic matter preservation: an assessment and  
214 speculative synthesis. *Mar. Chem.* **49**, 81–115 (1995).
- 215 7. Mayer, L. M. Relationships between mineral surfaces and organic carbon concentrations in  
216 soils and sediments. *Chem. Geol.* **114**, 347–363 (1994).
- 217 8. Vogel, C. *et al.* Submicron structures provide preferential spots for carbon and nitrogen  
218 sequestration in soils. *Nat. Comms.* **5**, 2947 (2014).
- 219 9. Kellerman, A. M., Kothawala, D. N., Dittmar, T. & Tranvik, L. J. Persistence of dissolved  
220 organic matter in lakes related to its molecular characteristics. *Nat. Geosci.* **8**, 454–457  
221 (2015).
- 222 10. Hedges, J. I., Cowie, G. L., Ertel, J. R., James Barbour, R. & Hatcher, P. G. Degradation of  
223 carbohydrates and lignins in buried woods. *Geochim. Cosmochim. Acta* **49**, 701–711  
224 (1985).
- 225 11. Rothman, D. H. & Forney, D. C. Physical model for the decay and preservation of marine  
226 organic carbon. *Science* **316**, 1325–1328 (2007).
- 227 12. Keil, R. G. & Mayer, L. M. “Mineral matrices and organic matter,” in *Treatise on*  
228 *Geochemistry* (eds. Holland, H. & Turekian, K.) 337–359 (Elsevier Ltd., 2014).
- 229 13. Torn, M. S., Trumbore, S. E., Chadwick, O. A., Vitousek, P. M. & Hendricks, D. M. Mineral  
230 control of soil organic carbon storage and turnover. *Nature* **389**, 170–173 (1997).
- 231 14. Six, J., Conant, R. T., Paul, E. A. & Paustian, K. Stabilization mechanisms of soil organic  
232 matter: implications for C-saturation of soils. *Plant and Soil* **241**, 155–176 (2002).

- 233 15. Hunter, W. R. *et al.* Metabolism of mineral-sorbed organic matter and microbial lifestyles in  
234 fluvial ecosystems. *Geophys. Res. Lett.* **43**, 1582–1588 (2016).
- 235 16. Hedges, J. I. *et al.* Evidence for non-selective preservation of organic matter in sinking  
236 marine particles. *Nature* **409**, 801–804 (2001).
- 237 17. Barber, A. *et al.* Preservation of organic matter in marine sediments by inner-sphere  
238 interactions with reactive iron. *Sci. Rep.* **7**, 366–377 (2017).
- 239 18. Lalonde, K., Mucci, A., Ouellet, A. & Gelinias, Y. Preservation of organic matter in  
240 sediments promoted by iron. *Nature* **483**, 198–200 (2012).
- 241 19. Wagai, R. & Mayer, L. M. Sorptive stabilization of organic matter in soils by hydrous iron  
242 oxides. *Geochim. Cosmochim. Acta* **71**, 25–35 (2007).
- 243 20. Leinweber, P. & Schulten, H. R. Advances in analytical pyrolysis of soil organic matter. *J.*  
244 *Anal. Appl. Pyrol.* **49**, 359–383 (1999).
- 245 21. Plante, A. F., Fernández, J. M. & Leifeld, J. Application of thermal analysis techniques in  
246 soil science. *Geoderma* **153**, 1–10 (2009).
- 247 22. Hemingway, J. D., Rothman, D. H., Rosengard, S. Z. & Galy, V. V. Technical note: An  
248 inverse method to relate organic carbon reactivity to isotope composition from serial  
249 oxidation. *Biogeosciences* **14**, 5099–5114 (2017).
- 250 23. Williams, E. K., Rosenheim, B. E., McNichol, A. P. & Masiello, C. A. Charring and non-  
251 additive chemical reactions during ramped pyrolysis: applications to the characterization  
252 of sedimentary and soil organic material. *Org. Geochem.* **77**, 106–114 (2014).
- 253 24. Boudreau, B. P. & Ruddick, B. R. On a reactive continuum representation of organic matter  
254 diagenesis. *Am. J. Sci.* **291**, 507–538 (1991).
- 255 25. Grant, K. E., Galy, V. V., Chadwick, O. A. & Derry, L. A. Thermal oxidation of carbon in  
256 organic matter rich volcanic soils: insights into SOC age differentiation and mineral  
257 stabilization. *Biogeochemistry in revision*, (2019).
- 258 26. Lehmann, J. & Kleber, M. The contentious nature of soil organic matter. *Nature* **113**, 143–9  
259 (2015).
- 260 27. Arnosti, C., Repeta, D. J. & Blough, N. V. Rapid bacterial degradation of polysaccharides in  
261 anoxic marine systems. *Geochim. Cosmochim. Acta* **58**, 2639–2652 (1994).
- 262 28. Schmidt, M. W. I. *et al.* Persistence of soil organic matter as an ecosystem property. *Nature*  
263 **478**, 49–56 (2011).

- 264 29. Keil, R. G., Montluçon, D. B., Prahl, F. G. & Hedges, J. I. Sorptive preservation of labile  
265 organic matter in marine sediments. *Nature* **370**, 549–552 (1994).
- 266 30. Kennedy, M., Pevear, D. & Hill, R. J. Mineral Surface Control of Organic Carbon in Black  
267 Shale. *Science* **295**, 657–660 (2002).
- 268 31. Kennedy, M. J. & Wagner, T. Clay mineral continental amplifier for marine carbon  
269 sequestration in a greenhouse ocean. *PNAS* **108**, 9776–9781 (2011).

270

271 **Acknowledgements:** This work is dedicated to John Hayes, whose knowledge, support, and  
272 inspiration was always invaluable. We thank the National Ocean Sciences Accelerator Mass  
273 Spectrometer staff, especially A. McNichol, A. Gagnon, and M. Lardie-Gaylord for RPO  
274 assistance; P. Raymond, E. Kyzivat, R. Spencer, and A. Stubbins for supplying sample material;  
275 B. Rosenheim, C. Schafer, and X. Zhang for providing access to raw data from published  
276 manuscripts; A. Pearson, D. Johnston, and A. Piasecki for comments and discussions. This  
277 research was supported by: the NSF Graduate Research Fellowship Program grant number  
278 2012126152 (J.D.H.); NASA Astrobiology grant number NNA13AA90A and NSF grant number  
279 EAR-1338810 (D.H.R.); the NSF-IGERT in Cross Scale Biogeochemistry and Climate at Cornell  
280 University (K.E.G.); and the WHOI Independent Study Award (V.V.G.).

