Glacier meltwater and monsoon precipitation drive Upper Ganges Basin dissolved organic matter composition

Jordon D. Hemingway [a,b,1,2], Robert G. M. Spencer[c], David C. Podgorski[d,3], Phoebe Zito[d,3], Indra S. Sen[e], Valier V. Galy[a]

[a]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, MA 02543, USA

[b]Massachusetts Institute of Technology – Woods Hole Oceanographic Institution Joint Program in Oceanography and Applied Ocean Science and Engineering, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[c]National High Magnetic Field Laboratory Geochemistry Group and Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL 32306, USA

[d]National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, FL 32310, USA

[e]Department of Earth Sciences, Indian Institute of Technology Kanpur, Kanpur, UP 208016, India

[1]To whom correspondence should be addressed. Tel.: +1 508 289 2821. E-mail address: jordon_hemingway@fas.harvard.edu (J.D. Hemingway)

[2]Current Address: Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, MA 02138, USA

[3]Current Address: Department of Chemistry, University of New Orleans, 2000 Lakeshore Drive, New Orleans, LA 70148, USA

KEYWORDS: dissolved organic matter | eco-hydrology | glaciers | Himalaya | monsoon

ABSTRACT

Mountain glaciers store dissolved organic carbon (DOC) that can be exported to river networks and subsequently respired to CO_2. Despite this potential importance within the global carbon cycle, the seasonal variability and downstream transport of glacier-derived DOC in mountainous river basins remains largely unknown. To provide novel insight, here we present DOC concentrations and molecular-level dissolved organic matter (DOM) compositions from 22 nested, glaciated catchments (1.4 – 81.8 % glacier cover by area) in the Upper Ganges Basin, Western Himalaya over the course of the Indian summer monsoon (ISM) in 2014. Aliphatic and peptide-like compounds were abundant in glaciated headwaters but were overprinted by soil-derived phenolic, polyphenolic and condensed aromatic material as DOC concentrations increase moving downstream. Across the basin, DOC concentrations and soil-derived compound class contributions decreased sharply from pre- to post-ISM, implying increased relative contribution of glaciated headwater signals as the monsoon progresses. Incubation experiments further revealed a strong compositional control on the fraction of bioavailable DOC (BDOC), with glacier-derived DOC exhibiting the highest bioavailability. We hypothesize that short-term (*i.e.* in the coming decades) increases in glacier melt flux driven by climate change will further bias exported DOM toward an aliphatic-rich, bioavailable signal, especially during the ISM and post-ISM seasons. In contrast, eventual decreases in glacier melt flux due to mass loss will likely lead to more a soil-like DOM composition and lower bioavailability of exported DOC in the long term.

1. INTRODUCTION

Mountainous river basins experience rapid rates of erosion, rock weathering, and organic carbon (OC) export, and are thus major drivers of the biogeochemical carbon cycle [*Milliman and Syvitski*, 1992; *Gaillardet et al.*, 1999; *Galy et al.*, 2015]. Despite this importance, the source and fate of dissolved OC (DOC) in mountainous rivers remains poorly constrained. Glaciated catchments are of particular interest since glaciers have been shown to provide nutrients and compositionally unique, highly bioavailable dissolved organic matter (DOM) to headwater streams [*Hood et al.*, 2009; *Singer et al.*, 2012; *Stubbins et al.*, 2012; *Spencer et al.*, 2014a; 2014b; *Hood et al.*, 2015]. It is estimated that mountain glaciers worldwide store approximately 70 Tg of DOC, with resulting meltwater runoff providing ≈ 0.6 Tg DOC yr^{-1} to fluvial networks [*Hood et al.*, 2015]. However, mountain glaciers are subject to major retreat and mass loss, both over glacial-interglacial cycles and in the coming centuries in response to climate change [*Bolch et al.*, 2012; *Bliss et al.*, 2014; *Lutz et al.*, 2014], with unknown consequences for DOC cycling in mountainous rivers.

This inability to predict carbon-cycle responses to changing glacier conditions is, at least in part, due to our limited understanding of the climatic and geomorphic controls on mountainous river DOC cycling. For example, seasonal precipitation trends are likely important drivers of DOC dynamics, but these controls have not yet been fully assessed. Additionally, downstream changes in catchment erosion rate, soil thickness, and soil pore-water residence time could influence the degree to which headwater DOC signals are overprinted by downstream soil inputs. However, riverine DOC signals are rarely interpreted within this geomorphic context, hindering our ability to isolate the role of glacier meltwater on carbon-cycle dynamics.

To provide novel insight, here we examined the spatial and seasonal evolution of DOC concentration, bioavailability, and DOM molecular composition in the Upper Ganges Basin. Located on the southern flank of the Western Himalaya, the Upper Ganges Basin is comprised of the Alaknanda and Bhagirathi Rivers, which combine to form the Ganges River (Fig. 1a). Both rivers are sourced from the Gangotri glacier group, one of the largest and best-monitored (in terms of area loss rate) glacier groups in the Himalayan range [*Bolch et al.*, 2012]. Additionally, the Upper Ganges Basin is strongly influenced by the Indian summer monsoon (ISM), which peaks in July and August and results in a roughly 5-fold increase in river discharge at this time [*Chakrapani and Saini*, 2009]. Both modeling and observational studies indicate that glacier meltwater contributes 10 to 30 % of total annual discharge in this system at the base of the Himalaya, while the remainder is derived primarily from ISM precipitation with supplemental snowmelt contribution during early summer months [*Bookhagen and Burbank*, 2010; *Maurya et al.*, 2010; *Immerzeel et al.*, 2013].

Extensive glacier coverage (Fig. 1b), combined with seasonal ISM influence (Fig. 1c), makes the Upper Ganges Basin an ideal location to assess the relative importance of precipitation and glacier melt as drivers of DOC concentration and DOM composition in mountainous rivers. To do so, here we report concentration and compositional results for samples collected from 22 main-stem and tributary locations spanning a ≈ 4 km elevation gradient starting at the Gangotri glacier terminus and ending in the Ganges River downstream of the Bhagirathi-Alaknanda confluence (Fig. 1a). Because main-stem geomorphic parameters such as catchment slope and soil thickness inherently co-vary with glacier coverage moving downstream, we interpret DOC results within a geomorphic context. By including a set of tributaries spanning a range of catchment areas, elevations, slopes, and glacial extents, our nested catchment approach aims to isolate the influence of glaciers on riverine DOC dynamics. In addition to evaluating spatial patterns, we investigated seasonal DOC and DOM variability by collecting samples across three seasons in 2014: pre-monsoon (April – May), ISM (June – September), and post-monsoon (October – December).

2. MATERIALS AND METHODS

2.1. Sample collection

Water was collected ≈ 5 m from the bank of each river and was immediately filtered through a pre-combusted (450 ºC, 4 hours) 0.45 μm glass fiber filter using an acid pre-leached (1.2 mol L^{-1} HCl, one week) Nalgene^{TM} filtration tower. Filtered water was transferred into either 60 mL polycarbonate (PC), 250 mL high density polyethylene (HDPE), or 1 L HDPE bottles (all acid pre-leached, 1.2 mol L^{-1} HCl, one week). The entire setup was rinsed (3×) with filtered river water before bottles were filled for sample collection. To constrain end-member DOC concentrations and DOM compositions, snowpack and glacier ice was additionally collected. During the pre-monsoon season, snowpack (2 locations) and glacier ice (1 location) samples were collected into 10 L bags using a pre-rinsed field hammer and immediately allowed to melt before being filtered as described above. At each snow/ice location, 4-5 aliquots were taken within a ≈ 1-2 m radius to provide a representative sample. All DOC samples were stored unacidified and were frozen within 48 h (typically < 24 h) and kept at -20 ºC until analysis. Because it is possible that small amounts of DOC were respired prior to sample freezing, concentrations reported here should be taken as conservative values.

2.2. DOC incubations

Seven pre-monsoon samples (1× glacier ice, 2× snowpack, 4× river water; Table S1) were subject to triplicate 28-day incubations as described previously [*Spencer et al.*, 2014b]. During sampling at each of these locations, filtered water (0.45 µm) was immediately transferred into 15× 20 mL pre-combusted (450 ºC, 4 hours) glass scintillation vials and allowed to incubate in the dark at room temperature (≈ 20 ºC). Because *in situ* temperatures varied significantly, incubating all samples at ≈ 20 ºC allows for more accurate comparisons of bioavailability between samples. Incubations were performed in the dark in order to inhibit growth of photoautotrophs, which would act to increase DOC concentrations and mask DOC losses due to heterotrophic respiration. At each time point (t = 0, 2, 7, 14, and 28 d), three vials were acidified dropwise using 12 mol L^{-1} HPLC-grade HCl until pH 2 was reached and were subsequently stored at room temperature until analysis. The initial time point (t = 0 d) for all samples was immediately acidified in the field. All waters were aerobic at the time of sampling and were unlikely to have become anaerobic during incubations. No biofilm formation or DOC flocculation was observed during incubations.

2.3. DOC quantification and extraction

All samples were measured for DOC concentrations (written [DOC]) via high-temperature combustion using a Shimadzu TOC-V organic carbon analyzer [*Mann et al.*, 2012]. After thawing at 4 ºC, water samples were acidified to pH 2 by adding 0.1% (v/v) concentrated HPLC-grade HCl to allow for removal of inorganic carbon (not necessary for incubation samples as they were previously acidified to pH 2) and were injected until the peak area coefficient of variance for three injections was < 2% (typically 3 – 5 injections). Areas were blank corrected using 18.2 MΩ Milli-Q water and were converted to [DOC] using a six-point standard calibration curve ranging from 0.10 – 10.00 mg L^{-1}. Both blanks and calibration standards were analyzed between every 10 samples. Long-term standard reproducibility indicates that results are precise to within ± 0.05 mg

L^{-1} (± 1σ) and that the detection limit for reliable quantification using this method is ≈ 0.10 mg L^{-1}. All results were thus rounded to the nearest increment of 0.05 mg L^{-1} and analytical uncertainty is assumed to be ± 0.05 mg L^{-1} throughout this study.

