



## A network perspective reveals decreasing material diversity in studies on nanoparticle interactions with dissolved organic matter

Nicole Sani-Kast, Jérôme Labille, Patrick Ollivier, Danielle Slomberg, Konrad Hungerbühler, Martin Scheringer

### ► To cite this version:

Nicole Sani-Kast, Jérôme Labille, Patrick Ollivier, Danielle Slomberg, Konrad Hungerbühler, et al.. A network perspective reveals decreasing material diversity in studies on nanoparticle interactions with dissolved organic matter. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114 (10), pp.E1756-E1765. 10.1073/pnas.1608106114 . hal-01519532

**HAL Id: hal-01519532**

**<https://amu.hal.science/hal-01519532>**

Submitted on 13 Apr 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A network perspective reveals decreasing material diversity in studies on nanoparticle interactions with dissolved organic matter

Nicole Sani-Kast<sup>a</sup>, Jérôme Labille<sup>b,c</sup>, Patrick Ollivier<sup>d</sup>, Danielle Slomberg<sup>b,c</sup>, Konrad Hungerbühler<sup>a</sup>, and Martin Scheringer<sup>a,e,1</sup>

<sup>a</sup>Institute for Chemical and Bioengineering, Eidgenössische Technische Hochschule Zürich, CH-8093 Zurich, Switzerland; <sup>b</sup>Aix-Marseille Université, CNRS, Institut de Recherche pour le Développement, Collège de France, Centre Européen de Recherche et d'Enseignement de Géosciences de l'Environnement, 13545 Aix-en-Provence, France; <sup>c</sup>International Consortium for the Environmental Implications of Nanotechnology, 13545 Aix-en-Provence, France; <sup>d</sup>Bureau de Recherches Géologiques et Minières, Water Environment and Ecotechnologies Division, 45060 Orleans, France; and <sup>e</sup>Research Centre for Toxic Compounds in the Environment, Masaryk University, 602 00 Brno, Czech Republic

Edited by Marguerite A. Xenopoulos, Trent University, Peterborough, Canada, and accepted by Editorial Board Member David W. Schindler January 24, 2017 (received for review May 21, 2016)

**Dissolved organic matter (DOM) strongly influences the properties and fate of engineered nanoparticles (ENPs) in aquatic environments. There is an extensive body of experiments on interactions between DOM and ENPs and also larger particles. [We denote particles on the nano- and micrometer scale as particulate matter (PM).] However, the experimental results are very heterogeneous, and a general mechanistic understanding of DOM–PM interactions is still missing. In this situation, recent reviews have called to expand the range of DOM and ENPs studied. Therefore, our work focuses on the diversity of the DOM and PM types investigated. Because the experimental results reported in the literature are highly disparate and difficult to structure, a new format of organizing, visualizing, and interpreting the results is needed. To this end, we perform a network analysis of 951 experimental results on DOM–PM interactions, which enabled us to analyze and quantify the diversity of the materials investigated. The diversity of the DOM–PM combinations studied has mostly been decreasing over the last 25 y, which is driven by an increasing focus on several frequently investigated materials, such as DOM isolated from fresh water, DOM in whole-water samples, and TiO<sub>2</sub> and silver PM. Furthermore, there is an underrepresentation of studies into the effect of particle coating on PM–DOM interactions. Finally, it is of great importance that the properties of DOM used in experiments with PM, in particular the molecular weight and the content of aromatic and aliphatic carbon, are reported more comprehensively and systematically.**

nanoparticles | dissolved organic matter | environmental fate | network analysis | experimental design

The development and use of nanotechnology results in the release of engineered nanoparticles (ENPs) to the environment (1, 2). Much effort in current research is put into assessing the environmental risks of ENPs, which requires the understanding of the fate and distribution of ENPs in different environmental media, as well as their potential toxicity (3–6). Dissolved organic matter (DOM) (e.g., humic substances, polysaccharides, and proteins) is ubiquitous in aquatic systems and is known to interact with ENPs and modify their surface properties (7). DOM has been shown to either stabilize or destabilize aqueous ENP suspensions, depending on DOM properties and medium composition (8, 9). The effects of DOM on the stability of ENP suspensions are expected to strongly influence the mobility and distribution of ENPs in the environment (10).

The extraordinary variability in DOM effects on ENPs originates from the fact that the interactions between DOM and ENPs depend on numerous interrelated factors, such as pH and ionic strength of the media, the particles' surface properties and size, and the chemical composition and concentration of the

DOM (11). Some types of DOM, such as fulvic and humic acids, electrostatically stabilize ENPs via the adsorption of the electrically charged DOM on the ENP surface (12–16), where the strength of the stabilization is correlated (although not linearly) with the amount of DOM adsorbed (17) and varies with its chemical structure (11, 12, 18, 19). In other settings, particularly in the presence of divalent cations, DOM such as polysaccharides and polypeptides enhances the aggregation of ENPs by forming bridges between the particles (20–24).

The complexity and high diversity of the interactions between DOM and ENPs makes it very difficult to predict the effect of DOM on ENPs in aqueous dispersions (25, 26). Louie et al. (25) proposed to use empirical correlations between the different experiments as an alternative to a first-principles mechanistic approach. To achieve this goal, the authors called for, among others, a broader variety of the DOM and ENP types used in the experiments (25). The need for higher diversity of materials has also been expressed by others (26, 27): Philippe and Schaumann (26) reviewed a large body of experimental work regarding the effect of DOM on ENP and colloids and pointed out that a rather small number of DOM types (mainly humic substances) dominates the experimental settings, which indicates a need to study new DOM types (26). However, the diversity of materials

## Significance

Understanding the environmental fate of engineered nanoparticles (ENPs) is essential to the risk assessment of nanotechnology. Dissolved organic matter (DOM) strongly influences the behavior of ENPs in aqueous systems. This influence depends on many factors and, in particular, the types of ENPs and DOM. Accordingly, recent reviews have called repeatedly to expand the range of DOM and ENPs studied. Here, we use a network perspective to assess, in terms of the material types studied, the current state of research into DOM–ENP interactions. We identify global tendencies in the choice of materials and experimental approaches that have contributed to an undesirable discrepancy between the ever-increasing number of published experiments and the persisting call for investigation of more material types.

Author contributions: N.S.-K., K.H., and M.S. designed research; N.S.-K. performed research; N.S.-K., J.L., P.O., D.S., K.H., and M.S. analyzed data; and N.S.-K., J.L., P.O., D.S., K.H., and M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. M.A.X. is a Guest Editor invited by the Editorial Board.

<sup>1</sup>To whom correspondence should be addressed. Email: [scheringer@chem.ethz.ch](mailto:scheringer@chem.ethz.ch).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1608106114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1608106114/-DCSupplemental).

(i.e., the number and range of different DOM–ENP combinations studied) was not explicitly addressed and reasons for the current low diversity were not discussed.

In response to this call for higher diversity of the materials studied, we set out to explicitly determine the diversity of ENPs and DOM investigated. Like Philippe and Schaumann (26), we consider DOM effects on both ENPs and larger particulate matter (PM) and use the term PM to denote particles in the nano- and micrometer range. We define diversity as the ratio of the different DOM–ENP combinations studied vs. the number of experiments conducted. We maintain that higher diversity of PM–DOM combinations is a necessary condition for a better mechanistic understanding of the interactions between PM and DOM.

However, it is important to keep in mind that high diversity of materials alone is not sufficient. Other aspects of diversity such as investigations of the same PM–DOM combination under different conditions are not covered by our definition of diversity, which focuses on the diversity of the materials.

High diversity of the PM and DOM types offers an additional benefit. Because of the high complexity of DOM–PM interactions, it will be useful to use nonmechanistic approaches to elucidate relationships between experimental conditions and the fate of ENPs, as suggested by Louie et al. (25). For example, machine learning methods can be used to infer such relationships and develop so-called “empirical models” (28). Two requirements for the development of empirical models are consistency (i.e., all data points have the same set of descriptors that were measured by comparable methods) and diversity (i.e., the dataset contains a broad range of values for each descriptor). Consistency enables the inclusion of all information in a coherent manner, and diversity provides that the inferred correlation accounts for a wide range of experimental settings. Therefore, sufficient diversity of the DOM and ENP types studied is needed as a basis for empirical models of the interaction of ENPs with DOM.

One striking feature of the experimental field of PM–DOM interactions is the huge number of potential DOM and PM combinations that can be studied, which makes it difficult to obtain, through a literature search, an overall description of the untested DOM–PM combinations and of the frequency of already studied ones in the published experiments. In response to this combinatorial challenge, we present here a perspective that quantifies the diversity of the DOM–PM combinations in a large set of experiments studying the effect of DOM on PM in aquatic media. We organize the different types of PM and DOM in a network, which makes it possible to visualize and quantify the overall diversity of materials by analyzing the topological features of the network. With this approach, we identify (i) a long-term decrease in the materials’ diversity over the last 25 y; (ii) significantly more experiments that do not explicitly consider the PM’s initial coating vs. ones that do, which is contrary to recent recommendations that call for the assessment of PM fate in the context of the particles’ life cycle, which often includes stages where the materials are coated; and (iii) dominant employment of DOM from fresh water origin and infrequent use of DOM from other environments (e.g., soils and sediments).