281

282 **Author Contributions:** J.D.H., D.H.R., and V.V.G. conceived the study; J.D.H. compiled data;  
283 J.D.H., K.E.G., and S.Z.R. performed laboratory measurements; J.D.H. and D.H.R. developed  
284 theoretical models and analyzed data; J.D.H., K.E.G., L.A.D., and V.V.G. provided samples;  
285 T.I.E. and L.A.D. contributed analytical tools and discussion; J.D.H., D.H.R., and V.V.G. wrote  
286 the manuscript with input from all authors.

287

288 **Author Information:** Reprints and permissions information is available at  
289 [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Correspondence  
290 and requests for materials should be addressed to [jordon\\_hemingway@fas.harvard.edu](mailto:jordon_hemingway@fas.harvard.edu).

291

292 **Data Availability:** All data that support the findings of this study are available on the “RPO  
293 Compilation Online Database” with the identifier doi:10.5281/zenodo.1158742.

294 **Figure Legends:**

295 **Figure 1 | Schematic representation of two potential preservation mechanisms.** Panel (a)  
296 shows a cartoon of each mechanism: (top) selective preservation, in which certain compounds  
297 are preferentially preserved and (bottom) mineral protection, in which OC becomes altered,  
298 attached to particles surfaces, and stabilized. For selective preservation, OC is initially composed  
299 of labile (white) and recalcitrant (black) compounds, which decay at different rates. In contrast,  
300 for mineral protection, both black and white circles have the same inherent lability. Either type  
301 can then become stabilized by clay particles (large parallelograms) and persist. Panel (b) shows  
302 how the amount (number of circles) and diversity (different colors and bound/free states) evolve  
303 with time for each mechanism.

304  
305 **Figure 2 |  $E$  and  $^{14}\text{C}$  age distributions.** (top) Each gray line represents the  $E$  distribution for one  
306 sample; the thick black line is the average  $E$  distribution for each sample type: (a) DOC (riverine  
307 and marine; average  $\mu_E = 163.3 \text{ kJ mol}^{-1}$ ; average  $\sigma_E = 14.8 \text{ kJ mol}^{-1}$ ), (b) soil OC (average  $\mu_E =$   
308  $153.8 \text{ kJ mol}^{-1}$ ; average  $\sigma_E = 20.1 \text{ kJ mol}^{-1}$ ), (c) riverine POC (average  $\mu_E = 160.1 \text{ kJ mol}^{-1}$ ;  
309 average  $\sigma_E = 25.3 \text{ kJ mol}^{-1}$ ), and (d) sedimentary OC (average  $\mu_E = 168.9 \text{ kJ mol}^{-1}$ ; average  $\sigma_E =$   
310  $24.0 \text{ kJ mol}^{-1}$ ). The area under each curve integrates to unity [*i.e.*,  $p(E)$  is a probability density  
311 function]. To compare all sample types side-by-side,  $p(E)$  units are not shown. (bottom) Each  
312 line represents the  $^{14}\text{C}$  reservoir age as a function of  $E$  for one sample; note that reservoir age is  
313 presented on a logarithmic scale. The number of samples of each type ( $n$ ) is shown in each panel.  
314 Reservoir age for each RPO fraction is equal to the  $^{14}\text{C}$  age for that fraction minus the youngest  
315  $^{14}\text{C}$  age of any fraction within a given sample (**Eq. 5**). OC in mineral-containing samples  
316 becomes older with increasing  $E$  whereas DOC samples exhibit either the opposite trend or no  
317 relationship.

318  
319 **Figure 3 | Predicted and observed  $p(E)$  evolution.**  $p(E)$  as a function of %OC for (a) selective  
320 preservation predictions and (b) measured observations. The area under each curve integrates to  
321 unity [*i.e.*,  $p(E)$  is a probability density function]. For panel (b), all mineral-containing samples  
322 for which %OC data exist ( $n = 55$ ) were binned by OC content and average  $E$  distributions were  
323 calculated for each bin: black,  $\text{OC} \geq 10\%$  (median OC = 15.5%;  $n = 6$ ); dark gray,  $10\% > \text{OC} \geq$   
324  $1\%$  (median OC = 1.7%;  $n = 26$ ); light gray,  $\text{OC} < 1\%$  (median OC = 0.8%;  $n = 23$ ). For panel

325 (a), the black line is a Gaussian distribution with %OC,  $\mu_E$ , and  $\sigma_E$  equal to those for the  
326 observed OC > 10% median distribution in panel (b). This distribution then evolves following  
327 selective preservation (**Methods**) until OC content reaches 1.7% (dark gray) and 0.8% (light  
328 gray), the median values for measured 10% > OC  $\geq$  1% and OC < 1% bins in panel (b). Selective  
329 preservation cannot explain observed trends.

330

331 **Figure 4 | Predicted and observed  $\mu_E$  and  $\sigma_E$  evolution.** Measured (black circles) and  
332 predicted (dotted line, gray line) (a)  $\mu_E$  and (b)  $\sigma_E$  evolution as a function of %OC. To visualize  
333 global trends, all mineral-containing samples for which %OC data exist ( $n = 55$ ) were binned and  
334 plotted at the mean value for each bin (11 bins containing 5 samples each). Additionally,  
335 autotrophic biomass samples (cellulose, woody tissue, and O horizon soil;  $n = 3$ ) were binned  
336 and plotted as a white circle at OC = 50 %, a canonical value for photosynthetic biomass. Black  
337 bars are  $\pm 1\sigma$  uncertainty for each bin. The dotted line shows selective preservation predictions  
338 starting from the measured phototrophic biomass composition. The gray line shows selective  
339 preservation predictions starting from an initial Gaussian composition with  $\mu_E$  and  $\sigma_E$  that were  
340 chosen to best fit the observed data (**Methods**). Both predictions clearly cannot explain observed  
341 trends, especially for  $\sigma_E$ . The %OC axis is presented on a logarithmic scale and is reversed to  
342 emphasize that OC content generally decreases with time.