After quantification, all samples (*n* = 58; excluding incubations) were prepared for FT-ICR-MS analysis via solid-phase extraction (SPE) using 50 mg Bond Elut (Agilent Technologies) styrene-divinylbenzene copolymer (PPL) columns [*Dittmar et al.*, 2008]. Columns were cleaned and primed by soaking in HPLC-grade methanol overnight, rinsing with 2× cartridge volumes of 18.2 MΩ MilliQ water, 1× cartridge volume of methanol, and finally 2× cartridge volumes of acidified (pH 2) MilliQ water. Acidified samples (pH 2) were then eluted, and sample volumes were adjusted such that 10 μg of extractable carbon was loaded onto each column (assuming an average 50 % extraction efficiency). Lastly, columns were rinsed with 2× cartridge volumes of acidified MilliQ water, dried under a stream of ultra-high purity N_2 gas, and eluted with 250 μL HPLC-grade methanol into pre-combusted (450 ºC, 4 hours) vials. Similar to previous studies focusing on DOC-poor, glacier streams [*Spencer et al.*, 2014b], PPL extraction efficiencies could not be calculated due to limited sample volumes. However, extraction efficiencies are generally between 40 and 60 %, depending on sample source [*Dittmar et al.*, 2008]. We therefore assume that all samples analyzed here exhibited extraction efficiencies between 40 and 60 % despite compositional differences. Some samples did not contain sufficient volume to reach the 10 μg target due to low [DOC] (minimum of 25 % target mass; Table S1). To test if this range of PPL-extracted DOC mass affects FT-ICR-MS results, we additionally extracted one sample for which there existed sufficient material (glacier ice; Table S1) at 4× target volume, as discussed in Section 3.3, below.

2.4. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS)

The molecular-level composition of PPL-extracted DOM was determined using a custom-built 9.4 T FT-ICR MS equipped with a 22 cm horizontal bore ICR cell located at the National High Magnetic Field Laboratory (NHMFL, Florida State University, Tallahassee, FL) [*Kaiser et al.*, 2011a; 2011b; 2013]. Samples were directly infused to the mass spectrometer via an electrospray ionization (ESI) source at a flow rate of 0.5 μL min^{-1} to generate negatively-charged molecular ions. Negative ion mode results in deprotonation of acidic functional groups that are abundant in natural samples and is therefore best suited for untargeted DOM analysis. We note that biases against highly hydrophilic material during PPL extraction, combined with the poor ionization efficiency of these compounds, potentially biases resulting mass spectra. However, these effects have been shown to be minor in natural freshwater DOM samples [*Raeke et al.*, 2016]. Experimental parameters were optimized for DOM analysis (-2.5 kV needle voltage, -300 V tube lens, 8 W heated metal capillary) [*Stenson et al.*, 2003]. Ion accumulation time per scan was adjusted following *O'Donnell et al.* (2016) to account for differences in PPL-extracted [DOC] due to sample limitation and variable extraction efficiency, with longer integration times for dilute samples leading to approximately constant total ion current across all samples. Each mass spectrum was the sum of 100 individual co-added spectra. Samples were measured in a random order, and reproducibility was estimated by analyzing an arbitrarily chosen subset ($n = 5$) of samples in triplicate.

Molecular formulae were assigned to signals > 6σ root mean square baseline noise and with mass errors below 500 ppb [*O'Donnell et al.*, 2016]. Formulae were determined using the EnviroOrg© ™ (*Corilo*, 2015) following published rules [*Koch et al.*, 2007], and all elemental combinations within $C_{1\text{-}45}H_{1\text{-}92}N_{0\text{-}4}O_{1\text{-}25}S_{0\text{-}2}$ were considered for assignment. To classify formulae within compound classes, a modified version of the aromaticity index (AI_{mod}) first presented by *Koch and Dittmar* [2006] was calculated for each formula as

$$AI_{mod} = \frac{1+C-S-0.5(O+H+N)}{C-0.5O-N-S}. \qquad \text{(Eq. 1)}$$

Formulae were then classified based on elemental stoichiometries and AI_{mod} values as follows: condensed aromatic, $AI_{mod} > 0.67$; polyphenolic, $0.67 \geq AI_{mod} > 0.5$; unsaturated phenolic high oxygen content, $H/C < 1.5$, $O/C \geq 0.5$; unsaturated phenolic low oxygen content, $H/C < 1.5$, $O/C < 0.5$; peptide-like, $H/C \geq 1.5$, $N \geq 1$; aliphatic, $H/C \geq 1.5$, $N = 0$ [*Santl-Temkiv et al.*, 2013; *Spencer et al.*, 2014b]. We note that peptide assignments can be ambiguous since N-containing compounds are additionally present in alternative isomeric arrangements. Additional classification constraints specifically incorporating phosphorus content have recently been calibrated using biomass extract and have been shown to increase classification accuracy [*Rivas-Ubach et al.*, 2018]. However, because phosphorus-containing compounds are typically low in abundance and not easily resolved within mountainous headwater DOM samples [*Spencer et al.*, 2014b], here we retain the classification scheme of *Santl-Temkiv et al.* (2013). Finally, the relative abundance of each compound class was determined by rescaling peak intensities such that the total ion count for the entire mass spectrum is equal to unity and calculating the intensity-weighted sum of all peaks within a given compound class.

2.5. Geospatial and statistical analysis

Geospatial data for all sites were analyzed using the Geographic Resources Analysis Support System software (GRASS v7.2). Catchment areas and geomorphic parameters upstream of each sampling location were calculated using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) global digital elevation model with 90 m spatial resolution (Fig. 1a) [*Jarvis et al.*, 2008]. Average catchment slope was calculated as the mean value of the slope for all pixels within a given catchment area. Catchment relief ratio at each sampling location was determined as the maximum elevation difference divided by the upstream main-stem distance. Glacial extent within each catchment was estimated using the Randolph Glacier Inventory (RGI) v5.0 database [*RGI Consortium*, 2015] and converted to percent areal coverage (Fig. 1b). Precipitation estimates were generated using re-analyzed tropical rainfall monitoring mission (TRMM) data following *Bookhagen and Burbank* (2010) (Fig. 1c).

Samples were divided into seasonal groups and temporal variability in DOC concentration/composition was assessed using one-way analysis of variance (ANOVA). Reported *p*-values for temporal trends represent the probability of falsely rejecting the null hypothesis that there exists no seasonal variability. Because we do not expect DOC concentration/composition to be a linear function of catchment properties *a priori*, and because FT-ICR MS compositional results are only semi-quantitative, all spatial trends were assessed using non-parametric rank correlation unless otherwise stated. Resulting Spearman correlation coefficients (ρ_S) denote the strength of any monotonically increasing/decreasing relationship, and corresponding *p*-values represent the probability that no relationship exists.

3. RESULTS

3.1. Geomorphic parameters

All geomorphic parameters are reported in Table S1. Sample sites represent a ≈ 4 km elevation transect, ranging from a minimum of 338 meters above sea level (masl) to a maximum of 3961 masl. Mean catchment elevation upstream of each sample site correspondingly ranged from 2900 masl to 5400 masl, while catchment area varied by roughly two orders of magnitude from 172 km^2 to 21,789 km^2. Tributary sampling locations (n = 7 sites) spanned ≈ 70 % of the main-stem sample elevation range (653 masl to 3175 masl; n = 15 sites), although tributary catchment areas only reached a maximum of 3026 km^2 (*i.e.* 14 % of the most down-stream main-stem sampling location).

Tributary and main-stem catchment relief ratio exhibited similar variability, ranging from 45 m km^{-1} to 119 m km^{-1} (average = 90 ± 30 m km^{-1}) and 21 m km^{-1} to 94 m km^{-1} (average = 49 ± 22 m km^{-1}), respectively. However, because all sites are contained within the Himalayan range, catchments were consistently steep and resulting average slope exhibited only modest variability, ranging from 24.2º to 31.9º (average = 28.1 ± 1.5 º). Catchment slope was thus uncorrelated with catchment area, sampling elevation, and relief ratio, but did exhibit a slight negative correlation with mean catchment elevation ($\rho_s = -0.53$; $p = 9.7\times10^{-3}$) and with glacier coverage ($\rho_s = -0.55$; $p = 6.0\times10^{-3}$). This negative correlation results from the fact that high-elevation, highly glaciated sites contained significant areas of low-slope glacial valleys, thus lowering the mean catchment slope.

Main-stem sites spanned a wide range in glacial coverage, from 10.3 to 81.8 % (average = 25.8 ± 18.8 %), while tributaries ranged from 1.4 to 44.8 % (average = 13.0 ± 15.3 %; Fig. 1b). Because main-stem glacial coverage inherently decreases as sample sites move downstream, there exists significant non-linear covariance with geomorphic parameters such as catchment area ($\rho_s = -0.99$; $p = 2.1\times10^{-33}$), mean elevation ($\rho_s = 0.99$; $p = 8.2\times10^{-36}$), and relief ratio ($\rho_s = 0.94$; $p = 6.3\times10^{-18}$). However, by including tributary streams in addition to main-stem samples, our nested catchment approach allows for separation of glacier coverage and geomorphic parameters. Tributary glacial coverage exhibited no significant correlation with catchment area or relief ratio, and considerably weaker correlation with mean catchment elevation ($\rho_s = 0.58$; $p = 6.6\times10^{-3}$). Thus, when considering the entire sample set (*i.e.* both tributaries and main-stem sites), glacier coverage was uncorrelated with both relief ratio and catchment area, allowing us to independently assess the influence of these controls on resulting DOM signals.