Our analysis provides explanations of the discrepancy between the repeated call for an increase in materials’ diversity, on the one hand, and the ever increasing number of published experiments, on the other hand. In particular, our results indicate a “lock-in” effect in the choice of materials (i.e., a positive reinforcement of the choices of DOM and PM in new experiments according to combinations already studied). Finally, our analysis allows researchers to identify the combinations of materials studied and their prevalence in the experiments and, thereby, provides a basis for the planning of future experimental efforts.

## DOM and PM Categorization

First, we constructed a database of DOM and PM that were studied in combination. The identities of the DOM and PM were obtained from experimental studies that investigate the effects of DOM on inorganic PM in aquatic media and were published between 1977 and 2015. Here, we consider an experiment as a pair of DOM and PM studied in combination. For example, a publication that separately studied the effect of river humic acid and alginate on  $\text{TiO}_2$  for different pH values and ionic composition contains two experiments: one for “river humic acid– $\text{TiO}_2$ ” and another for “alginate– $\text{TiO}_2$ .” In total, our database contains 951 experiments from 271 peer-reviewed publications that used 94 different PM types and 133 different DOM types.

Here, we differentiate between three groups of DOM in the database: (i) group 1 DOM consists of isolates from environmental samples such as humic substances and hydrophilic acids, which are mixtures of various substances and only characterized in terms of averaged properties (48 DOM types); (ii) group 2 DOM consists mostly of individual substances such as synthetic polymers, low-molecular-weight ligands, isolated proteins, and polysaccharides (68 DOM types); and (iii) group 3 DOM contains whole-water samples such as lake water, river water, and sewage treatment plant effluents (18 DOM types). Group 1 DOM, in particular humic substances, are frequently used in PM–DOM experiments (537 of the 951 experiments) and, at the same time, difficult to characterize. Therefore, we focus on this group in the following.

In the absence of a consensus regarding the extraction and characterization methods for group 1 DOM (29), and lack of consistency in the reported properties (*SI Appendix, Consistency of Reported DOM Parameters*), a DOM categorization fully based on chemical composition or chemical properties is not currently feasible. Therefore, in most cases, we characterize the group 1 DOM in our database as either unfractionated DOM, humic acid, or fulvic acid and in addition in terms of the environmental sources from which they were sampled (i.e., river, soil, peat, etc.).

In a next step, we add information on chemical composition, which is possible because humic substances from similar sources often tend to have similar chemical composition, carbon distributions, and molecular weight distributions (14, 30–36). Below we investigate this relationship between DOM source and DOM properties in more detail.

DOM properties that correlate with the type and extent of DOM–PM interactions and are, therefore, relevant here include molecular weight, aromaticity, aliphaticity, and carbonyl content (16, 37–39).

Here, we focus on aliphatic, aromatic, and carbonyl carbon content estimated from solid-state  $^{13}\text{C}$ -NMR spectra. This method is often used to characterize humic substances (35, 40). Solid-state  $^{13}\text{C}$ -NMR spectra are readily available for most humic substances provided by the International Humic Substances Society (IHSS) ([www.humicsubstances.org](http://www.humicsubstances.org)).

To investigate the relationship between the environmental sources of group 1 DOM and their carbon composition, we compiled a set of 80 different samples of group 1 DOM along with data on their content of aromatic carbon, aliphatic carbon, and carbonyl carbon. To increase the diversity of this set, we used the group 1 DOM from our database in combination with additional types of group 1 DOM (*Dataset S2*). We then performed a principal-component analysis (PCA) to map the 80 DOM types onto the 2D space spanned by the first two principal components (PCs). An earlier PCA study of the  $^{13}\text{C}$  NMR spectra of 8 DOM samples discussed the differentiation between humic and fulvic acids but not the relationship between carbon distribution and the environmental sources of the DOM (41).

The first two PCs explain 99.9% of the variability in the carbon composition of the DOM samples analyzed. The first PC

(denoted PC1) is a linear combination of the aromatic carbon and carbonyl carbon contents in relationship to the aliphatic carbon content (PC1 weights: 0.67,  $-0.57$ , and  $-0.47$  for aliphatic, aromatic, and carbonyl carbon, respectively). The second PC (denoted PC2) is the difference between carbonyl carbon content and aromatic carbon content (PC2 weights: 0,  $-0.59$ , and  $0.81$  for aliphatic, aromatic, and carbonyl carbon, respectively) (*SI Appendix, Annex A*).

The first two PCs, therefore, divide the space of group 1 DOM carbon distribution into four quadrants (quadrants **a–d** in Fig. 1). Fig. 1 shows for each quadrant, derived from the DOM that falls into the quadrant, the distributions of aliphatic (red), aromatic (green), and carbonyl (blue) carbon.

Based on the PC weights listed above, DOM samples with a high PC2 score have low aromatic carbon content (often less than 20% of total carbon), which is similar to their content of carbonyl carbon. These materials belong to quadrants **a** or **b**. DOM that falls in quadrants **c** or **d** has a low PC2 score and is characterized by high content of aromatic carbon (mostly above 20% and up to 60%). Finally, DOM samples that score high in PC1 have high aliphatic carbon content (more than 50%), and therefore fall in quadrants **b** or **c**.

The distribution of all 80 group 1 DOM samples in the 2D space spanned by the first two PCs is depicted in *SI Appendix, Fig. S6*. Most aquatic DOM has relatively low aromatic carbon content and either low or high aliphatic carbon content (quadrants **a** and **b**). More specifically, all river fulvic acids share similar fractions of aliphatic, aromatic and carbonyl carbon, which agrees with earlier findings (35) (*SI Appendix, Fig. S7*). Generally, freshwater DOM shows relatively little variability in its carbon distribution. DOM from soil, in contrast, shows the highest variability and is found in all four quadrants. Soil humic acids,

specifically, have a high aromatic carbon content and therefore show up in quadrants **c** or **d**. Some soil samples, primarily fulvic acids, have carbon distributions similar to that of aquatic DOM and are found in quadrants **a** and **b**. Extraction and purification processes of soil DOM are a potential cause to this large variability in its carbon composition (35, 42), but natural heterogeneity can also be a cause. Overall, humic substances show important differences in their carbon distribution. In experiments with PM, humic substances should be chosen in such a way that they actually reflect the type of environment that is in the focus of the experiment. For example, many experiments use Aldrich humic acid as a generic DOM. In the PC1–PC2 graph, Aldrich humic acid lies in the bottom part of quadrant **c**, typical of coal and soil humic acids (*SI Appendix, Fig. S6*), which makes this DOM type not representative of any aquatic environment. The map of humic substances in *SI Appendix, Fig. S7* can assist in the selection of humic substances for experiments with PM. For further details of the relationship between the composition and the sources of group 1 DOM, see *SI Appendix, Environmental Sources and Chemical Composition of Dissolved Organic Matter*.

In conclusion, our analysis allows us to differentiate between several group 1 DOM categories with substantially different fractions of aromatic, aliphatic and carbonyl carbon. We will use these categories (and DOM subgroups within the categories) to characterize the diversity of the group 1 DOM, such as humic substances, investigated in experiments with PM.

Regarding PM, the focus of the experiments is often material-specific [e.g., fate of  $\text{TiO}_2$  (43)]. Accordingly, the PM types in our database are defined by the chemical composition of both the core material and the coating (e.g., citrate-coated silver, silver, and titanium dioxide). Additionally, we also interpret the results of the analysis in terms of broader PM groups based on the core material (e.g., metal, metal oxide, metal sulfide, etc.). The complete list of DOM and PM used in the experiments and the respective references are given in *Dataset S1*.

