

# Dark Matter Enhances Interactions within Both Microbes and Dissolved Organic Matter under Global Change

Ang Hu,<sup>▽</sup> Fanfan Meng,<sup>▽</sup> Andrew J. Tanentzap, Kyoung-Soon Jang, and Jianjun Wang\*



Cite This: *Environ. Sci. Technol.* 2023, 57, 761–769



Read Online

ACCESS |



Metrics & More



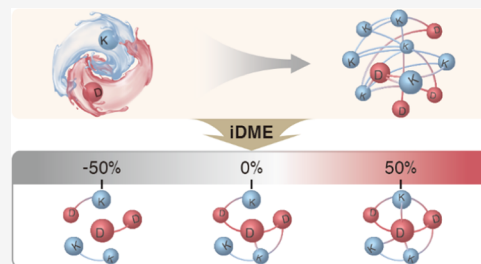
Article Recommendations



Supporting Information

**ABSTRACT:** There are vast but uncharacterized microbial taxa and chemical metabolites (that is, dark matter) across the Earth's ecosystems. A lack of knowledge about dark matter hinders a complete understanding of microbial ecology and biogeochemical cycles. Here, we examine sediment bacteria and dissolved organic matter (DOM) in 300 microcosms along experimental global change gradients in subtropical and subarctic climate zones of China and Norway, respectively. We develop an indicator to quantify the importance of dark matter by comparing co-occurrence network patterns with and without dark matter in bacterial or DOM assemblages. In both climate zones, dark matter constitutes approximately 30–56% of bacterial taxa and DOM metabolites and changes connectivity within bacterial and DOM assemblages by between  $-15.5$  and  $+61.8\%$ . Dark matter is generally more important for changing network connectivity within DOM assemblages than those of microbes, especially in the subtropical zone. However, the importance of dark matter along global change gradients is strongly correlated between bacteria and DOM and consistently increased toward higher primary productivity because of increasing temperatures and nutrient enrichment. Our findings highlight the importance of microbial and chemical dark matter for changing biogeochemical interactions under global change.

**KEYWORDS:** dark matter, microbes, dissolved organic matter, global change



## INTRODUCTION

Global change is altering Earth's biogeochemical cycles by rewiring the links between microbes and organic matter.<sup>1–4</sup> However, <1% of bacterial species have been cultured, and fewer than half of the identifiable major bacterial lineages or phyla include cultivated representatives.<sup>5</sup> Similarly, a large proportion of natural organic matter remains uncharacterized.<sup>6</sup> For instance, in sediments and water of global rivers, only 8.7 and 9.6% of 50,942 and 48,392  $m/z$  peaks of dissolved organic matter, measured using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), could be assigned to identifiable molecular formulae, respectively<sup>7,8</sup> (Table S1; the results from this study). These unknown majorities are colloquially called biological and chemical “dark matter,” like the as-yet-undetected mass in the universe.<sup>9</sup> Recent advances in next-generation sequencing technologies and ultrahigh-resolution mass spectrometry are now enabling new insights into the composition of microbes<sup>10,11</sup> and organic matter,<sup>12,13</sup> respectively. The unknown microbial taxa likely represent major evolutionary lines within the Tree of Life that are expected to play key ecological roles in their communities and environments.<sup>14</sup> Thus, revealing the ecological roles of biological dark matter alongside chemical dark matter can help unravel the full extent of biogeochemical cycles and their responses to global change. However, few approaches exist to quantify these dark matter effects.

Here, we assessed how dark matter influenced microbial communities and dissolved organic matter (DOM) assemblages. We did so by developing an indicator to quantify the effect of dark matter on taxa–taxa and metabolite–metabolite interactions, respectively (Figure 1). This indicator of dark matter effects (iDME) was based on ecological co-occurrence networks, which are a powerful tool to understand ecological roles and interactions within biological communities.<sup>15,16</sup> Network metrics such as degree can then be used to describe quantitatively the interconnectivity of microbial taxa or metabolites within an assemblage.<sup>17,18</sup> In each network, the nodes represent individual microbial taxa (or metabolites), and the edges identify the taxa–taxa (or metabolite–metabolite) interactions. The unknown (i.e., “dark”) nodes were identified as the taxa and metabolites that could not be assigned microbial taxonomy or molecular formulae, respectively. We first built two types of co-occurrence networks with either only known nodes (“KK” networks) or both dark and known nodes (“DK” networks) and then quantified the magnitude and direction (i.e., positive and negative) of the iDME as