3.2. DOC concentration

For the entire dataset, [DOC] ranged from a minimum of 0.10 mg L^{-1} to a maximum of 0.70 mg L^{-1} with a mean value of 0.29 ± 0.16 mg L^{-1} ($n = 55$; $\mu \pm 1\sigma$ uncertainty; Table S1). [DOC] displayed no statistically significant difference between main-stem and tributary sites (t-test for equal means: $p > 0.05$, $T = 0.51$, degrees of freedom = 35), with main-stem samples averaging 0.29 ± 0.17 mg L^{-1} ($n = 38$) and tributary samples averaging 0.31 ± 0.15 mg L^{-1} ($n = 17$). For all catchments, [DOC] decreased sharply from the pre-monsoon to the post-monsoon seasons. Mean values dropped from 0.39 ± 0.16 mg L^{-1} ($n = 19$) during the pre-monsoon to 0.18 ± 0.08 mg L^{-1} ($n = 17$) during the post-monsoon ($p = 2.8\times10^{-4}$; Fig. 2a), although we note that we were not able to sample all sites in all seasons. Still, the temporal [DOC] decrease remains statistically significant when only sites that were sampled in all seasons are considered ($n = 14$ sites; $p = 2.2\times10^{-4}$), indicating that this observed trend was not the result of sampling biases. Furthermore, the lack of statistically significant difference between main-stem and tributary [DOC] holds when data are separated by season (t-test for equal means: $p > 0.05$ in all cases), indicating that this result was not biased by the inclusion of data collected across multiple seasons.

3.3. DOM composition

FT-ICR MS resulted in the detection of 28,629 unique molecular formulae across our sample set, with individual samples containing between 7,392 and 15,198 formulae (average = 11,544 ± 1,917; $n = 58$; $\mu \pm 1\sigma$ uncertainty; Table S2). Triplicate measurements resulted in < 7.4 % variability (1–3 % for most samples) in the number of total detected formulae as well as the number of formulae assigned to each compound class, indicating minimal analytical uncertainty. Additionally, results from glacier ice analyzed at 1× and 4× concentration were identical within uncertainty (Table S1), indicating that the range of concentrations for samples presented in this study had no effect on peak detection and calculated compound class abundances.

For all riverine samples, DOM molecular diversity, as measured by formula number, decreased from pre- to post-monsoon ($p = 7.6\times10^{-8}$; Fig. 3a) and was positively correlated with [DOC] ($\rho_s = 0.76$; $p = 1.4\times10^{-11}$; Fig. 3b). The majority of DOM in all riverine samples was classified as unsaturated phenolic compounds with high oxygen content (average = 41.1 ± 7.9 %; $n = 55$; $\mu \pm 1\sigma$ uncertainty; Table S1) or unsaturated phenolic compounds with low oxygen content (average = 44.0 ± 8.5 %). Although lower in abundance than unsaturated phenolic compounds, aliphatic and polyphenolic material contributed up to 15.1 % (average = 4.9 ± 2.8 %) and 11.5 % (average = 7.4 ± 2.6 %) of fluvial DOM, respectively. Both condensed aromatic and peptide compound classes were significantly less abundant, contributing only 1.7 ± 0.8 % and 1.0 ± 1.0 %, respectively.

In contrast to fluvial samples, snowpack and glacier melt samples contained significantly lower contributions by high- and low-oxygen unsaturated phenolic compound classes at 18.4 ± 3.1 % and 32.5 ± 5.6 %, respectively (*n* = 3; Table S1). Rather, these samples were described by high relative contributions of aliphatic (average = 23.9 ± 5.8 %) and peptide-like material (average = 20.4 ± 3.0 %) and significantly lower contributions of condensed aromatic (average = 0.8 ± 0.1 %) and polyphenolic (average = 3.5 ± 0.6 %) compound classes.

4. DISCUSSION

4.1. Controls on Concentration

Our dataset reveals that Upper Ganges Basin [DOC] varied significantly as a function of season and glacier coverage (Fig. 2). Large seasonal hydrologic variability in this region likely exhibits a strong control on the relative contributions of glacier-, snow-, and soil-derived DOC to exported riverine signals. For example, warming air temperatures during early summer months, combined with expansive snow cover from late monsoon and winter precipitation, should lead to increased snowmelt-derived discharge at this time. Both observations [*Maurya et al.*, 2010; *Andermann et al.*, 2012] and modeling results [*Lutz et al.*, 2014] from this region indicate that up to ≈ 75 % of discharge during Apr-May-Jun is derived from surface runoff due to snowmelt. In contrast, ISM rainfall and glacier meltwater are expected to dominate monsoon-season discharge, when both temperature and precipitation reach annual maxima [*Andermann et al.*, 2012] (Fig. 1c). Observed seasonal [DOC] trends (Fig. 2a) are thus consistent with hydrologic variability. Elevated concentrations during the pre-monsoon season were likely a result of increased surface soil pore water residence time, as snowmelt is expected to slowly percolate through DOM-rich soil pore-waters and surface litter layers that have received organic matter inputs but have not yet been extensively flushed [*McGlynn and McDonnell*, 2003; *Inamdar et al.*, 2006; *Spencer et al.*, 2010]. Conversely, during the ISM, higher discharge and short hydraulic retention times on the landscape would result in a bias toward DOM-poor rainwater and glacier meltwater, thus diluting soil-derived inputs.

Post-monsoon samples exhibited the lowest [DOC] for all but one sampling location (Table S1). This result is unlikely to be caused by a simple dilution effect since pre- and post-monsoon seasons are described by nearly identical discharge regimes [*Chakrapani and Saini,* 2009]. Rather, it has been shown in nearby catchments that post-monsoon discharge is dominated by the flushing of transient fractured basement groundwater aquifers that have accumulated during the ISM [residence time ≈ 45 d; *Andermann et al*., 2012]. Low [DOC] at this time implies either that aquifer recharge was derived from DOC-poor sources such as rainwater and glacier meltwater [*Hood et al.* 2015], that groundwater has lost DOC during its ≈ 45 d transit through the bedrock (*e.g.* due to respiration), or a combination of both.

In addition to temporal variability, [DOC] exhibited a significant negative relationship with glacial coverage for all samples across all seasons ($\rho_s = -0.57$; $p = 5.5\times10^{-6}$; Fig. 2b). Proglacial streams and highly glaciated catchments exhibited the lowest [DOC] (0.10 mg L^{-1}), while modestly glaciated tributaries and downstream main-stem samples reached 0.60 mg L^{-1} and 0.70 mg L^{-1}, respectively. This relationship is non-linear, with [DOC] typically remaining below ≈ 0.30 mg L^{-1} until glacial coverage has dropped below ≈ 20 %. Because main-stem glacial coverage inherently decreases moving downstream (Fig. 1b), it remains possible that this correlation reflects a shift in soil inputs due to changing geomorphic parameters rather than glacier extent *per se*.

Because our nested sample approach includes samples from tributary sites whose geomorphic parameters are uncorrelated with glacier extent, we were able to independently assess the geomorphic and glacial controls on [DOC]. Following *Moore et al.* [1993], we chose catchment slope as a proxy for soil thickness and hydrologic retention time on the landscape. Catchment slope was uncorrelated with [DOC] across our dataset ($p > 0.05$), including both main-stem and tributary sites, suggesting that soil thickness alone cannot explain observed concentration trends. We further tested the effect of *in situ* processing during stream transit by treating relief ratio, defined as the change in elevation per unit of stream length, as a proxy for in-stream residence time. Catchment relief ratio was uncorrelated with glacier coverage ($p > 0.05$), making it an ideal independent geomorphic metric. While [DOC] did decrease with increasing relief ratio across the entire sample set ($\rho_s = -0.41$; $p = 2.2\times10^{-3}$; Fig. S1), this relationship was weaker than that with glacial coverage (Fig. 2b). Furthermore, unlike glacier extent, relief ratio was uncorrelated with [DOC] in tributary samples ($p > 0.05$). It is therefore unlikely that observed spatial

[DOC] trends simply reflect shifting geomorphic parameters. Rather, we conclude that DOC-poor glacier meltwater is an important driver of downstream DOC concentrations in the Upper Ganges Basin.

4.2. Compositional Trends

In addition to DOC concentration trends, we observed large spatiotemporal variability in DOM molecular composition within the Upper Ganges Basin (Figs. 3-5). Higher DOM molecular diversity with increasing [DOC] indicates the addition of a chemically unique downstream source, especially during the pre-monsoon season, while low post-monsoon diversity suggests increased relative contribution of headwater signals. Diversity trends are unlikely to be driven by photodegradation since glacier-fed headwater DOM is described by low UV-visible absorbance [*Stubbins et al.*, 2012], while high turbidity [*Chakrapani and Saini*, 2009] and short travel distances (≤ 206 km; Table S1) in these rivers further inhibit interaction with light.

To characterize DOM compositional trends, we examined changes in the relative abundances of formulae that were detected by FT-ICR MS in all river samples (n = 4,990, or 17 % of total formulae) when correlated with [DOC], glacier coverage, season, and relief ratio. Of these 4,990 formulae, 84 % were significantly correlated with [DOC] ($p < 0.05$), with an average absolute-value Spearman correlation coefficient (written as $|\rho_s|$) of 0.66 ± 0.17 ($\mu \pm 1\sigma$; Fig. 4a). Both the percentage of significantly correlated formulae and the average correlation strength decreased slightly when correlated with glacial coverage (76 %; $|\rho_s| = 0.55 \pm 0.15$; Fig. 4b) and season (72 %; $|\rho_s| = 0.49 \pm 0.13$; Fig. 4c) but decreased sharply when correlated with relief ratio (47 %; $|\rho_s| = 0.33 \pm 0.04$; Fig. S2). Any relationships between relative abundance and relief ratio do not simply reflect auto-correlation with glacial coverage, as our nested catchment approach ensured that there was no correlation between relief ratio and glacial coverage (Section 3.1.). Thus, the observation that relief ratio explains less variability than does glacial coverage, both in terms of formula number and correlation strength, indicates that compositional trends do not simply reflect downstream changes in catchment geomorphology. Rather, DOM molecular composition is a strong function glacier coverage due to the contribution of compositionally unique, low [DOC] glacier meltwater.