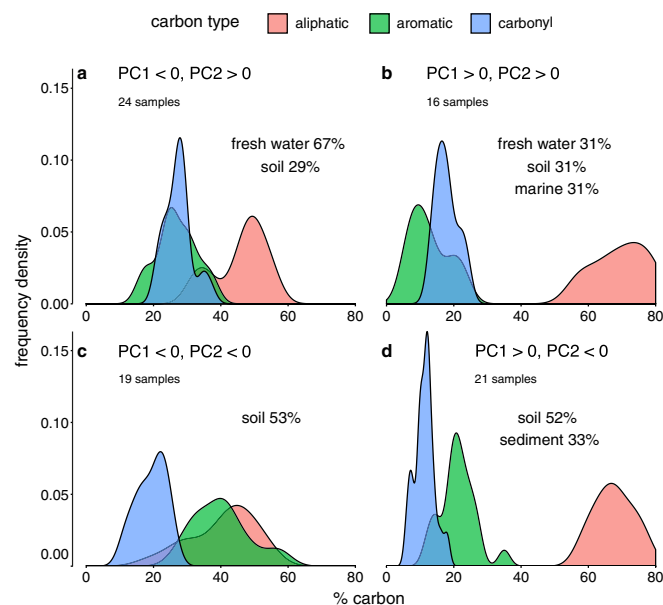
## Organizing the Experimental Data in a Network

In our analysis of the diversity of the PM–DOM combinations studied, the experimental field needs to be represented as a coherent entity. Such a representation can be attained by connecting the different types of DOM and PM that have been investigated in a network [i.e., a single structure that represents connections (links) between pairs of objects (nodes) (44)]. Here, we say a certain DOM and PM are connected by a link if they were studied together; the link's weight equals the number of experiments studying the connected DOM and PM (note that these experiments may still use different conditions of other parameters such as pH, ionic strength and ionic composition). The network is bipartite, which means that connections are allowed only between objects of different type (45), here DOM and PM.

## Network Topology

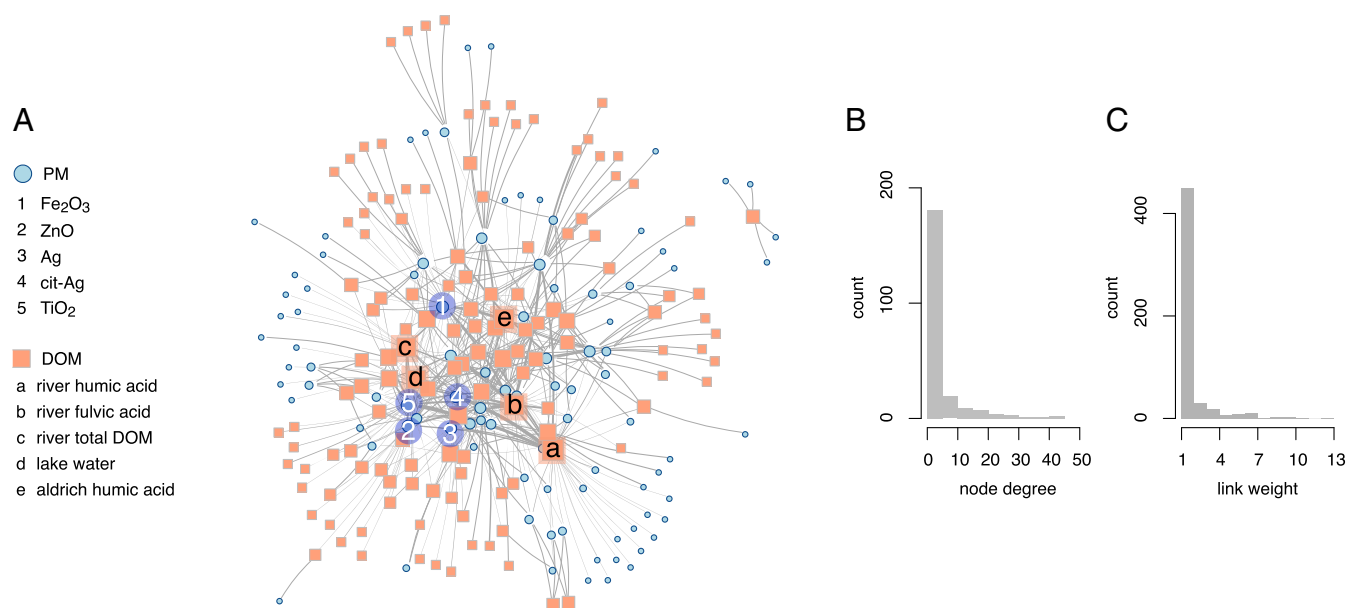
Fig. 24 shows the network derived from a set of 951 experiments with 94 PM nodes and 133 DOM nodes. In Fig. 24, the sizes of nodes is proportional to the number of counterparts a given material was studied with (e.g., the size of a PM node type is proportional to the number of DOM types the PM was studied with). The empirical network is called “complex” because it contains topological features absent from random networks of the same size (46) (*SI Appendix, Table S3* compares basic properties between the empirical and random networks, in which each DOM–PM combination is linked with a fixed probability).

Two important features of the empirical network are central nodes that have a large number of links (“hubs”) and relatively isolated, star-like shapes (“stars”) (the central and peripheral regions in Fig. 24, respectively). These features reflect different types of experimental studies. Hubs (central nodes in Fig. 24)



**Fig. 1.** Distributions of aliphatic, aromatic, and carbonyl carbon content in the four quadrants defined by the first two PCs, based on 80 group 1 DOM samples in total. “High” and “low” in the following are relative terms and specific to each type of carbon. Quadrant **a** is low content of both aliphatic and aromatic carbon and high content of carbonyl carbon. Quadrant **b** is low aromatic carbon and high carbonyl and aliphatic carbon content. Quadrant **c** is low content of carbonyl and aliphatic carbon and high content of aromatic carbon. Quadrant **d** is high content of aliphatic and aromatic carbon and low content of carbonyl carbon. In each quadrant, the percentages indicate how many DOM samples belong to the different types of environmental DOM sources.





**Fig. 2.** Important aspects of the empirical network. (A) The empirical network with nodes scaled according to their degree (i.e., number of links) and colored by material type as indicated in the legends. Five material types with high node degrees are indicated for each class, PM and DOM. Link widths are proportional to the number of experiments using the linked combination. (B) Distribution of the number of DOM or PM types studied with any given PM or DOM, respectively (node degrees). (C) Distribution of the number of experiments conducted per each unique combination of DOM and PM (link weights).

correspond to well-studied materials that are frequently investigated in combination with new materials (e.g., Fe<sub>2</sub>O<sub>3</sub>, TiO<sub>2</sub> and Sigma-Aldrich humic acid). The presence of star-like shapes is quantified by the network's negative degree assortativity, which means that the correlation between the number of links of neighboring nodes is negative (*SI Appendix, Table S3*). Materials forming the centers of stars were studied in combination with many others (edges of the stars), which in turn were studied mainly or exclusively with the central material. Stars indicate the introduction of new combinations of materials that were rarely studied in other experiments (e.g., lake humic acid, wetland total DOC, and CuS).

If materials are selected for new experiments that lead to hubs or peripheral stars, then the overall materials' diversity increases, because new links are added to the network. It remains unclear, however, whether the call for an increase in the materials' diversity intends to promote the systematic investigation of selected and well-studied materials with new counterparts (hubs), or whether the need is for general expansion of the network by using completely new materials (peripheral stars). The simultaneous presence of hubs and stars in the empirical network is reflected by the right-skewed degree distribution, which reflects the distribution of the number of material types with which each material was studied (see Fig. 2B).

The diversity of the experimental field increases with the number of DOM–PM combinations studied. In the empirical network, diversity is reflected by the density of the links (i.e., the fraction of existing links out of all possible links). The density of the empirical network is 4.3%, which means that 4.3% of the possible DOM–PM combinations (535 of 12,502 possible ones) have been studied in one or more experiments. However, some experiments study the same combinations of DOM and PM. Had each experiment studied a different DOM–PM combination (i.e., all links have a weight of 1), the overall number of links in the network would be the total number of experiments, which results in a density of 7.6% (951 DOM–PM combinations of 12,502 possible ones). The network's density is specific to the empirical network presented here, and addition of new materials (i.e., nodes) will result in a network with a different structure and density.

Because the number of experiments studying a given DOM–PM combination is the weight of the link connecting the two materials, high diversity of DOM–PM combinations studied will yield links with weights of 1 (under the assumption that the number of experiments is lower than the number of DOM–PM combinations). On the other hand, low diversity will result in links of higher weights. In our empirical network, most links have low weight, but there are a few links that have substantially higher weights (up to 13). The resulting right-skewed distribution of link weights (Fig. 2C) indicates that the topology of the empirical network is the result of two opposing trends in the choice of materials.

It is desirable that any given DOM–PM combination will be studied under different conditions by different researchers (i.e., the link connecting the given DOM and PM will have a weight >1). However, the fact that most links have a weight of 1 means that only a handful of DOM–PM combinations undergo such a more detailed investigation. For example, of the 951 experiments in our database, soil fulvic acid and Fe<sub>2</sub>O<sub>3</sub> were studied only once, whereas river humic acid in combination with citrate-coated silver was studied in 13 experiments.