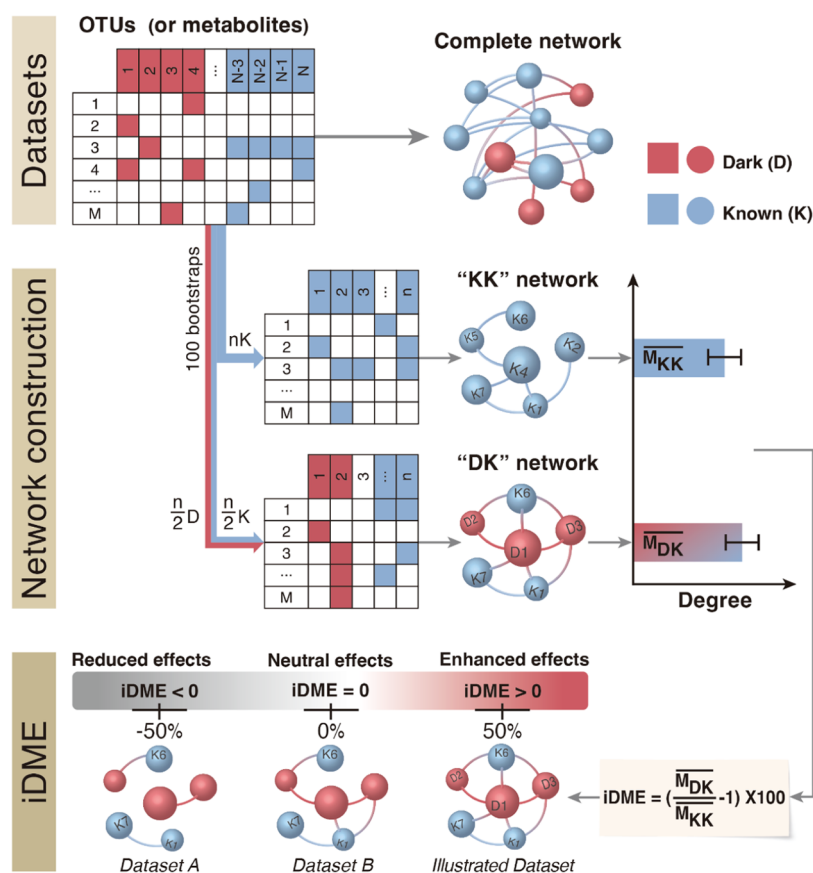
**Received:** July 13, 2022

**Revised:** November 28, 2022

**Accepted:** November 28, 2022

**Published:** December 14, 2022





**Figure 1.** Development of an indicator of dark matter effects (iDME) on microbial and metabolite networks. The indicator was developed based on ecological co-occurrence networks. In a network, nodes represent individual microbial OTUs or organic matter metabolites (i.e.,  $m/z$  peaks), and edges identify the interactions among OTUs or metabolites. For microbes, any uncultured, unassigned, or ambiguous OTUs at the genus level were designated as microbial dark matter. For organic matter,  $m/z$  peaks that could not be assigned to elemental formulae with a combination of C, H, O, N, P, and S were designated as elemental dark matter. The network between all known (K, blue circles) and dark (D, coral circles) nodes for either microbial or organic matter assemblages was considered the “Complete network.” There were then three primary procedures in developing the iDME. First, we built two types of reduced co-occurrence networks with either only known nodes (“KK” network) or both dark and known nodes (“DK” network). These two networks had the same number of randomly subsampled nodes from the entire microbial OTU or organic matter metabolite pool and were further bootstrapped 100 times. Second, we quantified network interaction metrics  $M$ , such as degree for each “KK” or “DK” network. Finally, we calculated the magnitude and direction of the iDME as the percentage change in the mean of network degree between “KK” ( $\overline{M}_{KK}$ ) and “DK” ( $\overline{M}_{DK}$ ) networks. Positive and negative iDME values indicate that dark matter enhances and reduces the network interactions, respectively, while an iDME of zero suggests a neutral effect. Larger absolute iDME values indicate dark matter has a greater potential influence on the known assemblages. In dataset A and the Illustrated dataset, iDME equals  $-50$  and  $+50\%$ , indicating that dark matter reduced and enhanced 50% of network interactions, respectively.

percentage change in the network metrics such as degree<sup>17,18</sup> between “DK” and “KK” networks. Positive and negative iDME values indicate that dark matter enhances and reduces network interactions within microbial communities or organic matter assemblages, respectively. We further evaluated the importance of bacterial and metabolite dark matter along experimental global change gradients in two contrasting climate zones.<sup>3,19</sup> Specifically, we established 300 aquatic microcosms composed of natural lake sediments and artificial lake water at five elevations on subtropical and subarctic mountainsides in China and Norway. At each elevation, we set up ten nutrient levels ranging from 0 to 36 mg N L<sup>-1</sup> in the overlying water.<sup>19</sup> This experiment allows us to examine the importance of dark matter across temperature and nutrient gradients in a simplified ecosystem under field conditions. The sediment bacteria and DOM assemblages were examined using high-throughput sequencing of 16S rRNA genes<sup>19</sup> and ultrahigh-resolution electrospray ionization FT-ICR MS,<sup>3</sup> respectively.

With this experiment, we aimed to examine the effects of dark matter on connectivity within bacterial or DOM assemblages using the indicator iDME and their responses to global change drivers. We showed that this novel indicator is robust and sensitive to quantify the importance of dark matter within microbes or DOM across contrasting climatic zones. We also revealed that dark matter consistently changes network connectivity within bacteria and DOM, and increasingly so under global change scenarios.

## ■ MATERIAL AND METHODS

**Experimental Design, Sample Collection, and Analyses.** The comparative field microcosm experiments were conducted in a subtropical region, Laojun Mountain in China (26.6959 N; 99.7759 E), and in a subarctic region, Balggesvarri Mountain in Norway (69.3809 N; 20.3483 E), in September–October and July 2013, respectively. The experimental design was first reported by Wang et al.<sup>19</sup> Briefly, we selected locations with five elevations on each mountainside. The

elevations were 3822, 3505, 2915, 2580, and 2286 m above sea level (m a.s.l.) on Laojun Mountain in China, and 750, 550, 350, 170, and 20 m a.s.l. on Balggesvarri Mountain in Norway. At each elevation, we established 30 aquatic microcosms (1.5 L bottle) composed of 15 g of sterilized lake sediments and 1.2 L of sterilized artificial lake water, which included ten nutrient levels of 0, 0.45, 1.80, 4.05, 7.65, 11.25, 15.75, 21.60, 28.80, and 36.00 mg N L<sup>-1</sup> of KNO<sub>3</sub> in the overlying water. Surface lake sediments less than 10 cm deep were obtained from the center of Taihu Lake, China, and were aseptically canned per bottle after autoclaving, as previously described in Wang et al.<sup>19</sup> To compensate for nitrate additions shifting stoichiometric ratios, KH<sub>2</sub>PO<sub>4</sub> was added to the bottles so that the N/P ratio of the initial overlying water was 14.93, which was similar to the annual average ratio in Taihu Lake during 2007 (that is, 14.49). It should be noted that we used “nutrient enrichment” to indicate a series of designed nutrient levels of both nitrate and phosphate, the former of which was used to represent nutrient enrichment in the statistical analyses due to the consistent N/P ratio. There were three replicated microcosms of each nutrient level at each elevation. The microcosms were left in the field for 1 month, allowing airborne microbes to freely colonize the sediments and water.

At the end of the experimental period, we aseptically sampled the overlying water and sediments of the 300 bottles (that is, 2 mountains × 5 elevations × 10 nutrient levels × 3 replicates) for the following analyses of bacterial communities and DOM composition. Details on field experiments, sample collection, environmental variables, and bacterial and DOM analyses were described in Wang et al.<sup>19</sup> and Hu et al.<sup>3</sup> Briefly, primary productivity variables such as water pH and sediment chlorophyll *a* were measured for each sample. Sediment bacteria were examined using high-throughput sequencing of 16S rRNA genes. The sequences were processed in QIIME (v1.9),<sup>20</sup> and operational taxonomic units (OTUs) were defined at 97% sequence similarity. Representative sequences from each OTU were aligned to the SILVA (v128) reference database<sup>21</sup> using PyNAST.<sup>22</sup> The taxonomic identity of each representative sequence was determined using the RDP Classifier,<sup>23</sup> and chloroplasts were removed. The bacterial sequences were rarefied to 20,000 per sample.