343 **Methods:**

344 **Ramped pyrolysis/oxidation instrumental setup.** The RPO instrument continuously heats OC,  
345 monitors evolved CO<sub>2</sub>, and traps CO<sub>2</sub> for isotope analysis. RPO methodology has been described  
346 previously<sup>22,32,33</sup>, and further information describing the detailed protocols used for each sample  
347 compiled in this study can be found in the original publications. In summary, all samples were  
348 either operated in oxidation mode (carrier gas: 92% He, 8% O<sub>2</sub>) or pyrolysis mode (carrier gas:  
349 100% He) with a flow rate of 35 mL min<sup>-1</sup> and a ramp rate of 5 °C min<sup>-1</sup>.

350

351 **Published data compilation.** We compiled all published RPO results that include <sup>14</sup>C  
352 measurements<sup>25,33-47</sup>. All thermograms and blank-corrected RPO fraction data (temperatures,  
353 masses, <sup>13</sup>C contents, and <sup>14</sup>C activities) can be accessed in the online database<sup>48</sup>. Samples  
354 included in this database span five continents and a latitude range between 70.5° N and 64.5° S  
355 (**Extended Data Fig. 1**). Samples were categorized as either: DOC, soil, riverine POC, or marine  
356 sediment. Sample collection and pre-treatment methods are described in detail in each of the  
357 original publications and are summarized in **Extended Data Table 1**.

358

359 **New samples analyzed for this study.** In addition to previously published results, we included  
360 in our database five samples presented for the first time in this study. These new samples were  
361 collected and treated as follows: (i) and (ii) Congo River DOC and surface POC collected near  
362 Brazzaville, Republic of Congo (4.3° S, 15.2° E) in April 2016. Approximately 100 L of water  
363 was filtered through 0.22 µm polyether sulfone (PES) membrane filters. Sediments were retained  
364 on the filters and 2 L of effluent water was collected in acid-leached HDPE bottles (10 % HCl, 1  
365 week) for DOC extraction. Sediments were freeze-dried at -40 °C, whereas DOC was extracted  
366 using bond-elut PPL cartridges, eluted in MeOH, transferred into pre-combusted quartz reactor  
367 inserts (850 °C, 5 hours), and dried under a stream of ultrapure N<sub>2</sub> gas<sup>49</sup>. (iii) Grassland soil  
368 overlying the Pololu lava flow located on the Kohala Peninsula of Hawaii (20.1° N, 155.7° W).  
369 Archived material from 80 to 90 cm depth was freeze-dried at -40 °C, homogenized, and sub-  
370 sampled for RPO analysis. (iv) Kosi River surface suspended sediment (sample LO765 in Galy  
371 *et al.*<sup>50</sup>). Surface water was collected near Chatra, Nepal (26.9° N, 87.2° E) in 2007 and was  
372 filtered through 0.22 µm PES membrane filters, de-carbonated over HCl fumes at 60 °C for 72  
373 hours, rinsed 3× with 18.2 MΩ MilliQ water, and freeze-dried at -40 °C. (v) Suwannee River

374 DOC sample NOM II obtained from the International Humic Substances Society (2R101N; 30.8°  
375 N, 82.4° W). Sample material was concentrated by reverse osmosis from large water volumes (~  
376 36,000 L), desalted via cation exchange, freeze-dried, and homogenized<sup>51</sup>. This sample was  
377 analyzed in freeze-dried powder form as provided by IHSS.

378

379 **Activation energy calculations.** To quantitatively compare OC bond strengths between samples,  
380 we converted observed thermograms into activation energy distributions. In addition to bond  
381 strength, thermograms depend on experimental conditions such as oven ramp rate<sup>52</sup>. This  
382 dependency can be removed by treating OC decay as a function of  $E$ , the activation energy  
383 required to volatilize a given carbon atom.  $E$  reflects the energy required to fully oxidize each  
384 carbon atom when exposed to a particular oxidation reaction pathway, including energy needed  
385 to break both covalent bonds and secondary interactions with mineral surfaces<sup>20,21,38</sup>, and is  
386 therefore a suitable quantitative proxy for bond strength. We emphasize that  $E$  values calculated  
387 here reflect thermal oxidation and are not quantitatively applicable to other (*e.g.*, enzymatic)  
388 oxidation pathways.

389 We calculated  $E$  distributions by finding the inverse solution to a set of parallel, non-  
390 isothermal, first-order kinetic decay reactions<sup>22</sup> as implemented in the ‘rampedpyrox’ package  
391 for Python v.3.5 (ref. 53). Model details, mathematical derivations, and sensitivity analyses can  
392 be found in Hemingway *et al.*<sup>22</sup>. As recommended, here we set the Arrhenius pre-exponential  
393 factor to be  $10^{10} \text{ s}^{-1}$  for all samples. Finally, to compare  $E$  between samples, we extracted two  
394 scalar metrics from each continuous distribution. First, the mean value for a given sample,  
395 termed  $\mu_E$ , was calculated as the first moment of  $E$ :

$$396 \mu_E = \int_0^{\infty} E p(E) dE. \quad (1)$$

397 Second, the standard deviation of  $E$  for a given sample, termed  $\sigma_E$ , was calculated as the sample  
398 variance of  $E$ :

$$399 \sigma_E = (\mu_{E^2} - [\mu_E]^2)^{1/2}. \quad (2)$$

400 **Radiocarbon calculations.** All radiocarbon measurements were initially reported as Fm, the  
401 “fraction modern,” and were subsequently converted to  $^{14}\text{C}$  years. Fm is equal to the  $^{14}\text{C}$  activity  
402 of a given sample (normalized to  $\delta^{13}\text{C} = -25.0 \text{ ‰}$  VPDB to remove fractionation effects) relative

403 to 95% of the  $^{14}\text{C}$  activity of NBS Oxalic Acid I in 1950 (ref. 54). The time that has passed since  
404 OC was initially fixed from  $\text{CO}_2$  was then calculated in  $^{14}\text{C}$  yr as a function of Fm according to:

$$405 \quad {}^{14}\text{C yr} = -8033 \ln(\text{Fm}), \quad (3)$$

406 where 8033 is the Libby  $^{14}\text{C}$  half-life divided by  $\ln(2)$ . We derived two different metrics to  
407 compare the spread in RPO-derived  $^{14}\text{C}$  ages between samples with different mean ages. First,  
408 we calculated the age variance for each sample as

$$409 \quad \text{Var}({}^{14}\text{C yr}) = E[({}^{14}\text{C yr} - \overline{{}^{14}\text{C yr}})^2], \quad (4)$$