Not all DOM types in the empirical network have the same environmental abundance (e.g., humic substances vs. proteins). To test the effect of different DOM types on the structure of the network, we built a “reduced” network only from experiments that use humic substances as DOM. Importantly, this reduced network has similar properties as the full empirical network. Particularly, both the full and the reduced networks are sparse (i.e., have low density), have a similar topology (i.e., core-periphery structure with hubs and star-like regions), and share many of the central nodes (*SI Appendix, Table S4 and Fig. S10*). Therefore, the reduced network reflects similar trends in the choices of materials as the full network, and the properties of the full network are not caused by mixing DOM types of different environmental abundance and relevance.

### Network Evolution from 1990 to 2015

Here, we investigate how temporal changes in the diversity of materials shape the global structure of the empirical network. To this end, we compare the empirical network to two simulated

networks that exhibit either high or low diversity, and track their structural changes over time.

The simulated networks were obtained by replacing in each publication in our database the DOM–PM combinations that were actually studied by DOM–PM combinations that were selected according to two rules: To obtain a network that exhibits low diversity, for each publication a DOM–PM combination was chosen that repeats one of those mentioned in the references cited in the publication. Correspondingly, choosing DOM–PM combinations different from the ones used in the cited references of each publication (by sampling uniformly at random from all combinations not studied in the cited references), resulted in a network with high diversity. We simulated 1,000 networks in each year for each network type (i.e., low- and high-diversity networks).

The simulated and empirical networks differ by the number of nodes and links as well as link weights. Because low diversity implies that over the years fewer DOM–PM combinations are studied, the corresponding network is smaller than the one exhibiting high diversity. In the high-diversity case, the network is constantly expanded by the introduction of new materials. The only quantity that is conserved across all networks is the total sum of link weights (i.e., the number of experiments performed in each year), which is used as a basis of comparison between the empirical and simulated networks: to compare the diversity of the networks, we measure how many DOM–PM combinations were studied compared with the number of experiments performed up to a given year. Accordingly, we define a diversity index,  $D_{\text{comb},i}$ , as:

$$D_{\text{comb},i} = \frac{n_{\text{comb},i}}{n_{\text{exp},i}}, \quad D_{\text{comb},i} \in \left[ \frac{1}{n_{\text{exp},i}}, 1 \right], \quad [1]$$

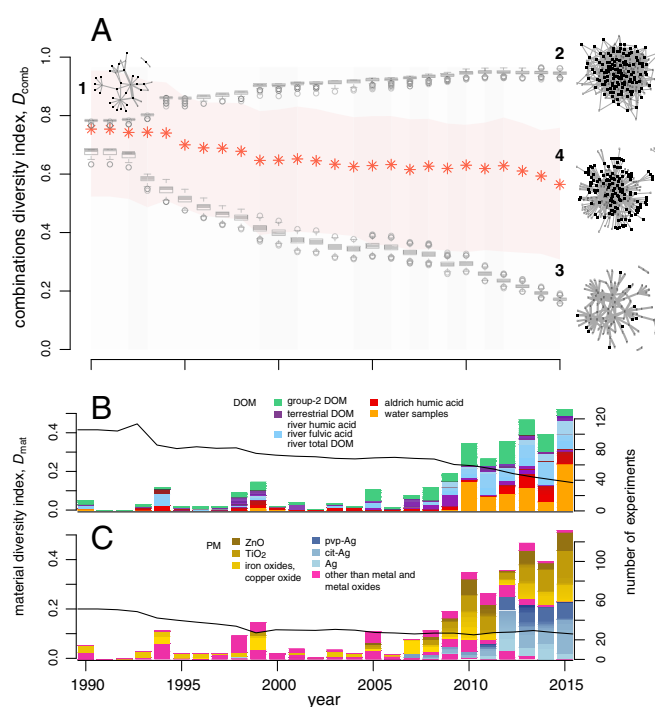
where  $i$  counts the years and  $n_{\text{comb},i}$  and  $n_{\text{exp},i}$  are the number of DOM–PM combinations studied and the number of experiments, respectively, in all years from 1990 to year  $i$ .

When the diversity is maximal, each experiment studies a different DOM–PM combination and  $D_{\text{comb},i}$  takes the value of 1. A decrease in  $D_{\text{comb},i}$  indicates that the number of combinations studied compared with the total number of experiments decreases. The minimum value for  $D_{\text{comb},i}$  is  $1/n_{\text{exp},i}$ , which corresponds to the case where all experiments up to year  $i$  studied the same combination of materials. Similarly, we define a diversity index for each class of material, DOM and PM:

$$D_{\text{mat},i} = \frac{n_{\text{mat},i}}{n_{\text{exp},i}}, \quad D_{\text{mat},i} \in \left[ \frac{1}{n_{\text{exp},i}}, 1 \right], \quad [2]$$

where  $n_{\text{mat},i}$  is the number of different types of DOM or PM studied in all years from 1990 to year  $i$ .

All simulations started from the empirical network in 1990 (Fig. 3A, network 1). As can be seen in Fig. 3A, in the networks with high diversity  $D_{\text{comb}}$  increases with time (network 2); when the diversity is low (network 3),  $D_{\text{comb}}$  primarily decreases with time. After 2007, there is a strong increase in the number of experiments per year (see bars in Fig. 3B and C), which correlates with a clear decrease in the  $D_{\text{comb}}$  of the low-diversity network (lower series of boxplots in Fig. 3A). In contrast, for the high-diversity network the number of experiments per year has almost no correlation with the  $D_{\text{comb}}$  values (upper series of boxplots in Fig. 3A). The reason for this finding is that the high-diversity network quickly approaches a high  $D_{\text{comb}}$  value, and new DOM–PM combinations increase both the numerator and the denominator in Eq. 1 and  $D_{\text{comb}}$  does not change much. The maximum value of 1 is not observed because at the starting point of the simulations (in 1990) several experiments already studied the same DOM–PM combinations. The boxplots represent the range of  $D_{\text{comb}}$  for each year, obtained from the 1,000 simulated high- and low-diversity networks (for a detailed discussion, see [SI Appendix](#)).



**Fig. 3.** Temporal changes in the diversity of the materials studied. (A) Diversity index of the DOM–PM combinations,  $D_{\text{comb}}$ , as a function of time. Network 1 represents the experimental field in 1990 and is the starting point for the simulated networks. Network 2 is the expected structure of the experimental field in 2015, when the diversity of the DOM and PM was simulated to be high during 1990–2015. Network 3 is the result of simulating low diversity during this period. Boxplots represent distributions of  $D_{\text{comb}}$  for 1,000 simulated networks in each year. Network 4 is the current empirical network; red asterisks represent the change in  $D_{\text{comb}}$  of the empirical network. The light red band shows the variation in  $D_{\text{comb}}$  as a result of either a finer (upper bound) or coarser (lower bound) categorization of DOM types. The left vertical axis in B shows the material diversity index,  $D_{\text{mat}}$ , of DOM as a function of time. Bars in B show the number of experiments in each year (right vertical axis); the color code indicates different DOM types, and the legend shows DOM types mentioned in the text. The left vertical axis in C shows the material diversity index,  $D_{\text{mat}}$ , of PM as a function of time. Bars in C show the number of experiments in each year (right vertical axis); the color code indicates different PM types, and the legend lists some of the most-used PM types between 2012 and 2015.

The  $D_{\text{comb}}$  of the empirical network (network 4 in Fig. 3A) decreases over time, however, not monotonously. Numerous factors influence the choice of materials (e.g., scientific interest and preferences of individual scientists, established scientific collaborations, trends in the field, costs and availability of materials, etc.), which results in the study of both new combinations and combinations studied before. Accordingly, the  $D_{\text{comb}}$  of the empirical network shows alternating periods of increase and decrease. Still, the significantly negative slope of a linear regression line fitted to the  $D_{\text{comb}}$  values (see Methods section) indicates that it is mainly a trend toward lower diversity of the materials studied that has shaped the structure of the experimental field over the last 25 y.

We performed several tests to check how stable the decreasing trend in  $D_{\text{comb}}$  is. Specifically, we analyzed how  $D_{\text{comb}}$  is influenced by potential missing publications in our database (e.g., not retrieved in the publication search), and by finer or coarser definitions of the DOM types. The decreasing trend in  $D_{\text{comb}}$  was stable in both tests: the decreasing diversity was also observed under random perturbations of the publications in our database, which we used to simulate the effect of missing publications. Decreasing diversity is also shown by the red band in Fig. 3A,

which shows the influence of finer and coarser definitions of the DOM types on  $D_{\text{comb}}$ . For a detailed discussion, see [SI Appendix, Resilience of the Experimental Network](#).

In conclusion, the observation of an overall decrease in the  $D_{\text{comb}}$  holds for a wide range of possible definitions of DOM types. Therefore, the decrease in diversity is unlikely to be an artifact of the DOM categorization scheme used.

