Global Change Biology

Global Change Biology (2016) 22, 613–626, doi: 10.1111/gcb.13094

Human activities cause distinct dissolved organic matter composition across freshwater ecosystems

CLAYTON J. WILLIAMS[1], PAUL C. FROST[1], ANA M. MORALES-WILLIAMS[2*], JAMES H. LARSON[3], WILLIAM B. RICHARDSON[3], AISHA S. CHIANDET[4] and MARGUERITE A. XENOPOULOS[1]

[1]*Department of Biology, Trent University, Peterborough, ON, Canada,* [2]*Environmental and Life Sciences Graduate Program, Trent University, Peterborough, ON, Canada,* [3]*U.S. Geological Survey, Upper Midwest Environmental Sciences Center, La Crosse, WI, USA,* [4]*Severn Sound Environmental Association, Midland, ON, Canada*

Abstract

Dissolved organic matter (DOM) composition in freshwater ecosystems is influenced by the interactions among physical, chemical, and biological processes that are controlled, at one level, by watershed landscape, hydrology, and their connections. Against this environmental template, humans may strongly influence DOM composition. Yet, we lack a comprehensive understanding of DOM composition variation across freshwater ecosystems differentially affected by human activity. Using optical properties, we described DOM variation across five ecosystem groups of the Laurentian Great Lakes region: large lakes, Kawartha Lakes, Experimental Lakes Area, urban stormwater ponds, and rivers (n = 184 sites). We determined how between ecosystem variation in DOM composition related to watershed size, land use and cover, water quality measures (conductivity, dissolved organic carbon (DOC), nutrient concentration, chlorophyll *a*), and human population density. The five freshwater ecosystem groups had distinctive DOM composition from each other. These significant differences were not explained completely through differences in watershed size nor spatial autocorrelation. Instead, multivariate partial least squares regression showed that DOM composition was related to differences in human impact across freshwater ecosystems. In particular, urban/developed watersheds with higher human population densities had a unique DOM composition with a clear anthropogenic influence that was distinct from DOM composition in natural land cover and/or agricultural watersheds. This nonagricultural, human developed impact on aquatic DOM was most evident through increased levels of a microbial, humic-like parallel factor analysis component (C6). Lotic and lentic ecosystems with low human population densities had DOM compositions more typical of clear water to humic-rich freshwater ecosystems but C6 was only present at trace to background levels. Consequently, humans are strongly altering the quality of DOM in waters nearby or flowing through highly populated areas, which may alter carbon cycles in anthropogenically disturbed ecosystems at broad scales.

Keywords: anthropogenic, carbon cycling, cultural eutrophication, fluorescence spectroscopy, land use, parallel factor analysis modeling, urban, UV-visible absorbance

Received 30 March 2015; revised version received 31 August 2015 and accepted 4 September 2015

Introduction

Humans have transformed global water resources and the state of aquatic ecosystems (Carpenter *et al.*, 2011). A variety of anthropogenic activities such as altering hydrologic flows, fertilizer application, soil destabilization, urbanization, and agriculture have led to declines in the habitat quality, species composition, water clarity, and biogeochemical function of freshwater ecosystems. One important human impact is the cultural eutrophication (i.e., pollution and nutrient enrichments due to human activities) of freshwater ecosystems, which leads to excess levels of nutrients, high phytoplankton biomass, structural losses to aquatic habitats (e.g., reduced macrophyte biomass), low oxygen availability and at times toxic water conditions (Rabotyagov *et al.*, 2014). This traditional view of human impacts on aquatic ecosystems, however, largely omits dissolved organic matter (DOM) cycles, a fundamental aspect of aquatic ecosystem function and global carbon cycles.

DOM is an abundant and mobile part of the aquatic carbon pool that includes molecules having a wide range of biogeochemical reactivities in aquatic

*Present address: Department of Environmental Science and Biology, The College at Brockport, State University of New York, Brockport, NY, USA