410 where  $\overline{{}^{14}\text{C yr}}$  is the mean  $^{14}\text{C}$  age for a given sample and  $E[x]$  denotes the expected value for  
411 variable  $x$ . Second, we determined the reservoir age, R, for each RPO fraction within a given  
412 sample relative to the most  $^{14}\text{C}$ -enriched fraction from that sample (typically the fraction  
413 collected over the lowest temperature range). Reservoir age was calculated as

$$414 \quad R_f = {}^{14}\text{C yr}_f - \min_f |{}^{14}\text{C yr}_f|, \quad (5)$$

$$415 \quad f = 1, \dots, n_f,$$

416 where  ${}^{14}\text{C yr}_f$  is the  $^{14}\text{C}$  age for RPO fraction  $f$ ,  $\min_f |{}^{14}\text{C yr}_f|$  is the minimum  $^{14}\text{C}$  age measured  
417 for any RPO fraction within a given sample, and  $n_f$  is the number of RPO fractions collected for  
418 that sample. Because all  $^{14}\text{C}$  decays at the same rate (*i.e.*, according to the  $^{14}\text{C}$  half-life), R is  
419 conserved with the passage of time<sup>55</sup>. Finally, R can also be expressed in “fractional” notation  
420 analogous to Fm according to:

$$421 \quad \text{F}^{14}\text{R} = e^{-R/8033}. \quad (6)$$

422 All OC aging trends discussed in this manuscript are reported in  $^{14}\text{C}$  yr and R notation, whereas  
423 sample screening and end-member mixing calculations utilize Fm and  $\text{F}^{14}\text{R}$  notation for  
424 mathematical convenience (see “Sample screening,” below). Finally, all  $^{14}\text{C}$  results reported here  
425 were corrected for instrument-specific RPO procedural blank contamination<sup>33,56</sup>.

426

427 **Uncertainty assessment.** Because our dataset contains samples analyzed using multiple  
428 methods, we assessed the influence of sample pre-treatment and RPO instrumental conditions on  
429 resulting  $E$  distributions. This was done by subjecting the same samples to various experimental  
430 conditions (**Extended Data Table 2**). Specifically, there are three methodological variables that



431 might influence OC thermal stability within our dataset: (i) whether or not a sample was  
432 acidified<sup>38,39</sup>, (ii) the type of acid treatment, and (iii) whether the RPO instrument was operated  
433 in pyrolysis or oxidation mode.

434 First, the overall influence of acidification was determined using 4 carbonate-free  
435 samples: one natural Canadian soil<sup>38</sup> (sample CaSa\_soil in **Extended Data Table 2**), two  
436 Hawaiian soils<sup>25,33</sup> (PuuEke4 and Pololu1), and one Amazon River suspended sediment<sup>39</sup>  
437 (AmRi\_ju2014\_54m). PuuEke4 was analyzed in both pyrolysis and oxidation mode, leading to 5  
438 acidified vs. non-acidified pairs. For these samples, acidification increased  $\mu_E$  values by  $5.6 \pm 5.6$   
439  $\text{kJ mol}^{-1}$  and  $\sigma_E$   $3.8 \pm 2.4 \text{ kJ mol}^{-1}$  ( $n = 5$ ).

440 Second, the difference between HCl fumigation<sup>33</sup> and liquid HCl treatment<sup>40</sup> was  
441 assessed using one Hawaiian soil (Pololu1), one Himalayan river suspended sediment (PB-60),  
442 and one pure rock-derived OC (OC<sub>petro</sub>) sample from Taiwan (LW\_08-03\_petro). Both PB-60  
443 and LW\_08-03\_petro were analyzed in both pyrolysis and oxidation modes, leading to 5 acid  
444 treatment pairs. Fumigation decreased  $\mu_E$  values by  $6.7 \pm 10.9 \text{ kJ mol}^{-1}$  and  $\sigma_E$  values by  $2.4 \pm$   
445  $4.0 \text{ kJ mol}^{-1}$  relative to liquid acid treatment in these experiments ( $n = 5$ ).

446 Finally, the difference between RPO pyrolysis and oxidation modes was determined  
447 using the abovementioned samples and acidification protocols. For these samples, pyrolysis  
448 mode analysis appeared to increase  $\mu_E$  values by  $5.9 \pm 4.8 \text{ kJ mol}^{-1}$  and  $\sigma_E$  values by  $2.1 \pm 2.3 \text{ kJ}$   
449  $\text{mol}^{-1}$  relative to oxidation mode ( $n = 7$ ). Additionally, Grant *et al.*<sup>25</sup> report only minor  
450 differences between pyrolysis and oxidation modes in a set of Hawaiian soil samples. This  
451 similarity indicates that complex kinetic reactions such as charring<sup>23</sup> occurring within the RPO  
452 instrument cannot fully explain our observed trends. If charring were driving RPO results, then  
453 we would expect significantly higher  $\mu_E$  values when the RPO instrument is operated in  
454 pyrolysis mode due to the lack of atmospheric O<sub>2</sub>; however, this is not observed. Furthermore,  
455 charring would act to shift low- $E$ , <sup>14</sup>C-enriched OC toward higher  $E$  values, thus artificially  
456 dampening  $E$  vs. <sup>14</sup>C signals and weakening observed trends.

457 Thus, none of these methodological differences can explain the signals observed in our  
458 dataset. Sample results span a  $\mu_E$  range of  $80 \text{ kJ mol}^{-1}$ , significantly broader than both the largest  
459 bias ( $6.7 \text{ kJ mol}^{-1}$ ) and the widest uncertainty ( $\pm 10.9 \text{ kJ mol}^{-1}$ ) observed in this assessment.  
460 Similarly, observed  $\sigma_E$  values in our dataset span a range of  $35 \text{ kJ mol}^{-1}$ , much broader than the  
461 maximum methodological bias ( $\pm 3.8 \text{ kJ mol}^{-1}$ ) and uncertainty ( $\pm 4.0 \text{ kJ mol}^{-1}$ ). We therefore

462 assume a conservative uncertainty of  $\pm 10.9 \text{ kJ mol}^{-1}$  for  $\mu_E$  and  $\pm 4.0 \text{ kJ mol}^{-1}$  for  $\sigma_E$ , resulting in  
463 a minimum signal-to-noise ratio above 7 for both metrics.