In Fig. 3B and C, temporal changes in the material diversity index ( $D_{\text{mat}}$ ) are depicted separately for DOM and PM, respectively ( $y$  axes on the left). For DOM,  $D_{\text{mat}}$  decreases over the entire period. In contrast, the decrease in PM diversity slows down after the year 2000. In the early 2000s, calls were made to assess the potential risks of nanotechnology (3, 47), which were followed by extensive research aimed at evaluating the fate of ENPs in the environment (48). Because these ENPs are more diverse (e.g., in terms of their surface properties and morphologies) than natural PM, ENP research introduced some additional diversity into the experiments. Thus, the overall decrease in the diversity of the experimental field is likely due to the constant decrease in DOM diversity.

There are several plausible explanations for the observed decrease in  $D_{\text{comb}}$ ; from the perspective of an individual scientist, low diversity between the experiments might be of a minor concern or is even desirable (e.g., for the purpose of reproducibility of experiments). The overall diversity and its importance is only evident when the entire experimental field is assessed as a whole. Therefore, the decrease in diversity may arise from the discrepancy between the individual and the collective perspectives. Additionally, even when individual researchers aim at studying new DOM–PM combinations to increase the diversity, the large number of possible DOM–PM combinations makes it difficult to assess (based on published experiments) which additional DOM–PM combinations should be investigated. Another possibility is that the increasing occurrence of certain materials used in the experiments may result in the perception of these materials as standards, creating drift toward the use of these materials in a positive feedback process.

The last period (2012–2015), which displays a decrease in  $D_{\text{comb}}$  for the empirical network (Fig. 3), is of special interest because the number of experiments carried out in this period comprises about 46% of the total number of experiments analyzed here (442 of 951). We will therefore analyze the origin of the decreasing trend, in particular in the last years, in more detail in the following section.

### Systematic Trends in Experimental Designs Explain the Reduction in the Diversity of Materials

Here, we investigate the temporal trends and frequency of use of specific materials, first for DOM and then for PM.

Fig. 3B shows that the various DOM types used in the experiments in recent years originate mainly from whole-water samples (e.g., lake water, river water, sea water, etc.; group 3 DOM) (orange parts of the bars in Fig. 3B) and from group 1 DOM sampled from river water (blue parts of the bars in Fig. 3B). The focus on group 1 humic substances from river water has become more pronounced in 2010–2015 and the fraction of experiments with group 1 DOM from other environment types such as peat, soil, seawater, and coal, has correspondingly decreased (purple part of bars in Fig. 3B; also see [SI Appendix, Fig. S3 B and C](#), where the groups of DOM are shown in more detail than in Fig. 3B). Other relevant parts of DOM visible in Fig. 3B are Aldrich humic acid (red parts of the bars) and group 2 DOM such as alginate (green).

The river DOM shown in blue in Fig. 3B mostly belongs to quadrant **a** on the map spanned by the first two PCs of the PCA (Fig. 1 and [SI Appendix, Fig. S6](#)). These are substances that have relatively low aliphatic and aromatic carbon content. On the

other hand, use of DOM that belongs to quadrant **c** constantly decreases. This DOM has low aliphatic and high aromatic carbon content and includes humic acids from peat, soil, and coal ([SI Appendix, Fig. S6](#)). Finally, the experiments in our database only rarely use DOM that contains high aliphatic and high aromatic carbon content, such as the ones that belong to quadrant **d** (e.g., sediment humic substances), and high aliphatic and low aromatic carbon content (i.e., marine humic substances in quadrant **b**).

The increasing focus on group 1 DOM types with similar carbon distribution (predominantly from quadrant **a**) parallels the decreasing diversity index of both DOM types used and DOM–PM combinations studied, primarily between 2012–2015 (Fig. 3A and B).

Soil DOM is highly heterogeneous and, at the same time, has low prevalence in the experiments (Fig. 3B). Accordingly, the information regarding the interaction of PM with soil DOM is still scarce. Because there is no “representative” soil DOM (see results of the PCA above), studies on the interaction between soil DOM and PM, for example in pore water, need to account for the heterogeneity of soil DOM, and use DOM extracted from different specific soil types.

Group 2 DOM, which primarily includes isolated and well-defined substances, was used in only 22% of all experiments (211 out of 951 experiments). Group 1 DOM is the major focus of the experimental effort, which confirms previous observations of high prevalence of humic substances as the DOM component in the DOM–PM experiments (26).

The DOM isolated from environmental media (blue and pink in Fig. 3B), on the one hand, and DOM in water samples without further processing (yellow in Fig. 3B), on the other hand, provide two complementary options for studying DOM–PM interactions. With isolated DOM, a deeper mechanistic understanding of the interaction between DOM and PM can be obtained. However, it is not always the case that the synthetic conditions created are environmentally relevant. Water samples, on the other hand, better represent environmental conditions but lack of control of the concentration and properties of the DOM present in a water sample limits the mechanistic understanding of the experimental output (25). Because these two approaches are complementary (49), it follows that, when researchers have obtained a mechanistic understanding, their insights could be expanded and corroborated in follow-up experiments with natural-water samples, for example see refs. 50 and 51.

Water-sample DOM types comprise a large number of the overall experiments (Fig. 3B). In the empirical network, however, water-sample DOM types occupy only a small region in the network (light blue squares in the network shown in [SI Appendix, Fig. S5A](#)). This finding indicates that the large number of experiments that use water samples are investigated in combination with only a small fraction of all PM types present in our database.

Therefore, there is a potential for more diversity through investigation of water-sample DOM with more PM types. However, in these experiments the DOM and also the experimental conditions in general would have to be characterized in a more detailed, more comprehensive, and more systematic way than what is found in many existing studies. Without that, the added value of more experiments with water samples will remain unclear.

For the PM in our database, it is the combinations of coating and core material that need to be investigated to understand the trend in diversity. The initial coating of ENPs strongly influences the interaction of the particles with DOM, which in turn can affect the stability of the ENPs (52–54). For this reason, every PM that was reported to have a distinct initial coating is a separate entry in our database. On the other hand, if the authors of an experimental study did not mention the exact PM coating (and if this information could not be found in the material description



of the supplier), or the authors explicitly referred to the PM as “bare,” we consider this PM as uncoated. Therefore, “uncoated” PM in the empirical network means that a possible initial coating of the PM was not considered as a parameter in the experiments, and therefore the experimental results do not account for the coating’s interaction with DOM.

PM types with novel coatings are continuously introduced into the network, which increases the diversity of the PM in the experimental network (because the PM types in our database are distinguished by both their core material and their coating (Fig. 3C, years 2011–2013) or keeps the decrease shallow (years 2014–2015). However, the diversity of the PM–DOM combinations further decreases (Fig. 3A, years 2012–2015) because the coated PM types are investigated only with a (very) limited number of DOM, as is demonstrated by the network shown in Fig. 4A.

In Fig. 4A, the 45 nodes that correspond to coated PM (orange dots) mainly occupy the periphery of the network, whereas those that correspond to uncoated PM (49 nodes) are mostly central (light blue circle in Fig. 4A). Particularly, coated PM is studied with significantly fewer DOM counterparts than uncoated PM (Fig. 4B). When the well studied citrate- and poly(vinylpyrrolidone) (PVP)-coated nanosilver as well as citrate-coated nanogold are not considered, the number of times a given coated PM was studied (i.e., number of experiments) is considerably smaller than a given uncoated PM (Fig. 4C).

We conclude that uncoated materials are often re-studied in combination with different DOM types and possibly varying other medium conditions. However, for coated PM, follow-up

experiments with different DOM types are performed only sporadically. As explained earlier, uncoated PM refers to any PM not explicitly described as being coated. The imbalance between the study of coated vs. uncoated PM implies that the role of the initial coating in determining the interaction of a PM with DOM is often not considered. However, ENPs are frequently coated, and, accordingly, coated ENPs are also likely to be released to the environment (52). Therefore, the underrepresentation of experiments with coated PM in the network indicates that more experiments that explicitly consider the coating of ENPs are needed to sufficiently characterize the interaction of ENPs with DOM.

Finally, in *SI Appendix*, Fig. S11, we observe that the network’s core, which contains the majority of the uncoated PM, is comprised mostly of recent experiments. Therefore, the low consideration of the effect of initial PM coating on DOM–PM interactions is a current phenomenon.

## Conclusions and Recommendations

Over the last 25 y, numerous experiments have been performed that aim to elucidate the effect of DOM on ENPs specifically and PM in general. Still, recent reviews call for additional experiments (25, 26), and it is important to consider carefully which DOM–PM combinations should be studied in future experiments. Here, the overall objective of the experiments plays an important role. One objective is to investigate the fate of PM in specific environments, such as ENPs in surface water. For experiments with this objective, DOM representative of the environment considered will be used (e.g., river and lake humic acids and fulvic acids for surface water but not Aldrich humic acid, which reflects soil DOM properties; *SI Appendix*, Fig. S7).