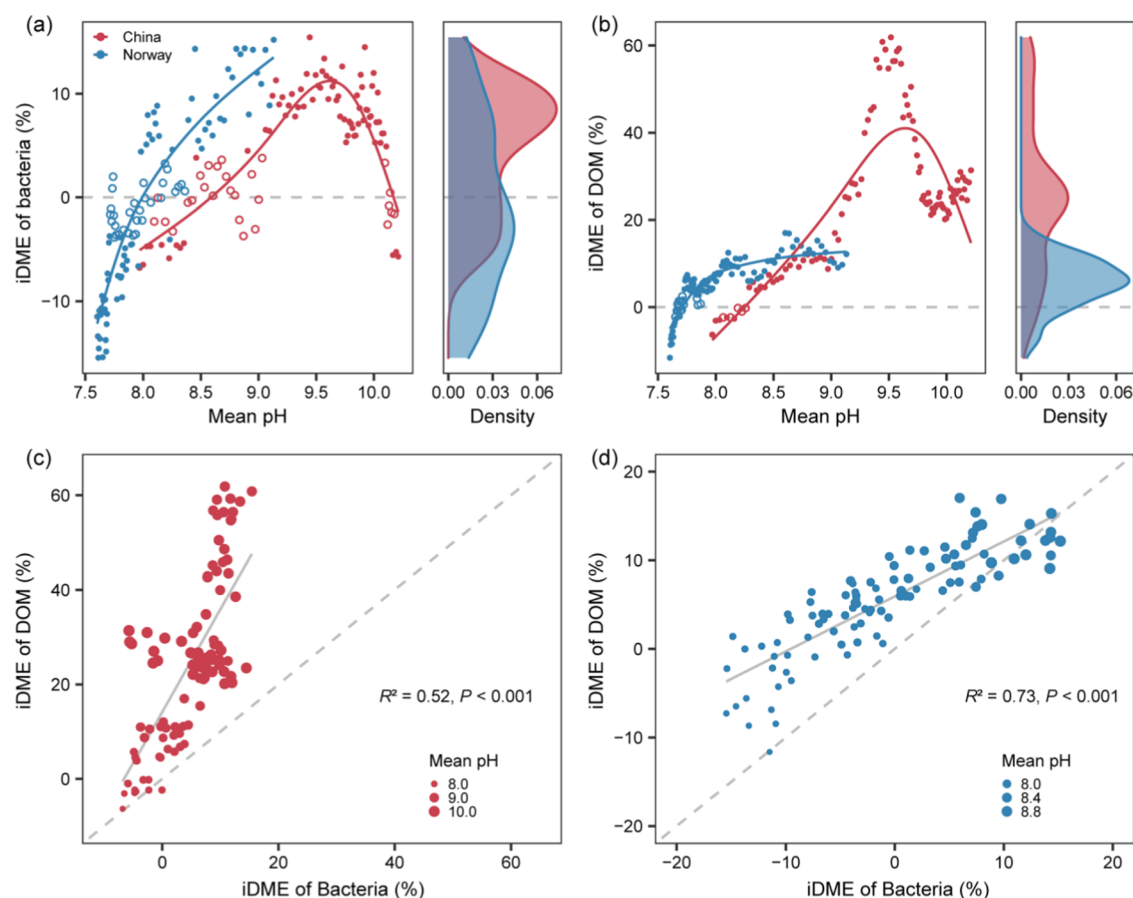
DOM within the sediment samples was solid-phase extracted (SPE) for FT-ICR MS measurement<sup>24</sup> with some modifications. Briefly, an aliquot of 0.7 g freeze-dried sediment was sonicated with 30 mL of ultrapure water for 2 h and centrifuged at 5000g for 20 min. To minimize cellular carbon release during sonication, we used flow-through tap water to maintain the consistent temperature of extracts. The sonication treatment combined with SPE method can increase the dissolution rates and extraction efficiency of DOC in a shorter period and thus minimize DOC decomposition during extraction. The extracted water was filtered through a 0.45 μm Millipore filter and further acidified to the pH of 2 using 1 M HCl. Cartridges were drained, rinsed with ultrapure water and methanol (ULC-MS grade), and conditioned with ultrapure water of pH 2. Calculated volumes of the extracts were slowly passed through the cartridges based on DOC concentration. The cartridges were rinsed with ultrapure water of pH 2 and dried with N<sub>2</sub> gas. Samples were finally eluted with methanol into precombusted amber glass vials, dried with N<sub>2</sub> gas, and stored at -20 °C until DOM analysis. Over 60% of DOM was expected to be recovered from the SPE procedure.<sup>24,25</sup> Highly accurate mass measurements of DOM

were conducted using a 15 Tesla solarix XR ultrahigh-resolution FT-ICR MS (Bruker Daltonics, Billerica, MA). The FT-ICR MS was coupled to an electrospray ionization interface, as demonstrated previously,<sup>26</sup> with some modifications. Data Analysis software (Bruker Daltonics version 4.2) was used to convert raw spectra to a list of *m/z* values using FT-MS peak picker with a signal-to-noise ratio threshold set to 7 and absolute intensity threshold set to the default value of 100. Putative chemical formulae were assigned using an in-house software Formularity,<sup>27</sup> following the Compound Identification Algorithm (CIA).<sup>28</sup> Briefly, CIA adopts established principles of formula assignment: peaks from spectra measured by FT-ICR MS are assigned with molecular formulae starting from the low *m/z* range searching CIA database, and high *m/z* compounds are assigned using formula expansion based on functional groups (e.g., CH<sub>2</sub>, H<sub>2</sub>, or O) because the number of formula candidates increases substantially as mass increases.<sup>27,28</sup>

We defined dark matter of bacteria as any uncultured, unassigned, or ambiguous OTUs (hereafter, dark OTUs) at any taxonomic level, such as the genus level. Sufficient known and dark nodes, such as >200 per environmental condition, are required to build the networks. At the genus level, 409 known and 438 dark OTUs in China, and 241 known and 308 dark OTUs in Norway were detected in more than 30% of the total samples (Table S1). Thus, genus level was used in the following analyses (Figure S1). We defined dark matter of DOM as any unassigned metabolites (i.e., *m/z* peaks) that could not be assigned to elemental formulae with a combination of C, H, O, N, P, and S (hereafter, dark metabolites or peaks).<sup>6</sup> In summary, 2118 known and 885 dark metabolites in China, and 1670 known and 922 dark metabolites in Norway were detected in more than 30% of the total samples (Table S1).

**Calculating the Importance of Dark Matter.** We developed the indicator iDME to assess how dark matter influenced microbial communities and organic matter assemblages. The iDME quantifies the effect size of dark matter on OTU–OTU (or metabolite–metabolite) interactions between the presence and absence of dark matter. Putative interactions between microbial OTUs (or DOM metabolites) were quantified using co-occurrence network analysis. There were three primary procedures in developing the iDME (Figure 1).