For all environmental parameters, formulae exhibiting similar ρ_s values were tightly clustered in van Krevelen space. Formulae with high H/C and low O/C were positively correlated with glacial coverage, season, and catchment relief ratio and were negatively correlated with [DOC], whereas low H/C and high O/C formulae displayed the opposite trend (Fig. 4, S2). This nearly identical compositional response to glacial coverage and season (Fig. 4b-c) strongly suggests that exported DOM becomes biased toward a glaciated, headwater signal during the ISM and, especially, post-monsoon seasons. This observed bias toward glaciated signals is consistent with previous studies that have related DOM composition to ^{14}C content and bioavailability in glacier-fed streams and have concluded that glacier-derived DOM is rich in highly bioavailable, aliphatic compounds (*i.e.* high H/C, low O/C) [*Hood et al.*, 2009; *Singer et al.*, 2012; *Spencer et al.*, 2014a; 2014b]. In contrast, the observed decreasing relative abundance of these compounds with increasing [DOC] (Fig. 4a) provides further evidence for downstream admixture of relatively unsaturated, aromatic DOM from surface litter and organic rich soil layers [*McGlynn and McDonnell*, 2003; *Inamdar et al.*, 2006; *Spencer et al.*, 2010].

To quantify spatiotemporal trends, we categorized each detected formula as aliphatic, peptide-like, unsaturated phenolic (both high- and low-oxygen content), condensed aromatic, or polyphenolic based on published classification schemes (see Section 2.4, above) [*Santl-Temkiv et al.*, 2013; *Spencer et al.*, 2014b]. Although the majority of compounds detected in stream samples (≥ 74 %, Table S1) were classified as unsaturated phenolic, here we focus on aliphatics, condensed aromatics, and polyphenolics since glacier and soil sources contain characteristic proportional contributions of these compound classes [*Singer et al.*, 2012; *Stubbins et al.*, 2012; *Spencer et al.*, 2014b]. For example, microbially derived DOM that is characteristic of glacier-sourced material is expected to be rich in aliphatics relative to soil-derived inputs [*Singer et al.*, 2012; *Stubbins et al.*, 2012; *Spencer et al.*, 2014b]. In contrast, DOM derived from the leaching of higher plant material in organic-rich soil horizons has been shown to exhibit higher proportions of phenolic and polyphenolic material [*O'Donnell et al.*, 2016; *Rivas-Ubach et al.*, 2018; *Stubbins et al.*, 2012].

Both polyphenolic and condensed aromatic relative abundances declined significantly as the ISM progressed ($p = 2.5\times10^{-5}$ and 3.4×10^{-6}, respectively; Fig. 5a-b). Additionally, these compound classes increased in relative abundance with increasing [DOC] ($\rho_s = 0.87$ and 0.83; $p = 5.1\times10^{-18}$ and 3.0×10^{-18}, respectively; Fig. 5d-e), decreased with increasing glacial coverage ($\rho_s = -0.62$ and -0.59; $p = 3.9\times10^{-7}$ and 2.1×10^{-6}; Fig. 5g-h), and displayed no correlation with catchment relief ratio ($p > 0.05$; Fig. S3a-b).

For highly glaciated catchments, both polyphenolic and condensed aromatic relative abundances approached their glacier/snowpack end-member values (0.8 ± 0.1 % condensed aromatic; 3.5 ± 0.6 % polyphenolic; $n = 3$; Fig. 5; Table S1), confirming that meltwater was the predominant headwater DOM source. In contrast, soil organic matter has been shown to contain high relative abundances of condensed aromatic (*e.g.* combustion products, black carbon) [*Jaffe et al.*, 2013] and polyphenolic (*e.g.* vascular-plant lignin) [*Stubbins et al.*, 2012; *O'Donnell et al.*, 2016] compounds. Strong enrichment in both classes with increasing [DOC] and decreasing glacier coverage further indicated downstream incorporation of soil-derived DOM and/or decomposition of glacier-derived DOM. In agreement with concentration (Fig. 2a) and chemical diversity (Fig. 3a) trends, temporal decreases in the relative abundance of these classes require that soil inputs become less important during the ISM and post-monsoon seasons. We therefore hypothesize that elevated precipitation during the ISM (Fig. 1c) increases surface flow rates and thus decreases hydraulic residence time in soil pore-waters, leading to less overprinting of headwater signals. Additionally, it has been shown that groundwater aquifers in this region are recharged during the ISM and exhibit a ≈ 45-day residence time [*Andermann et al.*, 2012]. Thus, while precipitation rates are low during the post-ISM season (Fig. 1c), large groundwater inputs could explain the continued decrease in soil-like DOM signatures at this time. This interpretation is consistent with seasonal [DOC] trends, which also reach minimum values during the post-ISM seasons (Section 4.1).

The relative contribution of aliphatic material increased with glacier cover ($\rho_s = 0.71$; $p = 1.1\times10^{-9}$) and decreased sharply with [DOC] ($\rho_s = -0.74$; $p = 9.8\times10^{-11}$; Fig. 5f, 5i). This trend agrees with previous studies showing that these compounds are abundant in depositional DOM sources and are produced in high quantities by active supraglacial, subglacial and proglacial microbial communities [*Sharp et al.*, 1999; *Bhatia et al.*, 2006; *Singer et al.*, 2012; *Stubbins et al.*, 2012; *Spencer et al.*, 2014b]. However, heavily glaciated catchments never reached the measured glacier/snowpack end-member value (23.9 ± 5.8 %; Fig. 5; Table S1), likely due to the high bioavailability of this material [*Hood et al.*, 2009; *Singer et al.*, 2012; *Spencer et al.*, 2014b] and large heterogeneity within glacier ecosystems (Fig. 5i) [*Bhatia et al.*, 2006; *Wilhelm et al.*, 2013]. Aliphatic compounds have also been shown to degrade rapidly in both glacier-derived [*Singer et al.*, 2012] and permafrost-derived DOM [*Spencer et al.*, 2015]. This is consistent with our observations and likely explains the lack of temporal trend in aliphatic abundance (Fig. 5c), in contrast to all other observed signals.

4.3. Bioavailability trends

Upper Ganges Basin DOC bioavailability additionally exhibited large variability. To compare with literature results [Hood *et al*., 2009; Singer *et al.*, 2012; Spencer *et al.*, 2014b], here we calculated bioavailable DOC (% BDOC) as the average relative decrease in [DOC] between $t = 0$ d and $t = 28$ d for triplicate samples. Intermediate time points were used to verify that DOC decay was first-order with respect to carbon concentration (*i.e.* exponential decay), as is expected for first-order decay processes such as biological utilization (Fig. 6). Consistent with other studies, incubations were terminated at $t = 28$ d in order to capture the entire decay profile. That is, the concentration of DOC remaining at $t = 28$ d approached an asymptotic value, as shown in Fig. 6.

BDOC ranged from 32.8 % (Ganges at Rhishikesh, corresponding to 0.23 mg C L^{-1}) to 59.7 % (Gangotri Glacier at Gomukh, corresponding to 0.06 mg C L^{-1}) for river samples and averaged 60.5 ± 6.1 % for glacier ice and snowpack samples (corresponding to 0.83 ± 0.32 mg C L^{-1}; $n = 3$; Table S1). Although photochemical processes could increase BDOC relative to light-free incubation results reported here, interaction with light is likely minimal in these streams due to the low UV-absorbance of mountainous headwater DOM, high turbidity [*Chakrapani and Saini*, 2009] and short *in situ* residence times. Similar to trends observed in previous studies [*Hood et al.*, 2009; *Singer et al.*, 2012; *Spencer et al.*, 2014b], % BDOC increased significantly with increasing glacial coverage (Fig. 7a) and was strongly correlated with DOM chemical composition. Interestingly, the BDOC range and relationships with chemical composition presented here are similar to those observed from the Tibetan Plateau [*Spencer et al.*, 2014b] despite the difference in filtration pore size (0.45 μm in this study; 0.7 μm in *Spencer et al.*, 2014b). Although future work is needed to more directly to test this result, this similarity suggests that small differences in heterotroph cell size do not exert a first-order control on DOC respiration rates in mountainous streams.

To assess bioavailability as a function of composition, we regressed % BDOC against polyphenolic relative abundance using ordinary least squares ($r^2 = 0.83$; $p = 4.4\times10^{-3}$; $n = 7$; Fig. 7b). We emphasize that FT-ICR MS results are only semi-quantitative due to, for example, ion suppression effects (see Section 2.4, above) and that resulting composition-bioavailability regression relationships are likely not truly linear. Nonetheless, the linear regressions performed here remain useful for qualitatively understanding seasonal BDOC variability in the absence of more quantitative measurements. We chose polyphenolic abundance as a pseudo-conservative tracer since it exhibits little variability in the glacier/snowpack end-member (3.1 – 4.2 %; $n = 3$; Fig. 5g; Table S1) and is likely to exhibit minimal degradation during transit in this system. In contrast, glacier/snowpack aliphatic abundance is highly variable (19.3 – 30.4 %; $n = 3$; Fig. 5i; Table S1) and behaves non-conservatively, likely due to rapid consumption [*Spencer et al.*, 2015]. Still, we note that regressing % BDOC against condensed aromatic relative abundance yields identical results to those calculated here within uncertainty, further supporting the idea that bioavailability is a function of chemical composition in these samples.

Assuming the observed BDOC vs. composition relationships hold for all seasons, we used the measured polyphenolic relative abundance for each sample to predict temporal changes in bioavailability. For all sites in which samples were collected for all seasons ($n = 14$), we find that BDOC increased from an average of 39 ± 4 % during the pre-monsoon to 54 ± 5 % during the post-monsoon (Table S1). This increase in bioavailability partially balances the observed decrease in [DOC] throughout the course of the ISM (Fig. 2a), leading to a modest decrease in BDOC concentration of only 0.06 ± 0.05 mg L^{-1} from pre- to post-monsoon seasons ($p = 1.9\times10^{-3}$). We again emphasize that predicted seasonal BDOC trends are based on composition-bioavailability regressions (Fig. 7b) and are thus likely subject to large, unknown uncertainty. Still, these results are consistent with previous studies [*Singer et al.*, 2012] and imply only minimal seasonal variability in BDOC concentrations throughout the Upper Ganges Basin despite large temporal [DOC] trends due to increased relative contribution of bioavailable headwater sources during the ISM and, especially, post-ISM seasons.