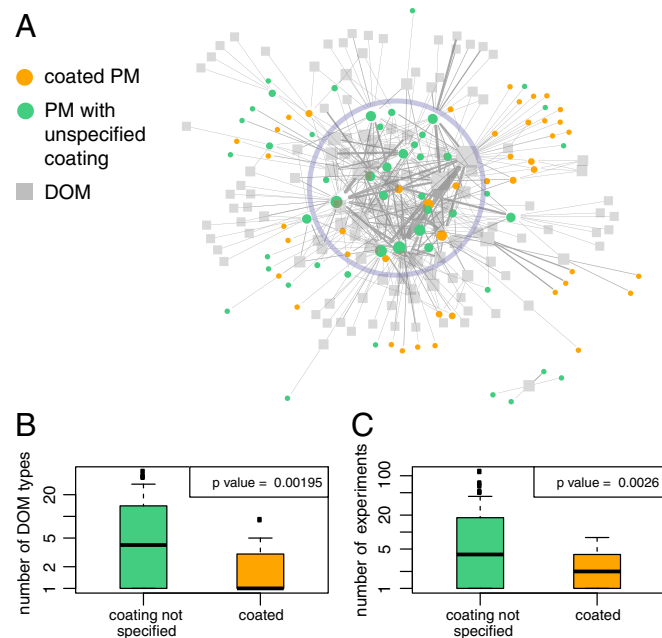
Another objective is to mechanistically study the PM–DOM interactions without a focus on a specific environment, but with the aim to improve the general mechanistic understanding of these interactions. For this kind of experiment, a broad range of material properties is needed.

To increase the diversity of materials, researchers can study entirely new materials (i.e., add new nodes to the empirical network) and/or study DOM–PM combinations that were not studied together before (i.e., add new links between existing nodes in the empirical network). For the selection of DOM for new experiments, the PC1–PC2 map in *SI Appendix*, Figs. S6 and S7 can guide the identification of suitable DOM. In any case, it is imperative that the experiments will be comparable in terms of the materials’ characterization to foster the development of both empirical models and a better mechanistic understanding.

The development of a DOM classification based on physicochemical properties and chemical functionality requires a consistent reporting of relevant DOM characteristics (chemical composition, carbon distribution, molecular weight, etc.). In 40% of the publications in our database that use one or more group 1 DOM, there is no information regarding the molecular weight nor the chemical composition nor the carbon distribution for at least one of the group 1 DOM studied. Similarly, less than half of the publications in our database that use DOM from the IHSS report the sample’s identifier (*SI Appendix*, Fig. S1). It is essential that the basic characteristics of the DOM used in experiments with PM are reported much more systematically and comprehensively.

What are possible reasons for the observed decrease in diversity? To some extent, the decrease has been caused by the many studies focusing on ENPs in aquatic systems in the last years (which is an intentional focus on an environmentally relevant problem).

In addition, because the experimental field is large and disparate, it is difficult to obtain a comprehensive overview of past experimental achievements. This situation may lead to a lock-in dynamics that reinforces the use of certain DOM–PM



**Fig. 4.** Differences between coated PM and PM where coating was not considered (uncoated). (A) Uncoated PM (green) are located mainly at the center of the empirical network (light blue circle), whereas the coated PM (orange) are located mainly at the periphery of the network. Size of nodes is proportional to their degree. (B) Boxplots showing the distributions of the number of DOM types studied with uncoated and coated PM (i.e., the degrees of the PM nodes). (C) Boxplots showing the number of experiments performed (i.e., sum of link weights of the PM nodes) with uncoated and coated PM (with citrate- and PVP-coated nanosilver and citrate-coated nanogold omitted). The borders of the boxplots are the 25th and the 75th percentiles. *P* values are for the two-sample, nonparametric Mann–Whitney *U* test (alternative hypotheses: “uncoated PMs are studied with more DOM types than coated PMs” and “uncoated PMs are studied in more experiments than coated PMs”).



combinations and makes other combinations less visible and likely to be studied.

Such a lock-in effect in favor of certain experimental settings or materials is not limited to the field of environmental fate of ENPs. It is known that in science in general an initial focus on certain experimental settings reinforces the use of these settings in the future. For example, in nanomaterial ecotoxicology it has been pointed out that the majority of the ecotoxicological data are confined to, among others, fresh water organisms, although more studies on terrestrial and marine organisms are needed (55). Another example is the various databases that summarize key properties of ENPs to facilitate ENP risk assessment. These databases often offer information on selected ENP types (e.g.,  $\text{TiO}_2$  and  $\text{ZnO}$ ) but lack the required information on others (e.g., Ag and  $\text{Fe}_2\text{O}_3$ ) (56). In both examples, the materials and/or experimental settings are criticized as being unrepresentative or too narrow (55, 56).

This problem is also discussed in a more general way in philosophy of science. Weisberg and Muldoon (57) point out that the ability of a scientific community to discover new knowledge depends on the balance between scientists who use established settings and scientists who apply new approaches to new settings (57). Our analysis shows how an imbalance between these two approaches can emerge over time.

A network perspective as shown here is useful because in a network the entire experimental field, consisting of diverse experimental settings, can be visualized as a single entity. This visualization provides a graphical summary of the current state of the science, but also a tool for modeling and understanding how the structure of the research field has developed; finally, it helps identify gaps in the body of experiments performed and future research needs. In the context of DOM and PM research, further research should (i) place more emphasis on the initial coating of the PM and its interaction with DOM; (ii) confront results from experiments with isolated DOM against studies with natural water samples; and (iii) strive to increase the diversity of the types of DOM and PM investigated.

## Materials and Methods

**Publication Search.** The empirical network is based on experiments that study the effects and/or interactions of DOM with PM in aquatic media. Experiments that report transport of PM in porous media were included if in addition they studied the DOM effects on the PM in aqueous media. Experiments where the PM is a naturally occurring colloid (e.g., suspended sediment) were not included. Relevant studies published until 2013 were taken from the review by Philippe and Schaumann (26). Studies between 2013–2015 were obtained from two search queries in SciFinder (58) performed on July 2, 2015 using the keywords (i) colloids and organic matter, yielding 352 results; and (ii) nanoparticles and organic matter, yielding 493 results. The search results were automatically filtered, using a designated R script, and the remaining publications contained in their abstract either one of the (stemmed) keywords: “stability,” “aggregation,” “adsorption,” “sorption,” “sedimentation,” “coating,” “deposition,” “mobility,” “surface,” “disaggregation,” and “precipitation.”

In a second search (July 12, 2016) additional publications were obtained from two more queries in SciFinder (58): “dissolved organic matter and nanoparticles” and “dissolved organic matter and inorganic colloids,” for experimental papers published between 1977–2015. This search resulted in 173 and 58 papers for the first and second query, respectively. Of those, 26 were relevant based on their title and abstract and were not already in our database. A third SciFinder search (December 12, 2016) was carried out for experimental papers published between 1977–2015 using four additional queries: “dissolved organic matter and nanoparticles dissolution,” “dissolved organic matter and nanoparticles precipitation,” “dissolved organic matter and colloids precipitation,” and “dissolved organic matter and colloids dissolution.” This search gave 53, 7, 28, and 95 hits for the first to fourth query, respectively. Of those, six were pertinent to our analysis and were not yet in the database. The remaining papers were studied in detail and kept if they report the effect of DOM on PM aggregation/dissolution/precipitation or if they study the coating of PM by DOM in aquatic media.

**Types of Studies.** Of the 271 publications, 172 investigate, among others, PM aggregation; 45 study, among others, PM dissolution and precipitation; and 58 study only DOM adsorption onto PM surfaces. A full list of publications and the used materials are given in [Dataset S1](#).

**PCA.** We analyzed the carbon distribution of different DOM types reported in the literature and on the IHSS website ([www.humicsubstances.org](http://www.humicsubstances.org); accessed November 8, 2016). The full list of materials and their properties are given in [Dataset S2](#). Because of the many missing values for different parameters (above), we could not use all DOM types listed in the references.

The publications analyzed here report the carbon distribution as obtained from  $^{13}\text{C}$ -NMR spectra. Different authors sometimes use different regions of the chemical shift to quantify different carbon types such as aliphatic, heteroaliphatic, acetal, aromatic, carboxyl and carbonyl carbon. To smooth out these differences, we combined smaller integration intervals to larger ranges that account for three basic carbon types: aliphatic carbon (i.e., heteroaliphatic and acetal carbon), aromatic carbon, and carbonyl (also accounts for carboxyl) (International Humic Substances Society and refs. 59 and 60): (i) aliphatic carbon (%), integrated area between 0 and 110 ppm; (ii) aromatic carbon (%), integrated area between 110 and 165 ppm; and (iii) carbonyl carbon, integrated area between 165 and 220 ppm.