First, we built two types of co-occurrence networks, that is, “KK” and “DK” networks. In each network, nodes represent individual bacterial OTUs or DOM metabolites, and edges identify the interactions among OTUs or metabolites. These two networks had an identical number of nodes that were randomly subsampled from the whole microbial OTU or organic matter metabolite pool and were further bootstrapped 100 times. The determination of node number could be found in the section of Robustness test of iDME. The co-occurrence network was inferred based on the SparCC (Sparse Correlations for Compositional data)<sup>29</sup> correlation matrix constructed with the R package SpiecEasi V1.0.7.<sup>30</sup> SparCC is based on the log-ratio transformation, such as the ratios of the fractions of two OTUs, which avoids the potential issue of spurious correlations. The rarefied OTU table or normalized metabolite table with signal intensities were used for calculating SparCC correlations. Bacterial OTUs or DOM metabolites observed in more than 30% of the total samples in China or Norway were retained for correlation calculations.



**Figure 2.** The importance of dark matter in microbial communities and DOM assemblages under global change. We quantified the indicator of dark matter effects (iDME) for network metric of degree on bacterial OTUs and DOM metabolites along a pH gradient in 101 windows with a fixed size of 50 samples. We plotted iDME against pH for bacteria (a) and DOM (b) in China (red) and Norway (blue). Solid and hollow points indicate iDME values were statistically ( $P \leq 0.05$ ) and nonstatistically ( $P > 0.05$ ) different from zero, respectively. Lines are generalized additive models with 4 knots. We also show the density distribution of iDME for bacteria (a) and DOM (b) in the two regions. The relationships between iDME of bacteria and DOM in China (c) and Norway (d) were visualized with linear regression models, and statistically significant model fits were indicated by solid lines ( $P \leq 0.05$ ). Dashed gray lines mark the 1:1 relationship. The mean pH in each window was indicated by dot size.

The threshold value of SparCC  $\rho$  correlations for generating co-occurrence networks was  $|\rho| = 0.30$  to filter the uncorrelated or weakly correlated interactions. Second, we measured network metrics ( $M$ ), such as degree, to quantify the connectivity for each “KK” or “DK” network with the R package igraph V1.2.6. Degree is defined as the number of edges that connects a focal node to other nodes.<sup>18</sup> The OTUs or metabolites with a higher degree are more interconnected within an assemblage. Finally, we calculated the magnitude and direction of the iDME, which is defined as a percentage change in the mean value of a given network metric  $M$  between “KK” and “DK” networks

$$\text{iDME (\%)} = \left( \frac{\overline{M_{\text{DK}}}}{\overline{M_{\text{KK}}}} - 1 \right) \times 100$$

Positive and negative iDME values indicate that dark matter enhances and reduces network interactions within microbial communities or organic matter assemblages, respectively, while zero iDME suggests a neutral effect. Larger absolute iDME values indicate dark matter has a greater potential influence on the known assemblages. It should be noted that percentage change can be converted from another effect size measure of

log response ratio (LRR) which is usually used in non-normalized data: percentage change =  $(e^{\text{LRR}} - 1) \times 100$ .

**Robustness Test of iDME.** To test the robustness of the iDME of the network degree, we determined how the network size and the ratio of dark/known nodes influenced the indicator performance. We randomly selected 50 samples out of 150 samples in each region with 20 bootstraps, and the bacterial OTUs or DOM metabolites observed in more than 30% of the total samples were retained for correlation calculations. The 50 samples were selected so that the sample number was consistent with the other statistical analysis (that is, the sample number in each window of the moving-window analysis). For each network size, we considered two ratios of dark/known nodes, that is, the 1:1 ratio and the observed ratio in an assemblage. For either ratio of dark/known nodes, we reconstructed networks that contained between 100 and 200 nodes in 20-node increments for bacteria and between 100 and 1000 nodes in 100-node increments for DOM. For bacteria, we used the maximum network size of 200 nodes, which was based on the number of known OTUs retained in the above bootstraps, respectively (Figure S2). For DOM, we used the maximum size of 1000 nodes for network analysis because there are computational limits on handling large networks (e.g., >1000 nodes) for the robustness test. The inconsistent

increments between DOM and bacteria will not affect our main conclusions as we were examining the overall trends toward large network size. More details are shown in the [Results and Discussion](#) section.

**iDME Partitioning Analysis.** We further disentangled whether the effects of dark matter were due to changes in links between dark–dark nodes or dark–known nodes. To do so, we partitioned the iDME into intra-iDME and inter-iDME, which were defined as the percentage change in the number of dark–dark and dark–known links, respectively

$$\text{intra-iDME(\%)} = \left( \frac{\overline{M_{DD}} - M_{K_2 K_2}}{\overline{M_{K_1 K_1}} + M_{K_2 K_2} + M_{K_1 K_2}} \right) \times 100$$

$$\text{inter-iDME(\%)} = \left( \frac{\overline{M_{DK_1}} - M_{K_1 K_2}}{\overline{M_{K_1 K_1}} + M_{K_2 K_2} + M_{K_1 K_2}} \right) \times 100$$

where “ $K_1$ ” are the half-selected nodes from “KK” network for “DK” network, “ $K_2$ ” are the remaining nodes in “KK” network, “D” are the dark nodes for “DK” network.  $M_{K_1 K_1} + M_{K_2 K_2} + M_{K_1 K_2}$  are the total number of links in “KK” network (i.e.,  $M_{KK}$ ) in each bootstrap;  $M_{K_2 K_2}$  and  $M_{K_1 K_2}$  are the number of intra- and inter-links in “KK” network in each bootstrap, respectively.  $M_{DD}$  and  $M_{DK_1}$  are the number of intra- and inter-links in “DK” network in each bootstrap, respectively.

**iDME along the Global Change Gradient.** To study further potential drivers of variation in the iDME of network degree, we used the proxy variable of primary productivity (i.e., water pH) to represent the joint outcome of rising temperature and nutrient enrichment that characterize global change, as previously described in Hu et al.<sup>31</sup> We used water pH as an easily measured *in situ* proxy for primary productivity, as opposed to laboratory-based measurements that were more removed from the field system. This is because the primary productivity of benthic algae (i.e., chlorophyll *a*) showed a strong positive correlation ( $R^2 = 0.65$  to  $0.84$ ,  $P < 0.001$ ) with water pH at almost all nutrient levels and elevations due to depleting dissolved  $\text{CO}_2$  upon algal growth.<sup>19</sup> It should be noted that we also tested sediment chlorophyll *a*, and it showed similar patterns in the iDME to water pH (Figure S3). To be consistent with our previous studies, we would like to show the results relevant to the pH gradient in the main text. In addition, we observed that water pH, that is, a proxy representing the joint results of kinetic (i.e., temperature) and potential (i.e., nutrients) energy supply, was more important (all  $P \leq 0.05$ ; Figure 2a,b) in driving the iDME than either type of energy gradient on its own. This result was supported by the observation that the iDME generally did not vary with either temperature or nutrient (most  $P > 0.05$ ; Figure S4).