464

465 **Sample screening.** Because our focus was on the preservation of naturally produced biospheric  
466 OC that has been continuously exposed to heterotrophic respiration since initial formation, we  
467 screened the database for samples containing other carbon sources or alternate diagenetic  
468 histories. Specifically, samples were omitted if they were shown to contain: (i) carbonates, (ii)  
469 anthropogenically produced biospheric OC, (iii) rock-derived or “petrogenic” OC ( $\text{OC}_{\text{petro}}$ ), or  
470 (iv) OC that has not been continuously exposed to diagenesis due to protracted storage in ice  
471 (permafrost) or due to the rapid shutdown of pedogenesis (paleosols)<sup>45</sup>. Screening was performed  
472 as follows:

473         First, all samples that were not acidified prior to RPO analysis were screened for  
474 carbonates. This includes PPL-extracted and ultrafiltered DOC<sup>33</sup>, Amazon River suspended  
475 sediments<sup>39</sup>, and a subset of soils<sup>25,38</sup>. Because DOC cannot contain carbonates by definition, all  
476 DOC samples were retained. For Amazon River sediments, Rosengard<sup>39</sup> analyzed one sample  
477 (AmRi\_ju2014\_54m) both un-acidified and after acid fumigation to directly test for the presence  
478 of carbonates. Both sample treatments contain nearly identical RPO-fraction mass-weighted <sup>14</sup>C  
479 activity (un-acidified:  $F_m = 0.820 \pm 0.002$ , fumigated:  $F_m = 0.770 \pm 0.002$ ) and <sup>13</sup>C content (un-  
480 acidified:  $\delta^{13}\text{C} = -29.3 \pm 0.1 \text{ ‰ VPDB}$ , fumigated:  $\delta^{13}\text{C} = -29.1 \pm 0.1 \text{ ‰ VPDB}$ ), a source tracer  
481 that is highly sensitive to the presence of carbonates. Because of this similarity, especially for  
482 <sup>13</sup>C content, we conclude that any carbonate contribution to these samples is negligible, as  
483 discussed in the original publication<sup>39</sup>. This is further supported by thermogram results, which  
484 show negligible CO<sub>2</sub> production within the carbonate temperature range for these samples  
485 (approximately 600 to 800 °C)<sup>32</sup>. Furthermore, all soils collected from the Kohala Peninsula of  
486 Hawaii<sup>25</sup> were retained for two reasons: (i) this landscape contains purely igneous bedrock,  
487 precluding the possibility of detrital carbonates, and (ii) RPO thermograms do not show  
488 significant CO<sub>2</sub> produced within the carbonate temperature range. In contrast, one unacidified  
489 Canadian soil likely contains pedogenic carbonates<sup>38</sup> and was removed.

490         Second, samples were screened for significant anthropogenic influence since the  
491 inclusion of agricultural (*e.g.*, fertilizers, pesticides) or urban (*e.g.*, petroleum-derived

492 compounds) OC would likely influence RPO results. This resulted in the exclusion of only a  
493 single Canadian soil sample collected from an agricultural plot<sup>38</sup>.

494 Third, samples were screened for OC<sub>petro</sub> contribution. Samples were removed if they met  
495 any of the following three criteria: (i) If the original publication discusses the influence of OC<sub>petro</sub>  
496 or quantifies OC<sub>petro</sub> content. This removes Taiwanese soils, Taiwanese river suspended  
497 sediments, and Himalayan river suspended sediments<sup>37,40</sup>. (ii) If the F<sub>m</sub> value for any RPO  
498 fraction was statistically identical to zero. Because OC<sub>petro</sub> is formed on timescales significantly  
499 longer than the <sup>14</sup>C half-life, it is <sup>14</sup>C-free by definition. Thus, an RPO fraction with F<sub>m</sub> = 0  
500 would indicate purely OC<sub>petro</sub> contribution to that fraction. This criterion further excludes  
501 Himalayan river suspended sediments originally presented in this study. (iii) If RPO results  
502 suggest that OC<sub>petro</sub> contributes greater than 5 % of total OC. OC<sub>petro</sub> contribution can be  
503 estimated using RPO results according to:

$$504 \quad \%OC_{\text{petro}}^{\text{estimated}} = \sum_{(\mu_{E,f} - 2\sigma_{E,f}) > 185 \text{ kJ mol}^{-1}} (1 - F^{14}R_f) \left( \frac{m_f}{m_{\text{tot}}} \right) \times 100 \%, \quad (7)$$

505 where  $\mu_{E,f}$ ,  $\sigma_{E,f}$ ,  $F^{14}R_f$ , and  $m_f$  are the mean and standard deviation of  $E$ ,  $F^{14}R$  value, and carbon  
506 mass for RPO fraction  $f$ ;  $185 \text{ kJ mol}^{-1}$  is the minimum OC<sub>petro</sub>  $E$  cutoff<sup>37</sup>; and  $m_{\text{tot}}$  is the total  
507 carbon mass for the sample. This estimate treats all OC contained within the OC<sub>petro</sub>  $E$  range as a  
508 mixture of OC<sub>petro</sub> and un-aged biospheric OC (*i.e.*,  $F^{14}R = 1$ ). Because it is likely that high- $E$   
509 biospheric OC is described by  $F^{14}R < 1$ , we treat %OC<sub>petro</sub> values calculated using **Eq. 7** as  
510 maximum estimates. This is supported by Himalayan river suspended sediments for which  
511 independently calculated %OC<sub>petro</sub> estimates exist<sup>57</sup> (samples KoRi\_sed\_LO765,  
512 NaRi\_sed\_LO752, and NaRi\_sed\_PB60). For all samples ( $n = 3$ ), %OC<sub>petro</sub> values calculated  
513 using **Eq. 7** are roughly double those reported previously<sup>57</sup>, confirming that **Eq. 7** maximizes  
514 predicted %OC<sub>petro</sub> values and that our screening procedure is therefore conservative. This  
515 criterion further removes one Mississippi River suspended sediment<sup>42</sup> and two Antarctic  
516 sediment samples<sup>44</sup>. Retained samples may contain minor amounts of OC<sub>petro</sub>, especially fluvial  
517 sediments due to their integrative nature. We therefore assessed the sensitivity of our results to  
518 minor OC<sub>petro</sub> contributions (see “OC<sub>petro</sub> sensitivity analysis,” below).