Next, we performed a PCA (on the scaled database). In short, a PCA finds the set of orthogonal vectors (also called the principal components or, in short, PCs) that define the linear combinations of the parameters that capture the highest variability in the database. Each data point is projected into the respective PC dimension by calculating the linear combination of the parameters, weighted by the respective PC coefficients. The analysis was done using the function *princomp* in R, and its output is presented in [SI Appendix, Annex A](#).

**Empirical Network Construction.** From the publications collected a list of all studied DOM–PM combination was extracted and converted to a bipartite network (45) using the *igraph* package 61, available in R, version 3.3.0 (62). This network has two types of nodes, PM or DOM, and links are only possible between nodes of different types. The network,  $G$ , is defined by

$$G = (U, V, E), \quad E \subseteq U \times V, \quad [3]$$

where  $U$  and  $V$  are the sets of all DOM and PM, respectively, and  $U \times V$  is the Cartesian product, i.e., the set of all possible links. The set  $E$  denotes all DOM–PM combinations studied. The weight of a given link is the number of experiments studying the DOM and PM connected by this link. Links are undirected because they imply the presence of both materials in the corresponding experiments.

Coated PM was regarded as a new material (i.e., citrate-coated silver NPs were regarded different from polyvinylpyrrolidone-coated silver NPs) because the coating/surface functionalization has been shown to alter the interaction of the particles with DOM. The broad PM groups are based on chemical composition of the core material. The different types of DOM were categorized according to source (e.g., humic acid from a river or a lake) and physicochemical properties (e.g., anionic and nonionic surfactant). Exceptions included commercial humic acids, often from an unspecified source (63), which were categorized according to the manufacturer's name (e.g., Aldrich humic acid). For experiments reporting the effects of the total DOM from a given source as well as the different DOM fractions, the entry in the database is for the unfractionated DOM (e.g., different molecular fractions of Suwannee River Natural Organic Matter (64) were given a single node in the network).

**Linear Regression on the Diversity Values Between 1990 and 2015.** A linear regression line was fitted to the  $D_{\text{comb}}$  values presented in the section Network evolution between 1990–2015. However, inference of regression parameters by means of ordinary least squares (OLS) requires independence of the error terms (i.e., residuals), which is not the case for the OLS residuals of the  $D_{\text{comb}}$  values. Rather, they exhibit a serial correlation that can be modeled by an autoregressive (AR) process of order 1. Therefore, the linear regression was carried out using generalized least squares with residuals correlation structure of the respective AR process, using the function *glsl()* from the package *nlme* (65) in R (62). The resulting regression line is given by

$$D_{\text{comb}} = 0.807 - 0.00893 \cdot t \quad (t: \text{years from 1 to 25}). \quad [4]$$

The 95% confidence interval for the slope is  $[-0.0133 - 0.00455]$ .

**Simulation of Networks with High and Low Diversity of DOM–PM Combinations.** For each case (i.e., high and low diversity of the studied DOM–PM combinations), 1,000 networks were simulated for each year from 1990–2015. Networks of high diversity were simulated by replacing the DOM–PM combinations used in a given publication by combinations sampled uniformly at random from all possible combinations that were not studied in the cited publications of this given publication. Networks of low diversity were simulated by replacing the studied DOM–PM combination of a given publication by the most frequent combinations used in the cited references of this publication; if all experimental combinations cited in a publication are unique, DOM–PM combinations were sampled uniformly at random from the ones studied in the cited references of the publication.

**Supporting Information.** *SI Appendix* contains results and figures from additional analyses. The analysis of the robustness of the diversity trend using bootstrap simulation was performed with the package *boot* (66) in R (62).

Some figures in the *SI Appendix* were created by the *ggplot* (67) and *ggrepel* (68) packages in R (62). *Dataset S1* contains the experiments (with references) used to create the empirical network. *Dataset S2* lists the carbon distribution of the DOM types used for the PCA. The analysis code and its description are publicly available at [https://github.com/nicolesanikast/DOM-PM\\_network](https://github.com/nicolesanikast/DOM-PM_network).

**ACKNOWLEDGMENTS.** We thank Eli Goldberg for helpful discussions and the Seminar for Statistics (Eidgenössische Technische Hochschule Zurich), in particular Claude Renaux, for the valuable statistic consulting, support, and guidance. Funding from the Swiss Federal Office for the Environment is gratefully acknowledged. This work is part of the NANOETER project under the European Area Network (ERA-NET) on Safe Implementation of Innovative Nanoscience and Nanotechnology (SIINN). M.S. acknowledges financial support from the Czech Ministry of Education, Youth, and Sports (Grant LM2015051) and Masaryk University (CETOCOEN PLUS Project Grant CZ.02.1.01/0.0/0.0/15.003/0000469).