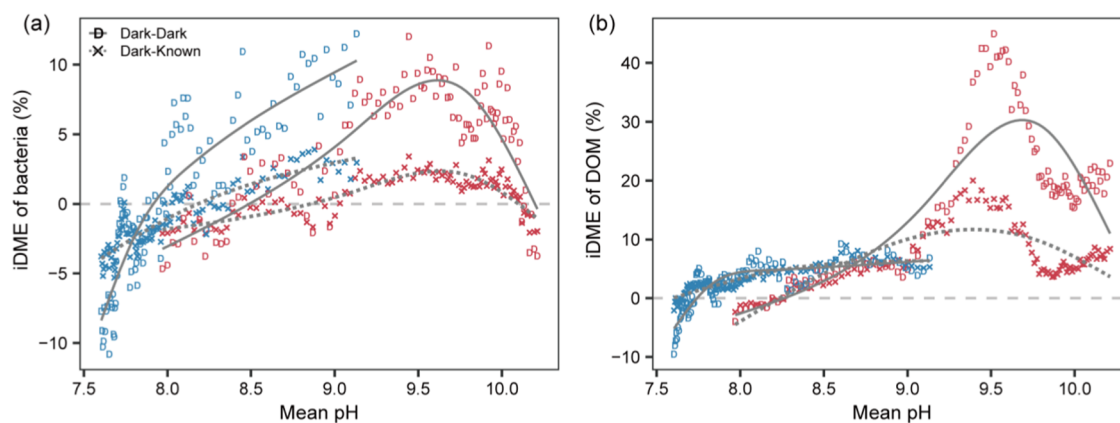
To test the effects of global change on the importance of dark matter in microbial or organic matter networks, we used a moving-window approach,<sup>32,33</sup> as it can identify continuous and/or sharp transitions in dark matter effects along a continuum. We first sorted the samples along the energy supply gradient, from minimum to maximum water pH, separately for China and Norway. We used one-third of the samples (that is, 50) as the window size, generating 101 windows (e.g., 1–50, 2–51, ..., 101–150 consecutive samples). We then calculated the mean pH for each window, resulting in a pH gradient ranging from 8.0 to 10.2 in China (gradient across individual samples from 7.5 to 10.8) and a

gradient from 7.6 to 9.1 in Norway (full gradient from 7.4 to 10.0).<sup>31</sup> For each window, we calculated the remaining bacterial OTUs and DOM metabolites observed in more than 30% of the total samples in each pH window (Figure S5). To ensure comparable estimates of iDME across the bacterial and DOM datasets and across the pH gradients, we constructed each network with a consistent setting, i.e., a network size of 200 nodes and a 1:1 ratio of dark and known nodes. To confirm the results of 1:1 ratio of dark and known nodes, we also considered the observed ratio for calculating the iDME. Our results showed similar patterns in the iDME for the two ratios of dark/known nodes along the pH gradient, indicated by strongly positive linear regressions ( $R^2 = 0.84$  to  $0.96$ ,  $P < 0.001$ ; Figure S6c,f).

## RESULTS AND DISCUSSION

Dark matter constituted 51.7 and 56.1% of bacterial OTUs, and 29.5 and 35.6% of DOM *m/z* peaks for China and Norway, respectively (Table S1). To test the robustness of the iDME, we determined how network size and the ratio of dark/known nodes influenced indicator performance. We reconstructed networks that contained between 100 and 200 nodes in 20-node increments for bacteria and between 100 and 1000 nodes in 100-node increments for DOM. For each network size, we considered two ratios of dark/known nodes, that is, a 1:1 ratio and the observed ratio in an assemblage. The latter ratio values varied from 1.01 to 1.38 for bacteria and from 0.33 to 0.61 for DOM across the global change gradients (Figure S5). In most cases, values of the iDME using the network metric of degree slightly changed between  $-0.07$  and  $0.31$  times with increasing network size (Figure S7). The changes also progressively slowed for DOM in Norway, with an increase of 5.26 times between 100 to 1000 nodes (Figure S6). In addition, iDME for the two ratios of dark/known nodes showed similar patterns with increasing network size and were strongly positively correlated ( $R^2 = 0.53$  to  $0.91$ ,  $P < 0.001$ ; Figure S7). These results indicate that iDME was more robust toward larger network sizes but was relatively insensitive to the ratio of dark/known nodes, even when it differed considerably as for DOM. Nonetheless, to ensure comparable estimates of iDME across the bacterial and DOM datasets and the environmental gradients, we quantified dark matter effects with a consistent setting, i.e., a network size of 200 nodes and a 1:1 ratio of dark and known nodes.

We found that dark matter substantially changed the network connectivity of both bacteria and DOM across experimental global change gradients. In both study regions, 71.3 and 89.6% of iDME for network metric of degree were significantly different from zero for bacteria and DOM. The iDME changed by  $-15.5$  to  $+15.4\%$  and from  $-11.6$  to  $+61.8\%$  in bacteria and DOM across the gradients, respectively (Figure 2a,b). Such large variation in the iDME values indicates that it was sensitive to different environmental contexts (Figure 2a,b). Interestingly, our results highlight that dark matter often increased the network connectivity for both bacteria and DOM, indicated by a positive iDME in 58.4 and 86.6% of cases, respectively. Previous studies have shown that microbial and chemical “dark matter” represents a major challenge for the exploration of global microbiome and metabolomes.<sup>9,34</sup> We however have limited knowledge about the dark matter of DOM, not mentioning the approaches to quantify these dark matter effects. Our iDME now offers, for the first time, a way to quantify the importance of microbial and metabolite dark



**Figure 3.** The partitioning of indicator of dark matter effects (iDME) within microbial and DOM assemblages under global change. We partitioned iDME into intra- and inter-iDME based on dark–dark links (solid lines) and dark–known links (dotted lines) along a pH gradient in 101 windows with a fixed size of 50 samples. We plotted intra- and inter-iDME against pH for bacteria (a) and DOM (b) in China (red) and Norway (blue). Lines are generalized additive models with 4 knots.

matter under global change such as temperature increase and nutrient enrichment. More importantly, by including a method to calculate the effect size, iDME can be compared between the biosphere and chemosphere and across contrasting climatic zones.