4.4. Carbon-cycle implications and global significance

The observed spatiotemporal influence of glacier-derived DOC on mountainous river carbon cycling is likely not limited to the Upper Ganges Basin. For example, similar to our results, *Spencer et al.* (2014b) showed that DOC in glaciers and glacial streams on the Tibetan Plateau contained 12 – 16 % aliphatic relative abundance and 46 – 69 % BDOC (Fig. 7a). We therefore hypothesize that increased contribution of bioavailable, glacier-derived DOC during the ISM and post-ISM seasons is a common feature within Himalayan rivers.

To assess DOC dynamics at the regional scale, we estimate the DOC flux exiting the Himalaya and entering the Ganges floodplain. Because discharge measurements for the year 2014 at our sampling locations are not available, we approximate DOC yields using season-specific discharge from 2002 – 2004 at nearby gauging stations [*Chakrapani and Saini*, 2009]. Although this approach will introduce large uncertainties, ISM precipitation and river discharge in the Himalaya exhibit minimal inter-annual variability [*Andermann et al.*, 2012], and resulting yield estimates are likely robust within an order of magnitude. Results are sparse ($n = 12$) yet show a consistent increase in DOC yield moving downstream and a general increase during the ISM season (Table S3). By combining all data points into a single rating curve (Fig. S4) and using an annual average discharge at our most downstream site of ~ 750 m^3 s^{-1} [*Chakrapani and Saini*, 2009], we estimate a flux of ~ 0.01 Tg DOC yr^{-1} and a yield of ~ 500 kg DOC km^{-2} yr^{-1} at the base of the Himalaya. Assuming a similar yield for nearby Himalayan rivers, this corresponds to ~ 0.1 Tg DOC yr^{-1} exported from the Himalayan Range into the Ganges Floodplain. This yield is roughly four-fold lower than for the entire Ganges-Brahmaputra (G-B) Basin [~2200 kg DOC km^{-2} yr^{-1}; *Ludwig et al.*, 1996], consistent with our interpretation that Himalayan DOC is dominated by low-concentration ISM precipitation and glacier meltwater sources, with little contribution from flushing of surface soils and litter layers except during the pre-monsoon season.

Furthermore, although more work is needed to reduce uncertainty, extend temporal records, and quantify DOC fluxes, our results can begin to inform predictions on future DOC cycling in the Upper Ganges Basin in particular and in Himalayan rivers more generally. Assuming secular trends mimic seasonal variability in terms of DOC source and composition, we expect future increases in glacier melt flux to bias exported DOM compositions toward aliphatic-rich, glaciated headwater signals in the short term (*i.e.* until ~ 2050; *Immerzeel et al.*, 2013). In contrast, continued warming will eventually lead to glacier mass loss and reduced meltwater fluxes [*Bolch et al.*, 2012; *Immerzeel et al.*, 2013; *Bliss et al.*, 2014; *Lutz et al.*, 2014], likely resulting in higher DOC concentrations and more soil-like composition in the long term (*i.e.* a bias toward pre-monsoon conditions). However, we suggest that concomitant decreases in the fraction of DOC that is bioavailable will dampen BDOC concentration variability, thus stabilizing the absolute flux of CO_2 produced from DOC respiration in this system.

Finally, the seasonal importance of glacier-derived DOM to headwater streams is likely not limited to those draining the Himalaya. For example, *Spencer et al.* (2014a) observed a depletion in ^{14}C and an enrichment in protein-like fluorescence of DOM exported from Mendenhall Glacier, southeast Alaska, during the glacial melt season relative to the rest of the annual cycle. Combined with the strong negative relationship between ^{14}C content and bioavailability in samples from the same location [*Hood et al.*, 2009], these temporal trends suggest increased relative contribution of highly bioavailable, glacier-derived DOM during the glacier melt season, consistent with our Upper Ganges Basin results.

5. CONCLUSION

Using samples collected throughout the Upper Ganges Basin in 2014, we show that DOC concentrations and DOM molecular compositions can exhibit large spatiotemporal variability in glaciated, mountainous headwater streams. Our results revealed a sharp decrease in DOC concentrations, aliphatic relative abundances, and condensed aromatic relative abundances with increasing glacial coverage across all seasons. In contrast, aliphatic relative abundances exhibited the opposite trend. Similar to spatial variability, DOC concentrations, aliphatic relative abundances, and condensed aromatic relative abundances decreased progressively from pre-ISM to ISM to post-ISM seasons. This observed similarity in spatial and temporal variability suggests increased downstream propagation of headwater-derived, glacier-influenced DOM as the monsoon progresses.

Previous studies have indicated that glacier meltwater provides highly bioavailable, aliphatic- and protein-rich DOM to headwater streams just below the glacier terminus. As our results demonstrate for the first time, these signals can propagate downstream for hundreds of kilometers, especially when monsoon rains decrease soil pore-water residence times and thus lower soil-derived DOM contributions. Furthermore, although more work is needed to better quantify seasonal shifts in bioavailability, our results imply that downstream soil-derived inputs are higher in DOC concentration but are less bioavailable than glacier-derived headwater signals. Consequently, we suggest that shifts in DOC concentration and bioavailability due to future glacier melt will largely counteract each other, thus stabilizing the absolute CO_2 emission flux from DOC respiration in Himalayan rivers.

ACKNOWLEDGEMENTS

We thank Britta Voss (WHOI) for assisting with sample collection; Travis Drake (FSU), and Ekaterina Bulygina (Woods Hole Research Center) for laboratory assistance; and the NHMFL ICR user program (NSF-DMR-1157490) for aiding in data acquisition and analysis. This study was partly supported by NSF-DEB-1145932 to R.G.M.S. J.D.H. was partially supported by the NSF Graduate Research Fellowship Program under grant number 2012126152, with additional support in the form of travel grants awarded by the MIT Houghten Fund and NHMFL. All data used in this study are available in the Supporting Information Tables S1 and S2.

REFERENCES

Andermann, C., L. Longuevergne, S. Bonnet, A. Crave, P. Davy, and R. Gloaguen (2012), Impact of transient groundwater storage on the discharge of Himalayan rivers, *Nat Geosci*, *5*, 127–132.

Bhatia, M., M. Sharp, and J. Foght (2006), Distinct bacterial communities exist beneath a high Arctic polythermal glacier, *Appl. Environ. Microbiol.*, *72*, 5838–5845.

Bliss, A., R. Hock, and V. Radić (2014), Global response of glacier runoff to twenty-first century climate change, *J. Geophys. Res.*, *119*, 717–730.

Bolch, T. et al. (2012), The State and Fate of Himalayan Glaciers, *Science*, *336*, 310–314.

Bookhagen, B., and D. W. Burbank (2010), Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge, *J. Geophys. Res.*, *115*, F03019.

Chakrapani, G. J., and R. K. Saini (2009), Temporal and spatial variations in water discharge and sediment load in the Alaknanda and Bhagirathi Rivers in Himalaya, India, *J. Asian Earth Sci.*, *35*, 545–553.

Corilo, Y. (2015) EnviroOrg. Florida State University, Tallahassee, FL, USA.

Dittmar, T., B. P. Koch, N. Hertkorn, and G. Kattner (2008), A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater, *Limnol. Oceanogr. Methods*, *6*, 230–235.

Gaillardet, J., B. Dupré, P. Louvat, and C. J. Allègre (1999), Global silicate weathering and CO_2 consumption rates deduced from the chemistry of large rivers, *Chem. Geol.*, *159*, 3–30.

Galy, V. V., B. Peucker-Ehrenbrink, and T. I. Eglinton (2015), Global carbon export from the terrestrial biosphere controlled by erosion, *Nature*, *521*, 204–207.

Hood, E., J. Fellman, R. G. Spencer, P. J. Hernes, R. Edwards, D. D’Amore, and D. T. Scott (2009), Glaciers as a source of ancient and labile organic matter to the marine environment, *Nature*, *462*, 1044–1047.

Hood, E., T. J. Battin, J. Fellman, S. O'Neel, and R. G. Spencer (2015), Storage and release of organic carbon from glaciers and ice sheets, *Nat Geosci*, *8*, 91–96.

Immerzeel, W. W., F. Pellicciotti, and M. F. P. Bierkens (2013), Rising river flows throughout the twenty-first century in two Himalayan glacierized watersheds, *Nat Geosci*, *6*, 742–745.

Inamdar, S. P., N. O'Leary, M. J. Mitchell, and J. T. Riley (2006), The impact of storm events on solute exports from a glaciated forested watershed in western New York, USA, *Hydrol. Process.*, *20*, 3423–3439.

Jaffe, R., T. Dittmar, Y. Ding, J. Niggemann, A. V. Vahatalo, A. Stubbins, R. G. Spencer, and J. Campbell (2013), Global charcoal mobilization from soils via dissolution and riverine transport to the oceans, *Science*, *340*, 345–347.

Jarvis A., H. I. Reuter, A. Nelson, E. Guevara (2008) Data from "Hole-filled SRTM for the globe Version 4." Available at: http://srtm.csi.cgiar.org.

Kaiser, N. K., A. M. McKenna, J. J. Savory, C. L. Hendrickson, and A. G. Marshall (2013), Tailored ion radius distribution for increased dynamic range in FT-ICR mass analysis of complex mixtures, *Anal. Chem.*, *85*, 265–272.

Kaiser, N. K., J. J. Savory, A. M. McKenna, J. P. Quinn, C. L. Hendrickson, and A. G. Marshall (2011a), Electrically compensated Fourier transform ion cyclotron resonance cell for complex mixture mass analysis, *Anal. Chem.*, *83*, 6907–6910.