- Karn B, Kuiken T, Otto M (2009) Nanotechnology and in situ remediation: A review of the benefits and potential risks. *Environ Health Perspect* 117(12):1823–1831.
- Gottschalk F, Nowack B (2011) The release of engineered nanomaterials to the environment. *J Environ Monit* 13(5):1145–1155.
- Colvin VL (2003) The potential environmental impact of engineered nanomaterials. *Nat Biotechnol* 21(10):1166–1170.
- Boxall ABA, Tiede K, Chaudhry Q (2007) Engineered nanomaterials in soils and water: How do they behave and could they pose a risk to human health? *Nanomedicine* 2(6):919–927.
- Hannah W, Thompson PB (2008) Nanotechnology, risk and the environment: A review. *J Environ Monit* 10(3):291–300.
- Batley GE, Kirby JK, McLaughlin MJ (2013) Fate and risks of nanomaterials in aquatic and terrestrial environments. *Acc Chem Res* 46(3):854–862.
- Beckett R, Le NP (1990) The role of organic matter and ionic composition in determining the surface charge of suspended particles in natural waters. *Colloids Surf* 44:35–49.
- Liu J, Legros S, von der Kammer F, Hofmann T (2013) Natural organic matter concentration and hydrochemistry influence aggregation kinetics of functionalized engineered nanoparticles. *Environ Sci Technol* 47(9):4113–4120.
- Baalousha M (2009) Aggregation and disaggregation of iron oxide nanoparticles: Influence of particle concentration, pH and natural organic matter. *Sci Total Environ* 407(6):2093–2101.
- Aiken GR, Hsu-Kim H, Ryan JN (2011) Influence of dissolved organic matter on the environmental fate of metals, nanoparticles, and colloids. *Environ Sci Technol* 45(8):3196–3201.
- Majedi SM, Kelly BC, Lee HK (2014) Role of combinatorial environmental factors in the behavior and fate of ZnO nanoparticles in aqueous systems: A multiparametric analysis. *J Hazard Mater* 264:370–379.
- Gallego-Urrea JA, Perez Holmberg J, Hassellöv M (2014) Influence of different types of natural organic matter on titania nanoparticle stability: Effects of counter ion concentration and pH. *Environ Sci Nano* 1(2):181–189.
- Abe T, Kobayashi S, Kobayashi M (2011) Aggregation of colloidal silica particles in the presence of fulvic acid, humic acid, or alginate: Effects of ionic composition. *Colloids Surf A Physicochem Eng Asp* 379(1–3):21–26.
- Erhayem M, Sohn M (2014) Effect of humic acid source on humic acid adsorption onto titanium dioxide nanoparticles. *Sci Total Environ* 470:471–482.
- Erhayem M, Sohn M (2014) Stability studies for titanium dioxide nanoparticles upon adsorption of Suwannee River humic and fulvic acids and natural organic matter. *Sci Total Environ* 468:249–257.
- Deonarine A, Lau BLT, Aiken GR, Ryan JN, Hsu-Kim H (2011) Effects of humic substances on precipitation and aggregation of zinc sulfide nanoparticles. *Environ Sci Technol* 45(8):3217–3223.
- Pan B, Xing B (2008) Adsorption mechanisms of organic chemicals on carbon nanotubes. *Environ Sci Technol* 42(24):9005–9013.
- Hesselink FT, Vrij a, Overbeek JTG (1971) Theory of the stabilization of dispersions by adsorbed macromolecules. II. Interaction between two flat particles. *J Phys Chem* 75(5):2094–2103.
- Napper D (1977) Steric stabilization. *J Colloid Interface Sci* 58(2):390–407.
- Chen KL, Mylon SE, Elimelech M (2007) Enhanced aggregation of alginate-coated iron oxide (hematite) nanoparticles in the presence of calcium, strontium, and barium cations. *Langmuir* 23(11):5920–5928.
- Fleer GJ, Lyklema J (1974) Polymer adsorption and its effect on the stability of hydrophobic colloids. II. The flocculation process as studied with the silver iodide-polyvinyl alcohol system. *J Colloid Interface Sci* 46(1):1–12.
- Pelssers E, Stuart M, Fleer G (1989) Kinetic aspects of polymer bridging: Equilibrium flocculation and nonequilibrium flocculation. *Colloids Surf* 38(1):15–25.
- John G (1988) Polymer adsorption and flocculation in sheared suspensions. *Colloids Surf* 31:231–253.
- Labille J, Thomas F, Milas M, Vanhaverbeke C (2005) Flocculation of colloidal clay by bacterial polysaccharides: Effect of macromolecule charge and structure. *J Colloid Interface Sci* 284(1):149–156.
- Louie SM, Tilton RD, Lowry GV (2016) Critical review: Impacts of macromolecular coatings on critical physicochemical processes controlling environmental fate of nanomaterials. *Environ Sci Nano* 3:283–310.
- Philippe A, Schaumann G (2014) Interactions of dissolved organic matter with natural and engineered inorganic colloids: A review. *Environ Sci Technol* 48(16):8946–8962.
- Grillo R, Rosa AH, Fraceto LF (2015) Engineered nanoparticles and organic matter: A review of the state-of-the-art. *Chemosphere* 119:608–619.
- Goldberg E, Scheringer M, Bucheli T, Hungerbuehler K (2015) Prediction of nanoparticle transport behavior from physicochemical properties: Machine learning provides insights to guide the next generation of transport. *Environ Sci Nano* 2:352–360.
- Filella M (2008) Freshwaters: Which NOM matters? *Environ Chem Lett* 7(1):21–35.
- Beckett R, Jue Z, Giddings JC (1987) Determination of molecular weight distributions of fulvic and humic acids using flow field-flow fractionation. *Environ Sci Technol* 21(3):289–295.
- Chin YP, Aiken G, O'Loughlin E (1994) Molecular weight, polydispersity, and spectroscopic properties of aquatic humic substances. *Environ Sci Technol* 28(11):1853–1858.
- Esteves VI, Otero M, Duarte AC (2009) Comparative characterization of humic substances from the open ocean, estuarine water and fresh water. *Org Geochem* 40(9):942–950.
- Perminova IV, et al. (2003) Molecular weight characteristics of humic substances from different environments as determined by size exclusion chromatography and their statistical evaluation. *Environ Sci Technol* 37(11):2477–2485.
- Malcolm RL (1990) The uniqueness of humic substances in each of soil, stream and marine environments. *Anal Chim Acta* 232:19–30.
- Malcolm RL (1990) Variations between humic substances isolated from soils, stream waters, and groundwaters as revealed by C-NMR spectroscopy. Humic Substances in Soil and Crop Sciences: Selected Readings, eds McCarthy P, Clapp EC, Malcolm RL, Bloom PR (Soil Science Society of America, Madison, WI), pp 13–35.
- Slomberg DL, et al. (2016) Insights into natural organic matter and pesticide characterisation and distribution in the Rhone River. *Environ Chem* 14:64–73.
- Chen W, Qian C, Liu XY, Yu HQ (2014) Two-dimensional correlation spectroscopic analysis on the interaction between humic acids and TiO<sub>2</sub> nanoparticles. *Environ Sci Technol* 48(19):11119–11126.
- Ghosh S, Wang ZY, Kang S, Bhowmik PC, Xing BS (2009) Sorption and fractionation of a peat derived humic acid by kaolinite, montmorillonite, and goethite. *Pedosphere* 19(1):21–30.
- Louie SM, Spielman-Sun ER, Small MJ, Tilton RD, Lowry GV (2015) Correlation of the physicochemical properties of natural organic matter samples from different sources to their effects on gold nanoparticle aggregation in monovalent electrolyte. *Environ Sci Technol* 49:2188–2198.
- Simpson AJ, Simpson MJ, Soong R (2012) Nuclear magnetic resonance spectroscopy and its key role in environmental research. *Environ Sci Technol* 46(21):11488–11496.
- Thomsen M, et al. (2002) Characterisation of humic materials of different origin: A multivariate approach for quantifying the latent properties of dissolved organic matter. *Chemosphere* 49(10):1327–1337.
- Worobey BL, Webster GRB (1981) Indigenous <sup>13</sup>C-NMR structural features of soil humic substances. *Nature* 292(5823):526–529.
- Stone V, et al. (2010) Nanomaterials for environmental studies: Classification, reference material issues, and strategies for physico-chemical characterisation. *Sci Total Environ* 408(7):1745–1754.
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev Soc Ind Appl Math* 45(2):167–256.
- Asratian AS, Denley TMJ, Hagkvist R (1998) Introduction to bipartite graphs. Bipartite Graphs their Applications (Cambridge Univ Press, Cambridge, UK), pp 7–22.
- Costa LDF, Rodrigues FA, Travieso G, Villas Boas PR (2007) Characterization of complex networks: A survey of measurements. *Adv Phys* 56(1):167–242.
- Williams R, Alivisatos P, eds (2000) Nanotechnology Research Directions: IWGN Workshop Report Vision for Nanotechnology in the Next Decade (Springer, The Netherlands).
- Lecoanet HF, Bottero JY, Wiesner MR (2004) Laboratory assessment of the mobility of nanomaterials in porous media. *Environ Sci Technol* 38(19):5164–5169.
- Wang Z, Zhang L, Zhao J, Xing B (2016) Environmental processes and toxicity of metallic nanoparticles in aquatic systems as affected by natural organic matter. *Environ Sci Nano* 3(2):240–255.
- Ottoufelling S, Von der Kammer F, Hofmann T (2011) Commercial titanium dioxide nanoparticles in both natural and synthetic water: Comprehensive multidimensional

- testing and prediction of aggregation behavior. *Environ Sci Technol* 45(23):10045–10052.
51. Navarro DAG, Watson DF, Aga DS, Banerjee S (2009) Natural organic matter-mediated phase transfer of quantum dots in the aquatic environment. *Environ Sci Technol* 43(3):677–682.
  52. Labille J, et al. (2010) Aging of TiO<sub>2</sub> nanocomposites used in sunscreen. Dispersion and fate of the degradation products in aqueous environment. *Environ Pollut* 158(12):3482–3489.
  53. Lau BL, Hockaday WC, Ikuma K, Furman O, Decho AW (2013) A preliminary assessment of the interactions between the capping agents of silver nanoparticles and environmental organics. *Colloids Surf A Physicochem Eng Asp* 435:22–27.
  54. Yin Y, et al. (2015) Particle coating-dependent interaction of molecular weight fractionated natural organic matter: Impacts on the aggregation of silver nanoparticles. *Environ Sci Technol* 49(11):6581–6589.
  55. Handy RD, Owen R, Valsami-Jones E (2008) The ecotoxicology of nanoparticles and nanomaterials: Current status, knowledge gaps, challenges, and future needs. *Ecotoxicology* 17(5):315–325.
  56. Hristozov DR, Gottardo S, Critto A, Marcomini A (2012) Risk assessment of engineered nanomaterials: A review of available data and approaches from a regulatory perspective. *Nanotoxicology* 6(8):880–898.
  57. Weisberg M, Muldoon R (2009) Epistemic landscapes and the division of cognitive labor\*. *Philos Sci* 76(2):225–252.
  58. *CPlus*, via *STN AnaVist*, version 1.1 (Chemical Abstracts Service: Columbus, OH), 2006; AN 1995:429860. Available at: <https://scifinder.cas.org>. Accessed July 2, 2015.
  59. Amir S, Hafidi M, Merlina G, Hamdi H, Revel JC (2004) Elemental analysis, FTIR and <sup>13</sup>C-NMR of humic acids from sewage sludge composting. *Agronomie* 24(1):13–18.
  60. Gonzalez-Vila F, Lüdemann HD, Martin F (1983) <sup>13</sup>C-NMR structural features of soil humic acids and their methylated, hydrolyzed and extracted derivatives. *Geoderma* 31(1):3–15.
  61. Csardi G, Nepusz T (2006) The igraph software package for complex network research. Available at <https://pdfs.semanticscholar.org/1d27/44b83519657f5f2610698a8ddd177ced4f5c.pdf>. Accessed February 10, 2017.
  62. R Core Team (2013) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna).
  63. Harbour PJ, Dixon DR, Scales PJ (2007) The role of natural organic matter in suspension stability. 1. Electrokinetic-rheology relationships. *Colloids Surf A Physicochem Eng Asp* 295(1–3):38–48.
  64. Louie SM, Tilton RD, Lowry GV (2013) Effects of molecular weight distribution and chemical properties of natural organic matter on gold nanoparticle aggregation. *Environ Sci Technol* 47(9):4245–4254.
  65. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2013) nlme: Linear and non-linear mixed effects models (R Foundation for Statistical Computing, Vienna).
  66. Canty A, Ripley B (2016) boot: Bootstrap R (S-Plus) functions (R Foundation for Statistical Computing, Vienna).
  67. Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis (Springer, New York).
  68. Slowikowski K (2016) ggrepel: Repulsive text and label geoms for 'ggplot2'. Available at <https://cran.rstudio.com/web/packages/ggrepel/ggrepel.pdf>. Accessed February 10, 2017.