To test the effects of global change on the importance of dark matter in the microbial or organic matter assemblages, we compared iDME of network degree across a pH gradient. We used a moving-window analysis to identify continuous and/or sharp transitions.<sup>32,33</sup> We found that the iDME for bacteria and DOM increased from negative to positive values along the pH gradient of 7.5–9.5 in both regions (Figure 2a,b). This result suggests that the dark matter can promote network connectivity for both bacteria and DOM as primary productivity increases. When primary productivity reached a very high level, such as at pH > 9.5 in China, the dark matter effects for DOM declined but remained positive (Figure 2b), while the effects for bacteria decreased to zero and eventually became negative (Figure 2a). This consistent pattern in dark matter effects along the primary productivity gradient is well supported by the strong correlation between iDME values for bacteria and DOM with  $R^2$  of 0.52 and 0.73 for China and Norway, respectively ( $P \leq 0.001$ ; Figure 2c,d). The positive contribution of dark matter at higher primary productivity, to a certain extent, agrees with a previous report showing that the exclusion of unknown taxa could significantly reduce bacterial network interactions in diverse extreme aquatic habitats such as hot springs and deep sea.<sup>14</sup>

Collectively, these results highlight that dark matter is similarly important between the biosphere and chemosphere and across contrasting climatic zones, and these effects were enhanced under global change scenarios of increasing temperatures and nutrient enrichment. This congruency would be expected as microbes are a primary driver of ecosystem metabolite transformations, and microbial composition is closely associated with organic matter composition.<sup>3,35–37</sup> Thus, as bacteria became dominated by dark matter, DOM also became dominated by dark matter and vice versa.

We further disentangled whether the effects of dark matter were due to changes in links between dark–dark nodes or dark–known nodes. To do so, we partitioned the iDME into intra-iDME and inter-iDME, which were defined as the percentage change in the number of dark–dark and dark–

known links, respectively. We found that for both bacteria and DOM, the intra-iDME was larger than the inter-iDME toward higher primary productivity, with the highest differences at pH 9.4–9.5 in China and 8.6–9.1 in Norway (Figure 3). These results indicate that the dark matter created its own distinct clusters (i.e., dark–dark links) and was also intermixed with known nodes (i.e., dark–known links), the former of which were consistent between bacteria and DOM and became more dominant toward higher primary productivity. Given the importance of such self-organized clusters among dark matter for network interactions in high primary productivity environments, the structure and functions of these dark matter clusters must be better characterized, particularly under future global change scenarios.

Notably, dark matter was generally more important for organic matter than bacteria and more so in the subtropical region. Specifically, values for the iDME of network degree were larger (ANOVA test,  $P < 0.001$  for both regions) and more positive for DOM than bacteria, with a mean value of +15.0 and +2.1%, respectively (Figure 2a,b). The larger effects of dark matter on DOM were stronger in the subtropical region, supported by the higher linear regression slope of 2.17 between the iDME of bacteria and DOM than in the subarctic region (Figure 2c,d). This difference between the importance of dark matter for DOM versus bacteria also increased in magnitude along the global change gradients, that is, with pH, in the subtropical region (Figure S8). We offer three nonmutually exclusive explanations for these results. First, organic matter assemblages are mainly processed by microbes and thus change faster than microbial community composition.<sup>38</sup> The resulting transformations of chemical dark matter could then result in more connections among metabolites and larger iDME. Second, microbial communities are highly functionally redundant,<sup>39</sup> so the network connections of dark matter species can be replaced by other counterparts, which are sometimes “known”. For organic matter, however, the linkages between metabolites potentially reflect molecular transformations that occur along conserved enzymatic pathways, and thus they are less replaceable if the metabolites involved in these transformations are dark matter. Finally, these results may be due to the methodological differences in levels of classification or taxonomic groups for defining dark matter. For instance, we could define the dark matter of

bacteria at either OTU or other broad taxonomic levels and the dark matter of DOM at different mass accuracy, e.g., 15 Tesla versus 21 Tesla FT-ICR MS. Further, there is a challenging situation for both microbes and DOM as there may be numerous bacterial strains within a defined OTU or tens of molecules with a defined formula.

Our findings have important implications for understanding the full extent of biogeochemical cycles under global change. We found that dark matter consistently changed the connectivity of microbial communities and the ecosystem metabolites they interact with, and increasingly so under global change scenarios. Thus, efforts to measure and predict biogeochemical cycles, especially under global change, are incomplete without revealing the ecological roles of microbial and chemical dark matter. At best, these efforts capture about two-thirds of the relevant ecological interactions. The role of chemical dark matter may be particularly challenging to reveal due to the faster turnover of DOM assemblages and the strong variation in dark matter effects of DOM across climate zones. It should be noted that microbial and organic matter data may be biased by their semi-quantitative nature due to the applied extraction methods and available analytical procedures. Further, some dark matter of bacteria and DOM are poorly captured by our approaches. For instance, a portion (10–40%) of DOM metabolites could leach from the solid-phase extracted processes<sup>24,25</sup> and should contain dark matter important for our understanding of microbial ecology and biogeochemical cycles. More advanced techniques like 21 Tesla FT-ICR MS and single-cell genetic analysis are now needed to help shine light on both the uncharacterized and uncovered chemical and microbial dark matter. Future studies on the indicator iDME is encouraged to examine how the degree to which microbial taxa and molecular formulae are identified may influence the actual effects of dark matter on microbial and DOM assemblages, largely because this identification is far from complete and unevenly distributed globally.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.est.2c05052>.

Additional information on the counts and percentage of known and dark bacterial OTUs at the genus level or DOM *m/z* peaks, counts of dark and known bacterial OTUs at different taxonomic levels, counts of dark and known bacterial OTUs (or DOM *m/z* peaks), changes in the iDME of network degree for bacteria and DOM along the gradients of sediment chlorophyll *a*, changes in the iDME of network degree for bacteria and DOM along the elevational and nutrient gradients, counts of dark and known bacterial OTUs (or DOM *m/z* peaks) and their ratio along a pH gradient, effects of global change on the iDME of network degree for bacteria and DOM, robustness test of the iDME of network degree for bacteria and DOM, and differences in the iDME of network degree between DOM and bacteria (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Jianjun Wang — State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology,

Chinese Academic of Sciences, Nanjing 210008, China; University of Chinese Academy of Sciences, Beijing 100049, China; [orcid.org/0000-0001-7039-7136](https://orcid.org/0000-0001-7039-7136); Email: [jjwang@niglas.ac.cn](mailto:jjwang@niglas.ac.cn)

### Authors

Ang Hu — College of Resources and Environment, Hunan Agricultural University, Changsha 410128, China; State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academic of Sciences, Nanjing 210008, China; [orcid.org/0000-0002-5755-2442](https://orcid.org/0000-0002-5755-2442)

Fanfan Meng — State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academic of Sciences, Nanjing 210008, China; University of Chinese Academy of Sciences, Beijing 100049, China

Andrew J. Tanentzap — Ecosystems and Global Change Group, School of the Environment, Trent University, Peterborough, Ontario K9L 0G2, Canada; Ecosystems and Global Change Group, Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom; [orcid.org/0000-0002-2883-1901](https://orcid.org/0000-0002-2883-1901)

Kyoung-Soon Jang — Bio-Chemical Analysis Team, Korea Basic Science Institute, Cheongju 28119, South Korea

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.est.2c05052>

### Author Contributions

<sup>‡</sup>A.H. and F.M. contributed equally to this paper. J.W. conceived the idea. J.W. carried out the field trips and provided the physiochemical and biological data. K.J. analyzed the DOM. F.M. performed the statistical analyses with the comments from A.H. and J.W. A.H. wrote the first draft of the manuscript. A.H. and J.W. finished the manuscript with the comments from A.T. All authors contributed to the intellectual development of this study.