519 Finally, samples were screened for alternate diagenetic histories. Specifically, a sample  
520 was removed if it has not been continuously exposed to heterotrophic remineralization since the  
521 time of OC formation due to (i) protracted storage under frozen conditions or (ii) rapid shutdown

522 of pedogenesis. Inclusion of such samples would conflate multiple preservation mechanisms; for  
523 example, since thawed, previously frozen samples would appear old in  $^{14}\text{C}$  space but have only  
524 exposed to heterotrophic respiration for a fraction of the time that has passed since OC  
525 formation. This resulted in the exclusion of permafrost-containing samples from the Colville  
526 River basin<sup>43,47</sup> as well as paleosol samples from Louisiana<sup>45</sup>. Still, we note that inclusion of  
527 these samples does not substantively change observed  $E$  or  $^{14}\text{C}$  trends.

528 In total, our screening procedure removed 56 potential samples from our analysis (mostly  
529 due to  $\text{OC}_{\text{petro}}$  contributions), resulting in a final database of 62 samples (6 DOC, 11 soils, 27  
530 fluvial sediments, and 18 continental margin and coastal marine sediments). Retained samples  
531 are denoted as such in **Extended Data Table 1**.

532  
533  **$\text{OC}_{\text{petro}}$  sensitivity analysis.** We performed a back-of-the-envelope sensitivity analysis to assess  
534 the influence of small  $\text{OC}_{\text{petro}}$  contributions (*i.e.*, less than 5 % of total OC) on  $E$  distributions of  
535 retained samples. Biospheric OC  $E$  was assumed to be Gaussian with  $\mu_E = 140 \text{ kJ mol}^{-1}$  and  $\sigma_E$   
536  $= 15 \text{ kJ mol}^{-1}$ , the minimum observed values for non-DOC samples in our dataset. We then  
537 included a 5%  $\text{OC}_{\text{petro}}$  contribution as a Gaussian with  $\mu_E = 230 \text{ kJ mol}^{-1}$  and  $\sigma_E = 20 \text{ kJ mol}^{-1}$ ,  
538 the maximum observed values for the  $\text{OC}_{\text{petro}}$  sample included in our uncertainty assessment.  
539 This approach maximized the potential  $\text{OC}_{\text{petro}}$  influence on  $E$  distributions. Resulting  $\mu_E$   
540 increased by  $5 \text{ kJ mol}^{-1}$  and  $\sigma_E$  increases by  $9 \text{ kJ mol}^{-1}$  relative to the purely biospheric OC case.  
541 However, this influence is likely significantly smaller in reality, especially for  $\sigma_E$  since this  
542 metric is sensitive to the difference in petrogenic and biospheric mean  $E$  values. Assuming a  
543 typical sediment with  $\mu_E = 170 \text{ kJ mol}^{-1}$  and  $\sigma_E = 25 \text{ kJ mol}^{-1}$ , adding 5 %  $\text{OC}_{\text{petro}}$  as above  
544 increased both  $\mu_E$  and  $\sigma_E$  by only  $3 \text{ kJ mol}^{-1}$ , well within the methodological uncertainty.

545 We similarly assessed the potential  $\text{OC}_{\text{petro}}$  influence on  $F^{14}\text{R}$  results. To do so, we  
546 assumed that a given sample contains 5 evenly-spaced RPO fractions (*i.e.*, each contains 20% of  
547 total OC), that biospheric OC  $F^{14}\text{R} = 1$  for all fractions, and that  $\text{OC}_{\text{petro}}$  (5% of total OC) is  
548 entirely contained within the highest- $E$  fraction. Thus,  $\text{OC}_{\text{petro}}$  contributes one quarter of OC in  
549 the highest- $E$  fraction ( $= 5\% \text{ OC}_{\text{petro}} \text{ total} / 20\% \text{ total OC per fraction}$ ), leading to  $F^{14}\text{R} = 0.75$   
550 and a reservoir age of  $R = 2,300 \text{ }^{14}\text{C yr}$ . These assumptions maximize the potential  $\text{OC}_{\text{petro}}$   
551 influence on  $F^{14}\text{R}$ , but still cannot explain observed reservoir ages as high as  $13,000 \text{ }^{14}\text{C yr}$ .

552           Nonetheless, inclusion of  $\text{OC}_{\text{petro}}$  does artificially increase  $\mu_E$ ,  $\sigma_E$ , and  $^{14}\text{C}$ -based reservoir  
553 ages relative to purely biospheric OC. Because samples retained in our dataset might contain  
554 small  $\text{OC}_{\text{petro}}$  contributions (*i.e.*, less than 5% of total OC), these results should be taken as  
555 maximum values.

556  
557 **Statistical analyses.** Because scatterplots generally contain uncertainty in both  $X$  and  $Y$   
558 variables, all linear relationships presented in this study (*i.e.*, **Extended Data Fig. 2**) were fit  
559 using reduced major axis regression<sup>58</sup>.

560  
561 **Selective preservation predictions.** Finally, we developed a model to predict the evolution of  
562  $\mu_E$  and  $\sigma_E$  if OC preservation were controlled exclusively by selective preservation. This model  
563 predicts  $\rho(t,E)$ , the mass of OC remaining at time  $t$  that is associated with activation energy  $E$ , as  
564 a function of  $G(t)$ , the total mass of OC remaining at time  $t$ . To generate the solid lines in **Fig. 4**,  
565 we then determined  $\mu_E$  and  $\sigma_E$  at each  $t$  following **Eqs. 1-2**. Mathematically, this model was  
566 derived as follows:

567           The mass of OC remaining at time  $t$  that is associated with activation energy  $E$  can be  
568 written as:

$$569 \quad \rho(t,E) = \rho(E)e^{-k(E)t}, \quad (8)$$

570 where  $\rho(E)$  is the initial mass of OC associated with  $E$  and  $k(E)$  is the first-order decay rate for  
571 OC associated with  $E$  (Ref. 22, 59, 60). OC associated with  $E$  described by lower  $k(E)$  will be  
572 selectively preserved relative to OC associated with  $E$  described by higher  $k(E)$ .