Kaiser, N. K., J. P. Quinn, G. T. Blakney, C. L. Hendrickson, and A. G. Marshall (2011b), A novel 9.4 Tesla FTICR mass spectrometer with improved sensitivity, mass resolution, and mass range, *J. Am. Soc. Mass Spectrom.*, *22*, 1343–1351.

Koch, B. P., and T. Dittmar (2006), From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter, *Rapid Commun. Mass Spectrom.*, *20*, 926–932.

Koch, B. P., T. Dittmar, M. Witt, and G. Kattner (2007), Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter, *Anal. Chem.*, *79*, 1758–1763.

Lutz, A. F., W. W. Immerzeel, A. B. Shrestha, and M. F. P. Bierkens (2014), Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation, *Nature Clim. Change*, *4*, 587–592.

Ludwig, W., J.-L. Probst, and S. Kempe (1996), Predicting the oceanic input of organic carbon by continental erosion, *Glob. Biogeochem. Cy.*, *10*, 23–41.

Mann, P. J., A. Davydova, N. Zimov, R. G. Spencer, S. Davydov, E. Bulygina, S. Zimov, and R. M. Holmes (2012), Controls on the composition and lability of dissolved organic matter in Siberia's Kolyma River basin, *J. Geophys. Res.*, *117*, G01028.

Maurya, A. S., M. Shah, R. D. Deshpande, R. M. Bhardwaj, A. Prasad, and S. K. Gupta (2010), Hydrograph separation and precipitation source identification using stable water isotopes and conductivity: River Ganga at Himalayan foothills, *Hydrol. Process.*, *25*, 1521–1530.

McGlynn, B. L., and J. J. McDonnell (2003), Role of discrete landscape units in controlling catchment dissolved organic carbon dynamics, *Water Res. Res.*, *39*, 1090.

Milliman, J. D., and J. Syvitski (1992), Geomorphic/tectonic control of sediment discharge to the ocean: The importance of small mountainous rivers, *J. Geol.*, *100*, 525–544.

Moore, I. D., P. E. Gessler, G. A. Nielsen, and G. A. Peterson (1993), Soil attribute prediction using terrain analysis, *Soil Sci. Soc. Am. J.*, *57*, 443–452.

O'Donnell, J. A., G. R. Aiken, K. D. Butler, F. Guillemette, D. C. Podgorski, and R. G. Spencer (2016), DOM composition and transformation in boreal forest soils: The effects of temperature and organic-horizon decomposition state, *J. Geophys. Res. Biogeosci.*, *121*, 2727–2744.

Raeke, J., O. J. Lechtenfeld, M. Wagner, P. Herzsprung, and T. Reemtsma (2016), Selectivity of solid phase extraction of freshwater dissolved organic matter and its effect on ultrahigh resolution mass spectra, *Environmental Science: Processes & Impacts*, *18*, 918–927.

Rivas-Ubach, A., Y. Liu, T. S. Bianchi, N. Tolić, C. Jansson, and L. Paša-Tolić (2018), Moving beyond the van Krevelen diagram: a new stoichiometric approach for compound classification in organisms, *Anal. Chem.*, *90*, 6152–6160.

RGI Consortium (2015), Randolph Glacier Inventory – A dataset of global glacier outlines: Version 5.0: Technical report, Global Land Ice Measurements from Space, doi:https://doi.org/10.7265/N5-RGI-50.

Santl-Temkiv, T., K. Finster, T. Dittmar, B. M. Hansen, R. Thyrhaug, N. W. Nielsen, and U. G. Karlson (2013), Hailstones: A window into the microbial and chemical inventory of a storm cloud, *PLoS ONE*, *8*, e53550–7.

Sharp, M., J. Parkes, B. Cragg, I. J. Fairchild, H. Lamb, and M. Tranter (1999), Widespread bacterial populations at glacier beds and their relationship to rock weathering and carbon cycling, *Geology*, *27*, 107–110.

Singer, G. A., C. Fasching, L. Wilhelm, J. Niggemann, P. Steier, T. Dittmar, and T. J. Battin (2012), Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate, *Nat Geosci*, *5*, 710–714.

Spencer, R. G., A. Vermilyea, J. Fellman, P. Raymond, A. Stubbins, D. T. Scott, and E. Hood

(2014a), Seasonal variability of organic matter composition in an Alaskan glacier outflow: insights into glacier carbon sources, *Environ. Res. Lett.*, *9*, 055005.

Spencer, R. G., P. J. Hernes, R. Ruf, A. Baker, R. Y. Dyda, A. Stubbins, and J. Six (2010), Temporal controls on dissolved organic matter and lignin biogeochemistry in a pristine tropical river, Democratic Republic of Congo, *J. Geophys. Res.*, *115*, G03013.

Spencer, R. G., P. J. Mann, T. Dittmar, T. I. Eglinton, C. McIntyre, R. M. Holmes, N. Zimov, and A. Stubbins (2015), Detecting the signature of permafrost thaw in Arctic rivers, *Geophys. Res. Lett.*, *42*, 2830–2835.

Spencer, R. G., W. Guo, P. Raymond, T. Dittmar, E. Hood, J. Fellman, and A. Stubbins (2014b), Source and biolability of ancient dissolved organic matter in glacier and lake ecosystems on the Tibetan Plateau, *Geochim. Cosmochim. Ac.*, *142*, 64–74.

Stenson, A. C., A. G. Marshall, and W. T. Cooper (2003), Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra, *Anal. Chem.*, *75*, 1275–1284.

Stubbins, A. et al. (2012), Anthropogenic aerosols as a source of ancient dissolved organic matter in glaciers, *Nat Geosci*, *5*, 198–201.

Wilhelm, L., G. A. Singer, C. Fasching, T. J. Battin, and K. Besemer (2013), Microbial biodiversity in glacier-fed streams, *The ISME Journal*, *7*, 1651–1660.

FIGURES & CAPTIONS

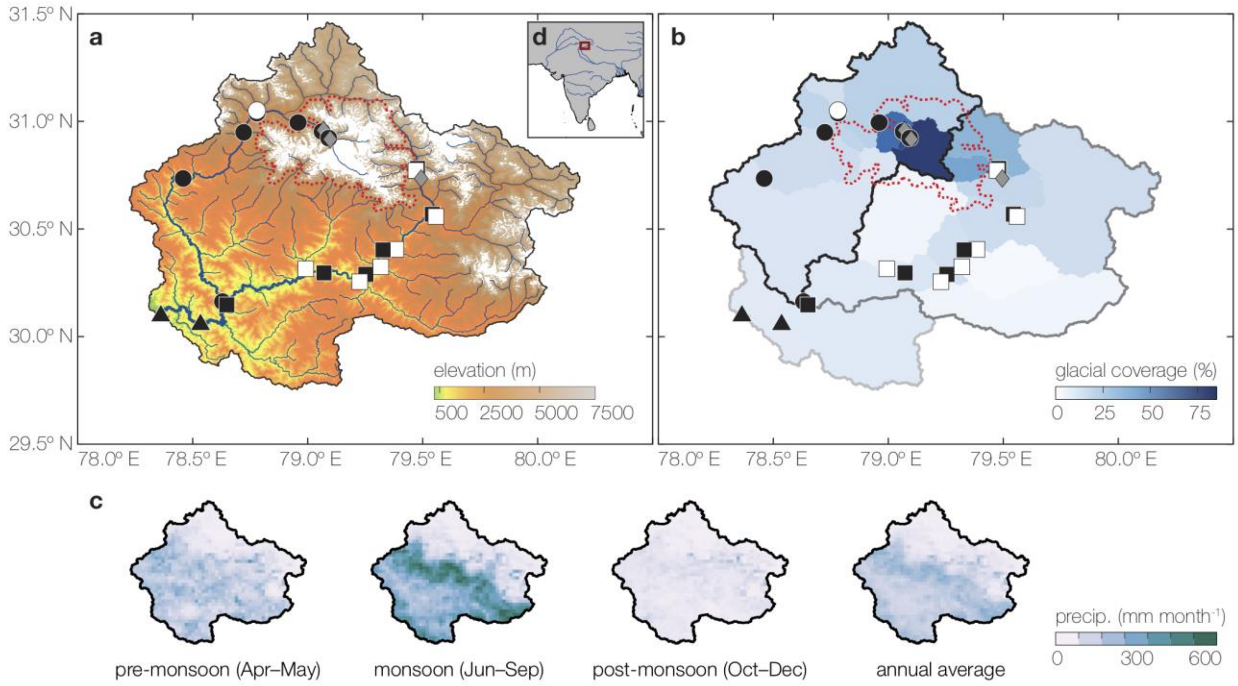

FIG. 1: Upper Ganges Basin map. **(a)** Elevation (colored pixels), glacier extent (white pixels), and river network (blue lines). **(b)** Areal percent of catchment upstream of each sampling location that is covered by glaciers. Named sub-catchments are identified by outline color in panel **(b)**: Bhagirathi (black), Alaknanda (dark gray), Ganges downstream of confluence (light gray). The Gangotri glacier group is outlined with a dotted red line in both panels. For both panels, river sampling locations are separated into main-stem (black) or tributary (white) for the Bhagirathi (circles), Alaknanda (squares), and Ganges (triangles) Rivers. Glacier and snowpack sampling locations are additionally shown as gray diamonds. **(c)** Seasonally (left three panels) and annually (right-most panel) averaged precipitation amounts throughout the basin based on 10 years of satellite measurements [*Bookhagen and Burbank*, 2010]. **(d)** Inset showing the study region location (red square) within South Asia.