Correspondence: Present address: Clayton Williams, Department of Environmental Science and Biology, The College at Brockport, State University of New York, Brockport, NY, USA, tel. + 1 585 395 5961, fax +1 585 395 5969,
e-mail: cjwillia@brockport.edu

ecosystems (Williamson *et al.*, 1999; Benner, 2003; Jaffé *et al.*, 2008). DOM serves as a fundamental link between terrestrial, freshwater, and oceanic carbon cycles given its general movement from upstream watersheds into downstream marine ecosystems through rivers and lakes (Opsahl & Benner, 1997; Cole *et al.*, 2007; Battin *et al.*, 2008; Tranvik *et al.*, 2009). Hence, aquatic DOM and mechanisms regulating DOM play an important role in global carbon cycles (Battin *et al.*, 2008; Larsen *et al.*, 2011; Stanley *et al.*, 2012). Global change and anthropogenic activities can influence the quality and quantity of DOM in aquatic ecosystems (Wilson & Xenopoulos, 2009; Larsen *et al.*, 2011; Yamashita *et al.*, 2011a; Stanley *et al.*, 2012). As agricultural systems are intensified and/or expanded and urbanization trends continue in response to human population growth, anthropogenic influences on DOM composition will likely become more widespread. Such human effects on DOM quality could have strong impacts on carbon cycles and need to be better understood.

The biogeochemistry of aquatic ecosystems varies across different land use and cover patterns as well as under different levels of anthropogenic ecosystem alteration and activity (Paul & Meyer, 2001; Grimm *et al.*, 2008; Yamashita *et al.*, 2011a; Butman *et al.*, 2015). Terrestrial watershed properties such as the presence of wetlands, agricultural intensity, and water flow path have been demonstrated to influence aquatic DOM quantity (i.e., dissolved organic carbon (DOC) concentration) and quality (McDowell & Likens, 1988; Dillon & Molot, 1997; Xenopoulos *et al.*, 2003; Frost *et al.*, 2006; Wilson & Xenopoulos, 2008; Larson *et al.*, 2014). Across temperate lakes, relationships between water color and phosphorus concentration suggest linked terrestrial supply of humic matter and nutrients (Dillon & Molot, 1997; Webster *et al.*, 2008). Anthropogenic disturbances alter land-water connections at local to regional scales and modify essential ecosystem functions from what is observed in more natural aquatic ecosystems (Grimm *et al.*, 2008), which can be associated with relatively greater proportions of protein-like and/or microbial-humic DOM within ecosystems (Williams *et al.*, 2010; Yamashita *et al.*, 2010; Petrone *et al.*, 2011; Hosen *et al.*, 2014). To determine how global carbon cycles will respond to anthropogenic pressures, the quality of DOM and how it relates to land use change, nutrient levels, and more generally human population across freshwater ecosystems must be understood in the context of these human altered ecosystems, which are increasing in area and intensity with global populations and water demand.

We compiled land use, water quality, and DOM optical chemical data from a variety of freshwater ecosystems within the Laurentian Great Lakes region. A multivariate approach was used to examine how DOM composition varied among ecosystem type and co-varied with DOC, land use and cover, water quality, and human population density. By including this diversity of aquatic ecosystems (lakes, rivers, and ponds) in contrasting landscapes, we are able to fully characterize the extent of variability in DOM composition in the Great Lakes region. We hypothesized that, across ecosystems, human activities alter DOM composition by increasing production of internally derived DOM and the processing of terrestrial DOM inputs. If so, this should be evident by a shift in DOM from allochthonous-like (e.g., higher plant derived, large sized, and aromatic) toward autochthonous-like (e.g., protein-like, small sized, and aliphatic) characteristics with higher levels of nutrients, human land use, and human population density in a watershed.