573           Here, we make the simplifying assumption that OC associated with the lowest RPO-  
574 derived  $E$  values would decay most rapidly under selective preservation; that is, higher- $E$  OC is  
575 preserved for longer. Although high- $E$  OC has been shown to decay on centennial timescales in  
576 some environments<sup>37</sup>, this assumption is supported by laboratory experiments indicating that  
577 thermal and microbial decay rates are positively correlated<sup>61</sup>. Nonetheless, we emphasize that  
578 this assumption will not qualitatively influence  $\sigma_E$  evolution (**Fig. 4b**); selective preservation  
579 requires that the distribution of  $\rho(t,E)$  *must* become narrower with time as some OC is  
580 preferentially respired. We therefore define  $k(E)$  following the Arrhenius equation:

$$581 \quad k(E) = -\omega e^{-E/RT}, \quad (9)$$

582 where  $\omega$  is the Arrhenius pre-exponential factor for microbial decay,  $R$  is the ideal gas constant,  
583 and  $T$  is temperature (here assumed to be 25 °C). The total mass of OC remaining at time  $t$  can  
584 then be calculated as<sup>60</sup>:

$$585 \quad G(t) = \int_0^{\infty} \rho(t, E) dE. \quad (10)$$

586 Because this function is not invertible, we cannot substitute  $G(t)$  into **Eq. 8** and calculate  $\rho(f^{-1}(G), E)$  directly. Rather, we choose a value for  $\omega$ , solve **Eq. 8** and **10** over some discretized  
587 vector  $\mathbf{t}$ , and calculate  $\mu_E$  and  $\sigma_E$  for each node in  $\mathbf{t}$ . Importantly, our choice of  $\omega$  has no effect on  
588 resulting  $E$  distributions nor %OC vs.  $\mu_E$  and %OC vs.  $\sigma_E$  trajectories. Because  $\omega$  is a scalar  
589 multiplier that does not change as a function of  $E$ , increasing  $\omega$  results in more rapid overall OC  
590 decay but does not influence the distribution of  $\rho(f^{-1}(G), E)$  for a given  $G(t)$ .  
591

592 To generate the  $E$  distributions shown in **Fig. 3**, we solved the model assuming that  $\rho(E)$   
593 follows a Gaussian distribution with the following initial conditions: OC = 15.5%,  $\mu_E = 150.3$  kJ  
594 mol<sup>-1</sup> and  $\sigma_E = 17.7$  kJ mol<sup>-1</sup>; that is, the median values for all retained samples in our database  
595 with OC  $\geq 10\%$  ( $n = 6$ ). We again emphasize that our choice of initial distribution shape will not  
596 qualitatively influence  $\mu_E$  and  $\sigma_E$  evolution. To generate the %OC vs.  $\mu_E$  and %OC vs.  $\sigma_E$   
597 trajectories shown in **Fig. 4**, we used the canonically assumed initial OC content for recently  
598 produced organic matter of 50%. We then simultaneously determined the initial  $\mu_E$  and  $\sigma_E$  values  
599 that minimize the root mean square error (RMSE) between theory and observations, again  
600 assuming that  $\rho(E)$  follows a Gaussian distribution. This resulted in best-fit initial values of  $\mu_E =$   
601 111.3 kJ mol<sup>-1</sup> and  $\sigma_E = 23.3$  kJ mol<sup>-1</sup>, with an RMSE of 11.2 kJ mol<sup>-1</sup>.  
602

### 603 **Methods references:**

- 604 32. Rosenheim, B. E. *et al.* Antarctic sediment chronology by programmed-temperature  
605 pyrolysis: methodology and data treatment. *Geochem. Geophys. Geosyst.* **9**, Q04005  
606 (2008).
- 607 33. Hemingway, J. D. *et al.* Assessing the blank carbon contribution, isotope mass balance, and  
608 kinetic isotope fractionation of the ramped pyrolysis/oxidation instrument at NOSAMS.  
609 *Radiocarbon* **59**, 179–193 (2017).

- 610 34. Bao, R., McNichol, A. P., McIntyre, C. P., Xu, L. & Eglinton, T. I. Dimensions of  
611 radiocarbon variability within sedimentary organic matter. *Radiocarbon* **60**, 775–790  
612 (2018).
- 613 35. Bao, R. *et al.* Tectonically-triggered sediment and carbon export to the Hadal zone. *Nat.*  
614 *Comms.* **9**, 121 (2018).
- 615 36. Bianchi, T. S. *et al.* Paleoreconstruction of organic carbon inputs to an oxbow lake in the  
616 Mississippi River watershed: effects of dam construction and land use change on regional  
617 inputs. *Geophys. Res. Lett.* **42**, 7983–7991 (2015).
- 618 37. Hemingway, J. D. *et al.* Microbial oxidation of lithospheric organic carbon in rapidly eroding  
619 tropical mountain soils. *Science* **360**, 209–212 (2018).
- 620 38. Plante, A. F., Beaupré, S. R., Roberts, M. L. & Baisden, W. T. Distribution of radiocarbon  
621 ages in soil organic matter by thermal fractionation. *Radiocarbon* **55**, 1077–1083 (2013).
- 622 39. Rosengard, S. Z. “Novel analytical strategies for tracing the organic carbon cycle in marine  
623 and riverine particles,” Ph.D. thesis, MIT/WHOI Joint Program in Oceanography,  
624 Cambridge, MA (2017).
- 625 40. Rosenheim, B. E. & Galy, V. V. Direct measurement of riverine particulate organic carbon  
626 age structure. *Geophys. Res. Lett.* **39**, L19703 (2012).
- 627 41. Rosenheim, B. E., Domack, E. W., Santoro, J. A. & Gunter, M. Improving Antarctic  
628 sediment <sup>14</sup>C dating using ramped pyrolysis: An example from the Hugo Island Trough.  
629 *Radiocarbon* **55**, 115–126 (2013).
- 630 42. Rosenheim, B. E. *et al.* River discharge influences on particulate organic carbon age  
631 structure in the Mississippi/Atchafalaya River system. *Global Biogeochem. Cy.* **27**, 154–  
632 166 (2013).
- 633 43. Schreiner, K. M., Bianchi, T. S. & Rosenheim, B. E. Evidence for permafrost thaw and  
634 transport from an Alaskan North Slope watershed. *Geophys. Res. Lett.* **41**, 3117–3126  
635 (2014).
- 636 44. Subt, C., Fangman, K. A., Wellner, J. S. & Rosenheim, B. E. Sediment chronology in  
637 Antarctic deglacial sediments: reconciling organic carbon <sup>14</sup>C ages to carbonate <sup>14</sup>C ages  
638 using Ramped PyrOx. *The Holocene* **26**, 265–273 (2016).