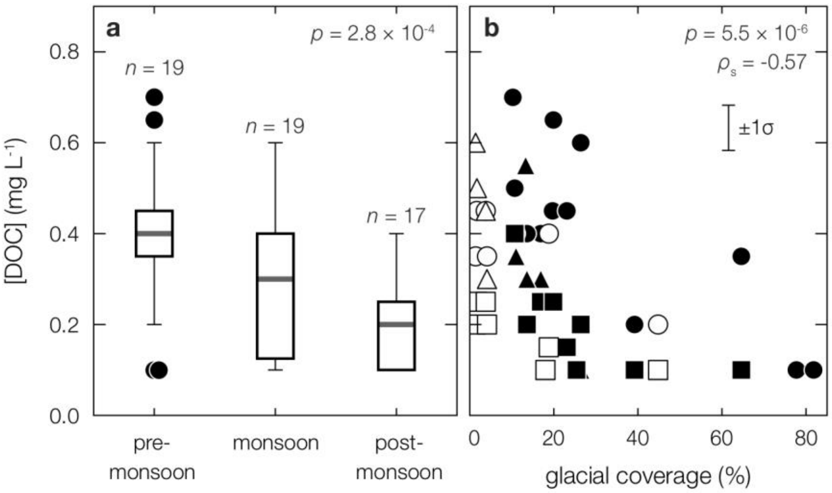

FIG. 2: Spatiotemporal trends in DOC concentration. **(a)** Box plots of [DOC] for all river samples separated by season, showing the median (thick gray line), inter quartile range (box), 95 % confidence interval (whiskers), and outliers (black circles). **(b)** [DOC] as a function of upstream glacial coverage. Markers are separated into pre-monsoon (circles), monsoon (triangles), and post-monsoon (squares) for main-stem (black) and tributary (white) samples. Analytical [DOC] uncertainty is additionally shown as ± 1σ in panel **(b)**.

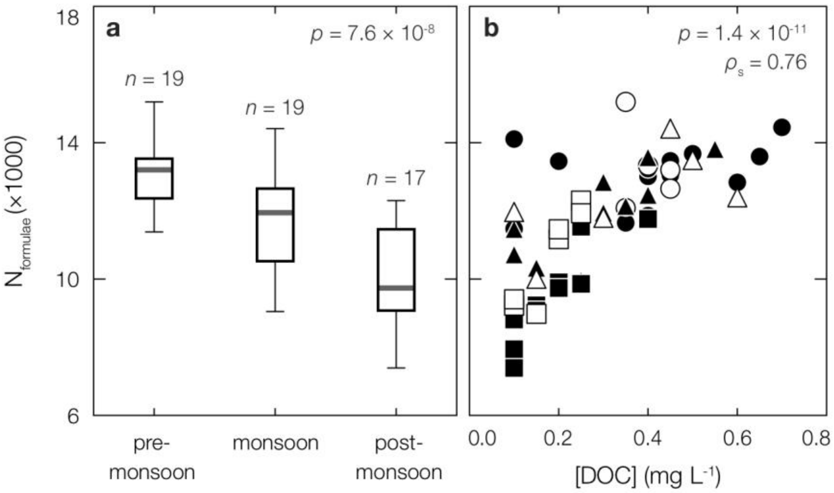

FIG. 3: Spatiotemporal trends in DOM chemical diversity. **(a)** Box plots showing the number of detected formulae for all river samples separated by season. Box plots represent the median (thick gray line), inter quartile range (box), and 95 % confidence interval (whiskers) for each population. **(b)** Scatter plot showing the number of detected formulae for all river samples as a function of [DOC]. Markers are separated into pre-monsoon (circles), monsoon (triangles), and post-monsoon (squares) for main-stem (black) and tributary (white) samples.

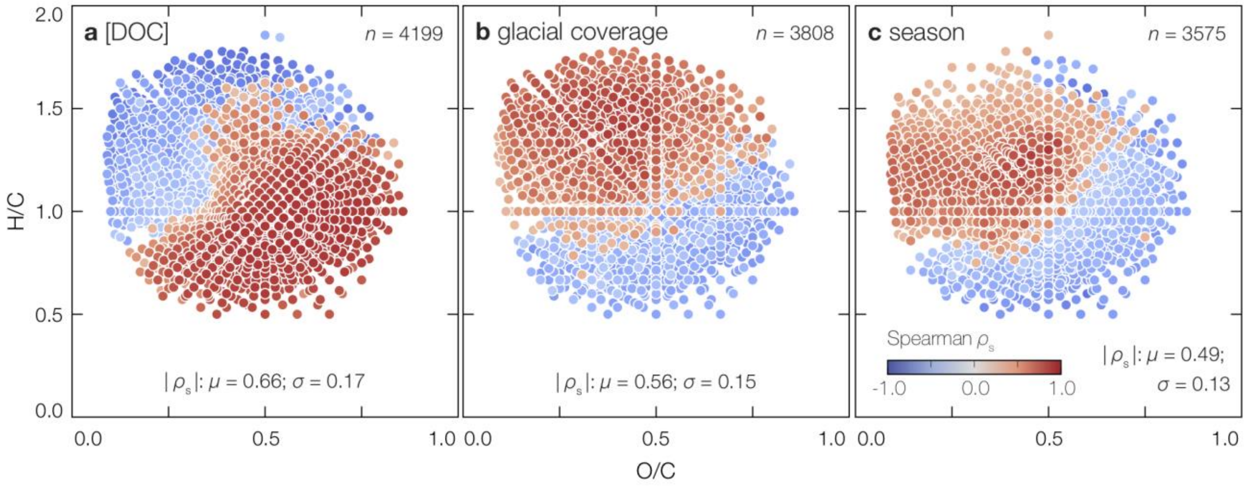

Fig. 4: DOM molecular composition as a function of **(a)** [DOC], **(b)** glacier coverage, and **(c)** season, plotted in van Krevelen space. Colors represent the correlation coefficient (ρ_s) between the relative intensity of each molecular formula as determined by FT-ICR MS and a given environmental variable [color bar in **(c)** applies to all panels]. Red formulae are more abundant in samples described by higher values of a given environmental variable whereas blue formulae are more abundant in samples described by lower values of a given environmental variable. For panel **(c)**, season has been replaced by a dummy variable (pre-ISM = 1, ISM = 2, post-ISM = 3). Only formulae that are detected in all river samples and are significantly correlated with a given environmental variable ($p \leq 0.05$) are shown. $|\rho_s|$ refers to the mean (μ) and standard deviation (σ) of the absolute value of ρ_s for all retained formulae in a given panel.

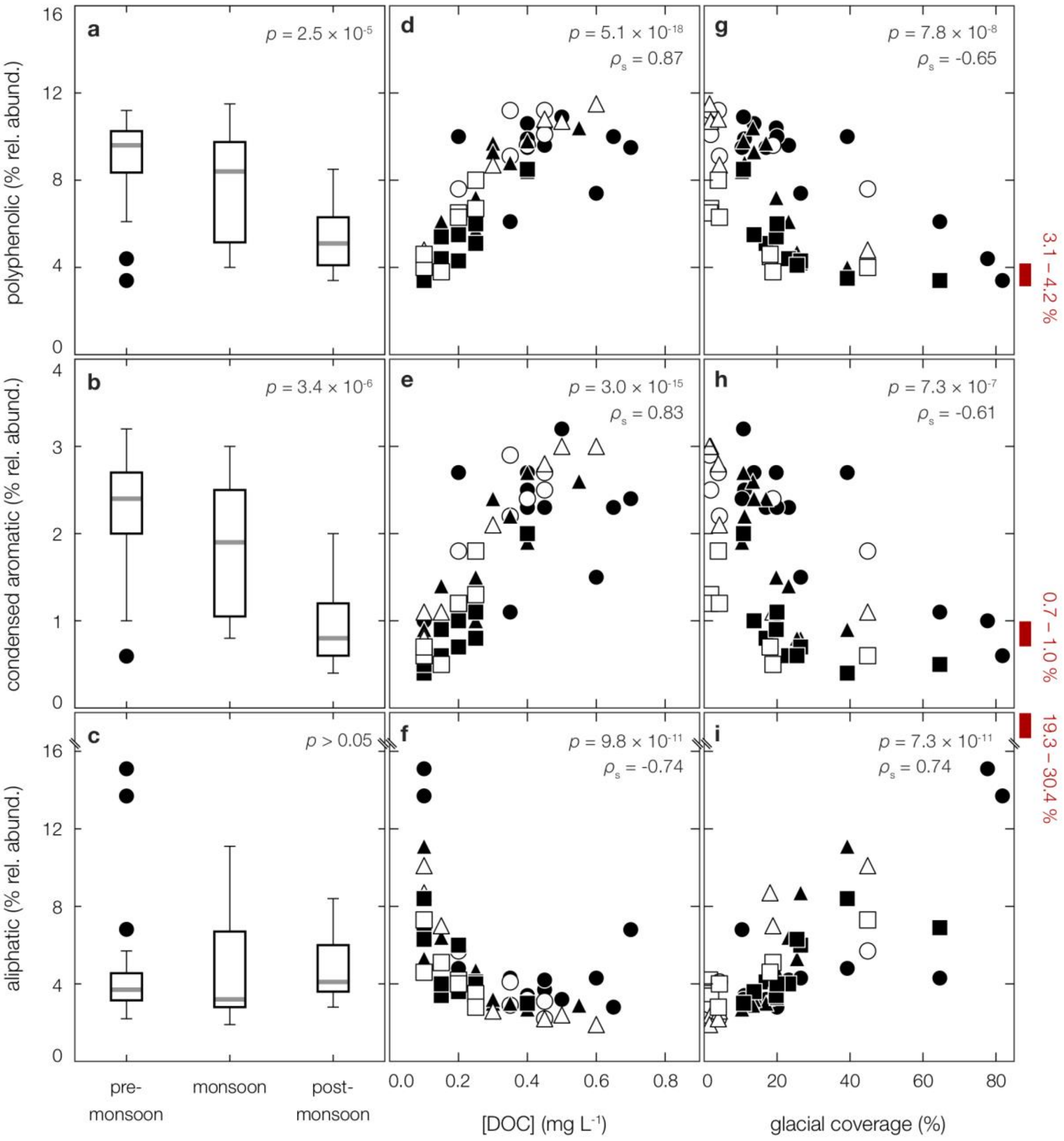

FIG. 5: Spatiotemporal trends in DOM composition. **(a-c)** Box plots showing the relative abundance of polyphenolic, condensed aromatic, and aliphatic formulae for all river samples separated by season. Box plots represent the median (thick gray line), inter quartile range (box), 95 % confidence interval (whiskers), and outliers (black circles) for each population. Scatter plots showing the relative abundance of each compound class for all river samples as a function of **(d-f)** [DOC] and **(g-i)** glacier coverage. Markers are separated into pre-monsoon (circles), monsoon (triangles), and post-monsoon (squares) for main-stem (black) and tributary (white) samples. The range of glacier/snowpack relative abundances for each compound class are additionally shown in panels **(g-h)** as red bars. Note broken *y* axes in panels **(c)**, **(f)**, and **(i)**.