### Notes

The authors declare no competing financial interest.

The scripts in calculating iDME are available in the R package iDOM at <http://github.com/jianjunwang/iDOM>. The microbial sequences and metadata generated in this study have been deposited in the MG-RAST database under accession code 17710 (Wang et al.).<sup>19</sup> Other data are available from the corresponding author upon reasonable request.

## ■ ACKNOWLEDGMENTS

The authors appreciate Lizhou Dai, Chengyan Zhang, Jinfu Liu, Xu Ma, Mira Choi, and Feiyun Pan for field sampling and laboratory analyses, Ji Shen, Qinglong Wu, and Yongqin Liu for kind supports, and Yunlin Zhang and Thorsten Dittmar for key introductions. This study was supported by the National Natural Science Foundation of China (42077052, 42225708, 92251304), the Second Tibetan Plateau Scientific Expedition and Research (STEP) Program (2019QZKK0503), Research Program of Sino-Africa Joint Research Center, Chinese Academy of Sciences (151542KYSB20210007), Scientific Research Fund of Hunan Provincial Education Department of China (21B0191), Science and Technology Planning Project of NIGLAS (NIGLAS2022GS09), and CAS Key Research Program of Frontier Sciences (QYZDB-SSW-DQC043). A.J.T. was supported by the Canada Research

Chairs Program and H2020 ERC Starting Grant (sEEInGDOM 804673). K.J. was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF-2021M1A5A1075510) and KBSI (C140440) grants.

## REFERENCES

- (1) Reich, P. B.; Hobbie, S. E.; Lee, T. D.; Rich, R.; Pastore, M. A.; Worm, K. Synergistic effects of four climate change drivers on terrestrial carbon cycling. *Nat. Geosci.* **2020**, *13*, 787–793.
- (2) Bradford, M. A.; McCulley, R. L.; Crowther, T. W.; Oldfield, E. E.; Wood, S. A.; Fierer, N. Cross-biome patterns in soil microbial respiration predictable from evolutionary theory on thermal adaptation. *Nat. Ecol. Evol.* **2019**, *3*, 223–231.
- (3) Hu, A.; Choi, M.; Tanentzap, A. J.; Liu, J.; Jang, K.-S.; Lennon, J. T.; Liu, Y.; Soininen, J.; Lu, X.; Zhang, Y.; Shen, J.; Wang, J. Ecological networks of dissolved organic matter and microorganisms under global change. *Nat. Commun.* **2022**, *13*, No. 3600.
- (4) Wang, J.; Hu, A.; Meng, F.; Zhao, W.; Yang, Y.; Soininen, J.; Shen, J.; Zhou, J. Embracing mountain microbiome and ecosystem functions under global change. *New Phytol.* **2022**, *234*, 1987–2002.
- (5) Rappé, M. S.; Giovannoni, S. J. The uncultured microbial majority. *Annu. Rev. Microbiol.* **2003**, *57*, 369–394.
- (6) Zark, M.; Christoffers, J.; Dittmar, T. Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. *Mar. Chem.* **2017**, *191*, 9–15.
- (7) Goldman, A. E.; Chu, R. K.; Danczak, R. E.; Daly, R. A.; Fansler, S.; Garayburu-Caruso, V. A.; Graham, E. B.; McCall, M. L.; Ren, H.; Renteria, L. *Environmental System Science Data Infrastructure for a Virtual Ecosystem, WHONDRS Summer 2019 Sampling Campaign: Global River Corridor Sediment FTICR-MS, NPOC, and Aerobic Respiration*, 2020.
- (8) Toyoda, J. G.; Goldman, A. E.; Chu, R. K.; Danczak, R. E.; Daly, R. A.; Garayburu-Caruso, V. A.; Graham, E. B.; Lin, X.; Moran, J. J.; Ren, H. *Environmental System Science Data Infrastructure for a Virtual Ecosystem, WHONDRS Summer 2019 Sampling Campaign: Global River Corridor Surface Water FTICR-MS and Stable Isotopes*, 2020.
- (9) da Silva, R. R.; Dorrestein, P. C.; Quinn, R. A. Illuminating the dark matter in metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 12549–12550.
- (10) Jiao, J.-Y.; Liu, L.; Hua, Z.-S.; Fang, B.-Z.; Zhou, E.-M.; Salam, N.; Hedlund, B. P.; Li, W.-J. Microbial dark matter coming to light: challenges and opportunities. *Natl. Sci. Rev.* **2021**, *8*, No. nwa280.
- (11) Solden, L.; Lloyd, K.; Wrighton, K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr. Opin. Microbiol.* **2016**, *31*, 217–226.
- (12) Kujawinski, E. B. High-Resolution Mass Spectrometry. In *Encyclopedia of Geochemistry: A Comprehensive Reference Source on the Chemistry of the Earth*; White, W. M., Ed.; Springer International Publishing, 2017; pp 1–5.
- (13) Kujawinski, E. B.; Freitas, M. A.; Zang, X.; Hatcher, P. G.; Green-Church, K. B.; Jones, R. B. The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Org. Geochem.* **2002**, *33*, 171–180.
- (14) Zamkovaya, T.; Foster, J. S.; de Crecy-Lagard, V.; Conesa, A. A network approach to elucidate and prioritize microbial dark matter in microbial communities. *ISME J.* **2021**, *15*, 228–244.
- (15) Banerjee, S.; Schlaeppi, K.; van der Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **2018**, *16*, 567–576.
- (16) Faust, K.; Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **2012**, *10*, 538–550.
- (17) Wasserman, S.; Faust, K. *Social Network Analysis: Methods And Applications*; Cambridge University Press, 1994.
- (18) Proulx, S. R.; Promislow, D. E.; Phillips, P. C. Network thinking in ecology and evolution. *Trends Ecol. Evol.* **2005**, *20*, 345–353.
- (19) Wang, J.; Pan, F.; Soininen, J.; Heino, J.; Shen, J. Nutrient enrichment modifies temperature-biodiversity relationships in large-scale field experiments. *Nat. Commun.* **2016**, *7*, No. 13960.
- (20) Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F. D.; Costello, E. K.; Fierer, N.; Pena, A. G.; Goodrich, J. K.; Gordon, J. I.; Huttley, G. A.; Kelley, S. T.; Knights, D.; Koenig, J. E.; Ley, R. E.; Lozupone, C. A.; McDonald, D.; Muegge, B. D.; Pirrung, M.; Reeder, J.; Sevinsky, J. R.; Tumbaugh, P. J.; Walters, W. A.; Widmann, J.; Yatsunenko, T.; Zaneveld, J.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336.
- (21) Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glockner, F. O. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **2012**, *41*, D590–D596.
- (22) Caporaso, J. G.; Bittinger, K.; Bushman, F. D.; DeSantis, T. Z.; Andersen, G. L.; Knight, R. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **2010**, *26*, 266–267.
- (23) Wang, Q.; Garrity, G. M.; Tiedje, J. M.; Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **2007**, *73*, 5261–5267.
- (24) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.: Methods* **2008**, *6*, 230–235.
- (25) Zou, C.; Li, M.; Cao, T.; Zhu, M.; Fan, X.; Peng, S.; Song, J.; Jiang, B.; Jia, W.; Yu, C.; Song, H.; Yu, Z.; Li, J.; Zhang, G.; Peng, P. Comparison of solid phase extraction methods for the measurement of humic-like substances (HULIS) in atmospheric particles. *Atmos. Environ.* **2020**, *225*, No. 117370.
- (26) Choi, J. H.; Jang, E.; Yoon, Y. J.; Park, J. Y.; Kim, T. W.; Becagli, S.; Caiazzo, L.; Cappelletti, D.; Krejci, R.; Eleftheriadis, K.; Park, K. T.; Jang, K. S. Influence of Biogenic Organics on the Chemical Composition of Arctic Aerosols. *Global Biogeochem. Cycle* **2019**, *33*, 1238–1250.
- (27) Tolić, N.; Liu, Y.; Liyu, A.; Shen, Y.; Tfaily, M. M.; Kujawinski, E. B.; Longnecker, K.; Kuo, L.-J.; Robinson, E. W.; Paša-Tolić, L.; Hess, N. J. Formularity: Software for Automated Formula Assignment of Natural and Other Organic Matter from Ultrahigh-Resolution Mass Spectra. *Anal. Chem.* **2017**, *89*, 12659–12665.
- (28) Kujawinski, E. B.; Behn, M. D. Automated Analysis of Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra of Natural Organic Matter. *Anal. Chem.* **2006**, *78*, 4363–4373.
- (29) Friedman, J.; Alm, E. J. Inferring Correlation Networks from Genomic Survey Data. *PLoS Comput. Biol.* **2012**, *8*, No. e1002687.
- (30) Kurtz, Z. D.; Müller, C. L.; Miraldi, E. R.; Littman, D. R.; Blaser, M. J.; Bonneau, R. A. Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Comput. Biol.* **2015**, *11*, No. e1004226.
- (31) Hu, A.; Jang, K.-S.; Meng, F.; Stegen, J.; Tanentzap, A. J.; Choi, M.; Lennon, J. T.; Soininen, J.; Wang, J. Microbial and Environmental Processes Shape the Link between Organic Matter Functional Traits and Composition. *Environ. Sci. Technol.* **2022**, *56*, 10504–10516.
- (32) Felipe-Lucia, M. R.; Soliveres, S.; Penone, C.; Fischer, M.; Ammer, C.; Boch, S.; Boeddinghaus, R. S.; Bonkowski, M.; Buscot, F.; Fiore-Donno, A. M.; Frank, K.; Goldmann, K.; Gossner, M. M.; Hölzel, N.; Jochum, M.; Kandeler, E.; Klaus, V. H.; Kleinebecker, T.; Leimer, S.; Manning, P.; Oelmann, Y.; Saiz, H.; Schall, P.; Schloter, M.; Schöning, I.; Schrumpf, M.; Solly, E. F.; Stempfhuber, B.; Weisser, W. W.; Wilcke, W.; Wubet, T.; Allan, E. Land-use intensity alters networks between biodiversity, ecosystem functions, and services. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 28140–28149.
- (33) Carlson, M. L.; Flagstad, L. A.; Gillet, F.; Mitchell, E. A. D. Community development along a proglacial chronosequence: are above-ground and below-ground community structure controlled more by biotic than abiotic factors? *J. Ecol.* **2010**, *98*, 1084–1095.