- 639 45. Vetter, L., Rosenheim, B. E., Fernandez, A. & Törnqvist, T. E. Short organic carbon turnover  
640 time and narrow  $^{14}\text{C}$  age spectra in early Holocene wetland paleosols. *Geochem.*  
641 *Geophys. Geosyst.* **18**, 142–155 (2017).
- 642 46. Williams, E. K., Rosenheim, B. E., Allison, M., McNichol, A. P. & Xu, L. Quantification of  
643 refractory organic material in Amazon mudbanks of the French Guiana Coast. *Mar. Geol.*  
644 **363**, 93–101 (2015).
- 645 47. Zhang, X. *et al.* Permafrost organic carbon mobilization from the watershed to the Colville  
646 River Delta: evidence from  $^{14}\text{C}$  ramped pyrolysis and lignin biomarkers. *Geophys. Res.*  
647 *Lett.* **91**, 385–10 (2017).
- 648 48. Hemingway, J. D., Schafer, C. & Rosenheim, B. E. RPO compilation online database.  
649 (2018). doi:10.5281/zenodo.1158742
- 650 49. Dittmar, T., Koch, B. P., Hertkorn, N. & Kattner, G. A simple and efficient method for the  
651 solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol.*  
652 *Oceanogr. Methods* **6**, 230–235 (2008).
- 653 50. Galy, V. V., Eglinton, T. I., France-Lanord, C. & Sylva, S. P. The provenance of vegetation  
654 and environmental signatures encoded in vascular plant biomarkers carried by the  
655 Ganges-Brahmaputra rivers. *Earth Planet. Sci. Lett.* **304**, 1–12 (2011).
- 656 51. Averett, R. C., Leenheer, J. A., McKnight, D. M. & Thorn, K. A. Humic substances in the  
657 Suwannee River, Georgia: interactions, properties, and proposed structures. *USGS Water-*  
658 *Supply Paper 2373* (1994).
- 659 52. Miura, K. & Maki, T. A simple method for estimating  $f(E)$  and  $k_0(E)$  in the distributed  
660 activation energy model. *Energy Fuels* **12**, 864–869 (1998).
- 661 53. Hemingway, J. D. *rampedpyrox*: open-source tools for thermoanalytical data analysis, 2016-.  
662 <http://pypi.python.org/pypi/rampedpyrox>, doi:10.5281/zenodo.839815 (2017).
- 663 54. Reimer, P. J., Brown, T. A. & Reimer, R. W. Discussion: Reporting and calibration of post-  
664 bomb  $^{14}\text{C}$  data. *Radiocarbon* **46**, 1299–1304 (2004).
- 665 55. Soulet, G., Skinner, L. C., Beaupré, S. R. & Galy, V. V. A note on reporting of reservoir  $^{14}\text{C}$   
666 disequilibria and age offsets. *Radiocarbon* **46**, 205–211 (2016).
- 667 56. Fernandez, A. *et al.* Blank corrections for ramped pyrolysis radiocarbon dating of  
668 sedimentary and soil organic carbon. *Anal. Chem.* **86**, 12085–12092 (2014).



- 669 57. Galy, V. V. & Eglinton, T. I. Protracted storage of biospheric carbon in the Ganges-  
670 Brahmaputra basin. *Nat. Geosci.* **4**, 843–847 (2011).
- 671 58. Rayner, J. Linear relations in biomechanics: the statistics of scaling functions. *J. of Zool.*  
672 *Lond.* **206**, 415–439 (1985).
- 673 59. Boudreau, B. P. & Ruddick, B. R. On a reactive continuum representation of organic matter  
674 diagenesis. *Am. J. Sci.* **291**, 507–538 (1991).
- 675 60. Forney, D. C. & Rothman, D. H. Inverse method for estimating respiration rates from decay  
676 time series. *Biogeosciences* **9**, 3601–3612 (2012).
- 677 61. Leifeld, J. & von Lützow, M. Chemical and microbial activation energies of soil organic  
678 matter decomposition. *Biol. Fertil. Soil.* **50**, 147–153 (2014).

679

680 **Extended Data Legends:**

681 **Extended Data Figure 1 | Sample locations.** Marine/riverine dissolved OC (black circles), soil  
682 OC (white squares), riverine particulate OC (red triangles), and marine sediment OC (blue  
683 diamonds). See **Extended Data Table 1** for GPS coordinates, sample collection information,  
684 and original publication references.

685

686 **Extended Data Figure 2 | Intra-sample <sup>14</sup>C age variances.** All individual mineral-containing  
687 samples (soil, white squares; river POC, red triangles; marine sediment OC, blue diamonds)  
688 plotted against %OC reveal a globally coherent power-law relationship. *r* is the reduced major  
689 axis correlation coefficient. Age variance is calculated as the sample variance of <sup>14</sup>C ages for all  
690 RPO fractions within a given sample (**Eq. 4**). Age variance and %OC are presented on  
691 logarithmic scales. Uncertainty ( $\pm 1\sigma$ ) is smaller than the marker for all data points. The %OC  
692 axis is reversed to emphasize that OC content generally decreases with time.

693

694 **Extended Data Table 1 | Sampling locations, sample collection details, screening procedure**  
695 **results, %OC, Fm, and original reference for all samples.**

696 Footnotes:

697 \*See Methods for screening procedure details

698 †Samples that originally contain CO<sub>3</sub> but have been acidified prior to RPO analysis are  
699 marked ‘n’

700 ‡Empty cells represent missing or nonexistent data

701 §If no uncertainty is reported, then std. dev. is taken to be the maximum of 0.01 or 1 % of  
702 measured %OC

703 ¶If no uncertainty is reported, then std. dev. is taken to be 0.005 Fm units

704 ¶¶P = Peat; E = Entisol; UB = Udic Boroll; A = Andisol

705

706 **Extended Data Table 2 |  $\mu_E$  and  $\sigma_E$  values for samples used to determine methodological**  
707 **uncertainty.**