Materials and methods

Freshwater ecosystems and water collection

Surface water quality (conductivity, nitrite+nitrate [NO_{2+3}^-), total phosphorus (TP), and chlorophyll a (CHL)], DOC, and optical DOM data were compiled from 184 sites in the Laurentian Great Lakes region. The study sites were sampled as part of several independent projects at different intervals between 2008 and 2012. The freshwater ecosystems included the five Great Lakes and their major tributaries (USA), Lake Simcoe and its tributaries (Canada), the Kawartha Lakes (Canada), several lakes at the Experimental Lakes Area (ELA; Canada), stormwater urban ponds in several municipalities of southern Ontario (Canada), and southern Ontario streams and rivers (all lotic ecosystems are henceforth designated as rivers; Fig. 1). Specific spatial and/or temporal observations within most of these projects have been described elsewhere (Williams *et al.*, 2010, 2012, 2013; McEnroe *et al.*, 2013; Larson *et al.*, 2014). Sites were grouped into freshwater ecosystems as follows: 1) large lakes, 2) Kawartha Lakes, 3) ELA lakes, 4) urban ponds (n = 21), and 5) rivers (n = 92; Fig. 1). Ecosystems were grouped from all available sites into lentic and lotic sites. Lentic sites were then subdivided into groups based on location, size, and/or origin (man-made or natural). For large lakes (n = 29), 19 sites were sampled once and 3 sites were sampled twice during the summer of 2011 across 5 Great Lakes (Lakes Erie, Huron, Michigan, Ontario, and Superior). In addition, Lake Simcoe was sampled at 7 sites 3 times during the open water season of 2010. For the Kawartha Lakes (n = 34), 11 lakes were sampled once at 34 sites during August 2009. Twelve of these sites in 9 lakes were sampled 5 more times from March to August 2010. For the ELA (n = 8), eight lakes (114, 221, 222, 224, 227, 239, 249, and 302) were sampled five times each between May to September 2012. For urban ponds (n = 21), 17 ponds were sampled once during August 2010 and 4 ponds were sampled 5 times throughout 2010. For rivers, 40 sites were sampled once and 52 sites were sampled at least twice between September 2008 and November 2011.

Overall, 43% of sites were sampled more than once and 57% of sites were sampled once. Of the singletons, 89% of the sites were sampled in summer months between June and August. These late summer samples occurred in the temperate seasonal cycle when phytoplankton levels are expected to be higher and soil-water infiltration rates and river discharge are expected to be lower due to warmer temperatures and increased rates of terrestrial evapotranspiration. Hence, lotic DOM samples should be fairly representative of base- and low flow conditions. Lentic DOM samples should be representative of warm conditions with higher rates of *in situ* processing. Here, we use the single point measurement (when sites were only sampled once) or the mean (when sites were sampled more than once) for each site to explore broad-scale patterns in DOM characteristics among freshwater ecosystems and determine across-ecosystem DOM relationships with land use, water quality, and human population density.

DOM composition was determined using optical chemistry at all sites. Water quality variables were not uniformly collected among projects. As such, the subset of water quality indicators that were measured across projects was used in the present study. For large lakes, subsurface euphotic-zone water samples were collected at open-water locations outside of visibly or chemically distinct plume introduced from nearby tributaries. For Kawartha Lakes sites water was sampled one meter below the water surface. ELA lakes were sampled from the center of the lake using a Van Dorn sampler to collect water from the middle of the epilimnion. For urban ponds, which are built to hold and process stormwater from upstream residential, business, and industrial catchments, water samples were collected at 0.2 m depth below the water surface. Surface samples in these ponds can be influenced by surface sheens of organic contaminants. During the time of our sampling, no visible evidence of organic contaminants was observed in collected water and the water and DOM chemistry of near bottom samples was generally indistinguishable from surface samples, suggesting contaminates if present were not detectible by our methods (Williams *et al.*, 2013). Finally, for independent sampling sites in rivers, water was collected just below the water's surface of headwater streams to 7th Strahler order rivers.

Specific conductivity was measured *in situ* at the specified sampling depth using YSI probes (YSI 600 XLM or YSI 30/10 FT, depending on project). Due to differences among projects, water sample storage methods differed among sites. Water samples for NO_{2+3}^-, DOM, and DOC were filtered on site through prerinsed 0.2 μm polycarbonate membrane filters for large lakes, Kawartha Lakes, and USA rivers. Canadian rivers, ELA, and urban pond water samples were stored in acid washed, sample rinsed polycarbonate or polyethylene bottles kept on ice and in the dark until filtering in the laboratory through prerinsed 0.2 μm polycarbonate membrane filters within 36 h of collection. NO_{2+3}^- samples were acidified and stored at 4 °C until analysis. DOC and DOM samples were stored in precombusted, sample rinsed amber glass vials or 24 h acid, base, and DI water soaked dark, polycarbonate bottles. Whole water samples for TP were stored frozen or acidified at 4 °C in polyethylene bottles until analysis. In addition, whole water samples were filtered in the lab through Whatman GF/F (0.7 μm nominal pore size) or Pall A/E (1 μm pore size) filters, and filters were frozen for CHL analysis. River sites