(34) Marcy, Y.; Ouverney, C.; Bik, E. M.; Lösekann, T.; Ivanova, N.; Martin, H. G.; Szeto, E.; Platt, D.; Hugenholtz, P.; Relman, D. A.; Quake, S. R. Dissecting biological “dark matter” with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11889–11894.

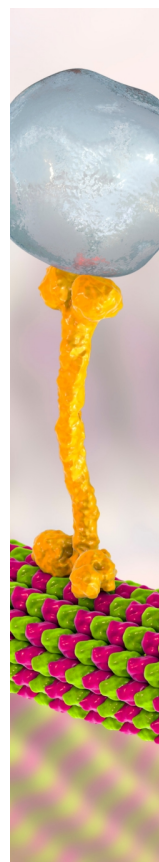
(35) Qin, S.; Chen, L.; Fang, K.; Zhang, Q.; Wang, J.; Liu, F.; Yu, J.; Yang, Y. Temperature sensitivity of SOM decomposition governed by aggregate protection and microbial communities. *Sci. Adv.* **2019**, *5*, No. eaau1218.

(36) Wieder, W. R.; Hartman, M. D.; Sulman, B. N.; Wang, Y. P.; Koven, C. D.; Bonan, G. B. Carbon cycle confidence and uncertainty: Exploring variation among soil biogeochemical models. *Global Change Biol.* **2018**, *24*, 1563–1579.

(37) Tanentzap, A. J.; Fitch, A.; Orland, C.; Emilson, E. J. S.; Yakimovich, K. M.; Osterholz, H.; Dittmar, T. Chemical and microbial diversity covary in fresh water to influence ecosystem functioning. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 24689.

(38) Danczak, R. E.; Chu, R. K.; Fansler, S. J.; Goldman, A. E.; Graham, E. B.; Tfaily, M. M.; Toyoda, J.; Stegen, J. C. Using metacommunity ecology to understand environmental metabolomes. *Nat. Commun.* **2020**, *11*, No. 6369.

(39) Louca, S.; Polz, M. F.; Mazel, F.; Albright, M. B. N.; Huber, J. A.; O'Connor, M. I.; Ackermann, M.; Hahn, A. S.; Srivastava, D. S.; Crowe, S. A.; Doebeli, M.; Parfrey, L. W. Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2018**, *2*, 936–943.



CAS BIOFINDER DISCOVERY PLATFORM™

## BRIDGE BIOLOGY AND CHEMISTRY FOR FASTER ANSWERS

Analyze target relationships,  
compound effects, and disease  
pathways

Explore the platform