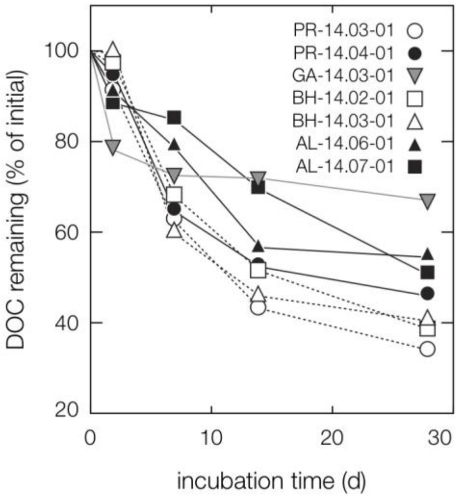

FIG. 6: [DOC] as a function of time for bioavailability incubations. Sample IDs correspond to those presented in Table S1. Error bars for triplicate measurements (± 1σ) are smaller than marker points (typically ± 1 – 2 %).

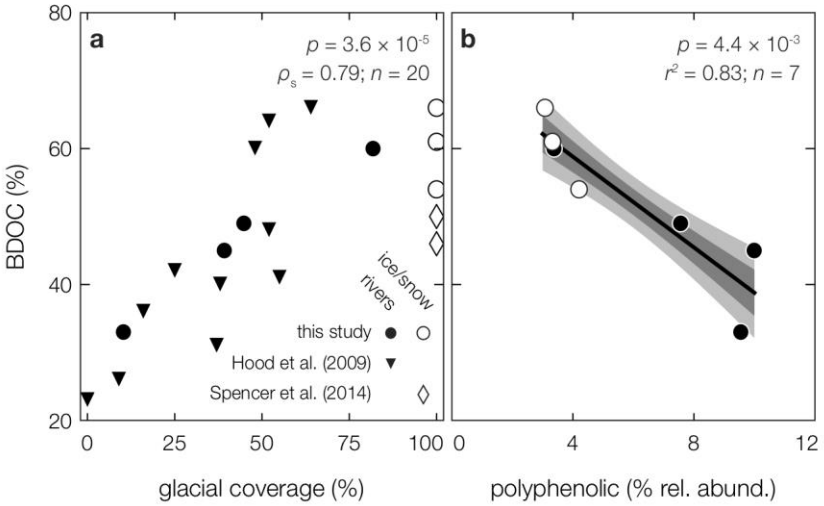

FIG. 7: Environmental and compositional controls on DOC bioavailability. Percent bioavailable DOC (% BDOC) during 28-day incubations as a function of **(a)** glacial coverage and **(b)** relative FT-ICR MS abundance of polyphenolic formulae. Markers are separated into river (black) and snow/ice (white) samples as reported in this study (circles), *Hood et al.* [2009] (Gulf of Alaska; triangles), and *Spencer et al.* [2014b] (Tibetan Plateau; diamonds). For panel **(b)**, solid black line is the ordinary least squares (OLS) regression line, dark gray envelope is the ± 1σ uncertainty, and light gray envelope is the 95 % confidence interval.

SUPPORTING INFORMATION TABLE CAPTIONS

Table S1: All DOC, geomorphic, and geospatial results for all samples in this study.

Table S2: Chemical formulae and intra-sample relative abundances for all detected compounds.

Table S3: Discharge and DOC flux estimates for rivers draining the Upper Ganges Basin.

SUPPORTING INFORMATION FIGURES & CAPTIONS

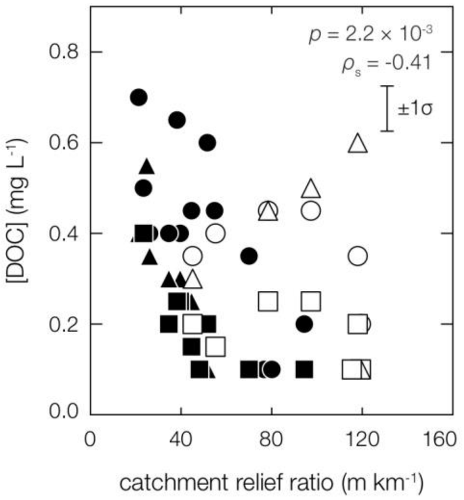

FIG. S1: [DOC] as a function of catchment relief ratio. Markers are separated into pre-monsoon (circles), monsoon (triangles), and post-monsoon (squares) for main-stem (black) and tributary (white) samples. Analytical [DOC] uncertainty is additionally shown as ± 1σ.

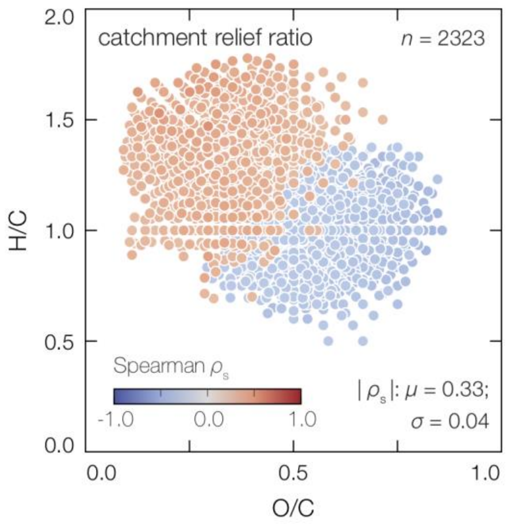

FIG. S2: DOM molecular composition as a function of catchment relief ratio, plotted in van Krevelen space. Colors represent the correlation coefficient (ρ_s) between the relative intensity of each molecular formula as determined by FT-ICR MS and relief ratio. Red formulae are more abundant in samples described by higher relief ratio whereas blue formulae are more abundant in samples described by lower relief ratio. Only formulae that are detected in all river samples and are significantly correlated with relief ratio ($p \leq 0.05$) are shown. $|\rho_s|$ refers to the mean (μ) and standard deviation (σ) of the absolute value of ρ_s for all retained formulae in a given panel.

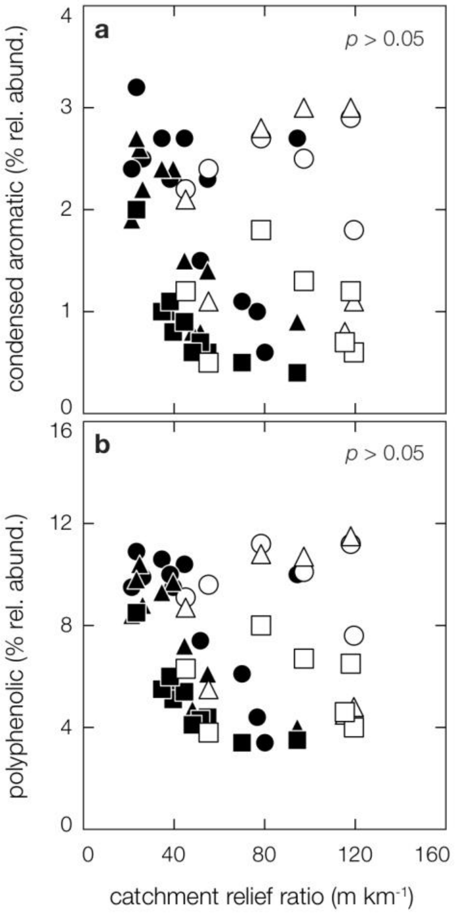

FIG. S3: Scatter plots showing the relative abundance of **(a)** polyphenolic and **(b)** condensed aromatic formulae as a function of catchment relief ratio. Markers are separated into pre-monsoon (circles), monsoon (triangles), and post-monsoon (squares) for main-stem (black) and tributary (white) samples.

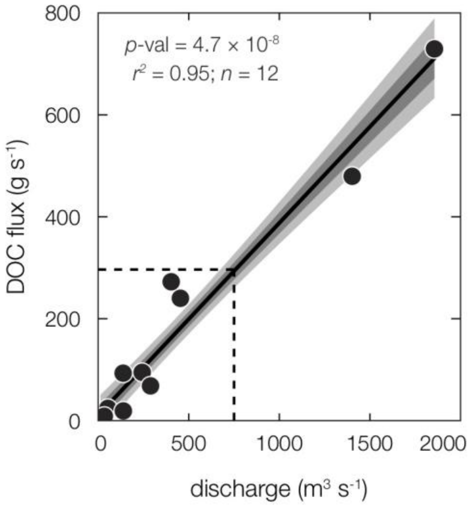

FIG. S4: Discharge vs. DOC flux rating curve for the Upper Ganges Basin using our [DOC] data and discharge data from nearby gauging stations from the years 2002 – 2004 [*Chakrapani and Saini*, 2009]. Dark gray envelope is the OLS regression ± 1σ uncertainty and light gray envelope is the 95 % confidence interval. Dashed line is the annual average discharge at our most downstream sampling location (~ 750 m^3 s^{-1}). Because discharge data are sparse and were collected 10 years prior to our DOC sample collection, resulting DOC fluxes contain large, unknown uncertainty and should only be interpreted within an order of magnitude.