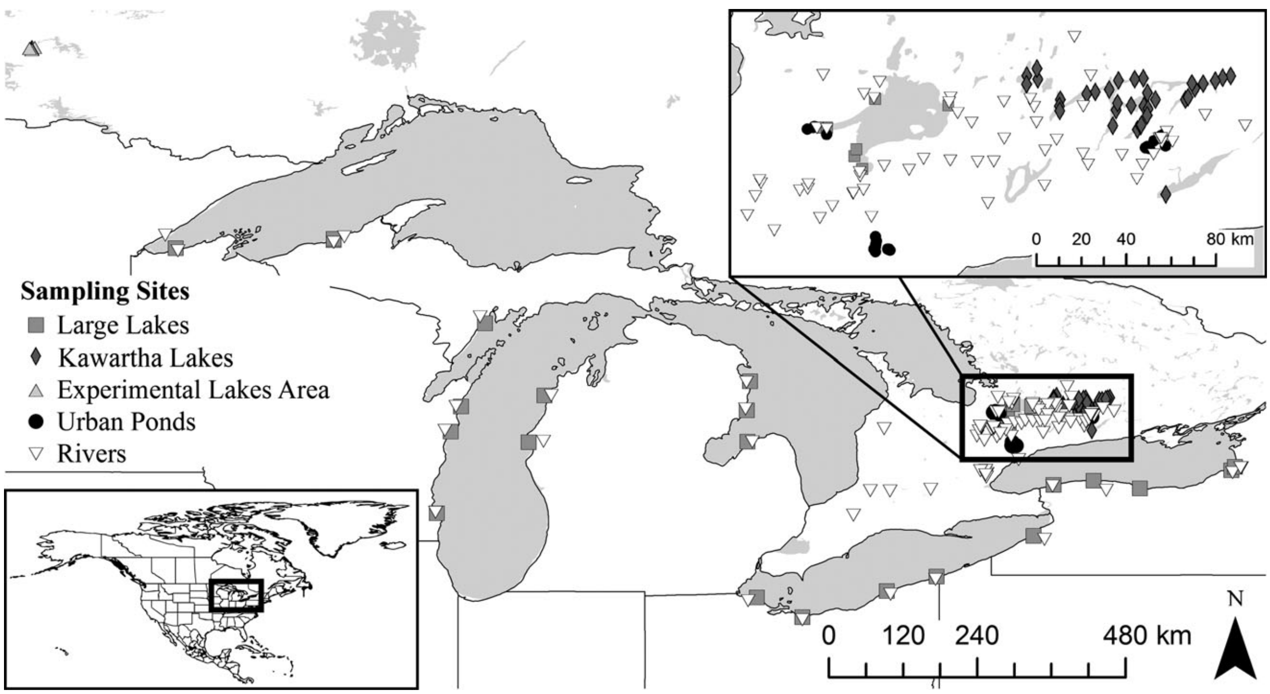

Fig. 1 Sampling site locations across US and CA. Inset map shows a close up of southern Ontario to reveal Lake Simcoe, Kawartha Lakes, and the distribution of river sites. Sampling sites are grouped into five freshwater ecosystems.

were not analyzed consistently for CHL and, consequently, CHL was not presented for rivers.

Water quality and dissolved organic matter analysis

An OI Aurora TOC Analytical Analyzer (Xylem Inc., College Station, TX, USA) with external TN unit was used to measure DOC concentration (mg C L^{-1}) as NPOC following sample acidification and combustion. TP concentration (μg P L^{-1}) was determined on whole water samples following persulfate digestion using the standard colorimetric assay (APHA, 2006). NO^{-}_{2+3} concentration (mg N L^{-1}) was measured colorimetrically following the standard cadmium reduction method (APHA, 2006) or through the second derivative method (Crumpton *et al.*, 1992). For some river, urban pond and Kawartha Lake samples, NO^{-}_{2+3} was not measured directly but estimated using a system specific relationship with total dissolved nitrogen (TDN, measured using the OI Aurora system or second derivative method). For rivers and urban ponds, the linear model: $NO^{-}_{2+3} = 1.0041*TDN - 0.3541$ ($R^2 = 0.97$, $n = 79$) was used to estimate NO^{-}_{2+3} concentration. For Kawartha Lakes, the equation: $\log_{10}$ (%NO^{-}_{2+3} of TDN) = $-1.3013*\log_{10}(TDN) - 0.6429$ ($R^2 = 0.83$, $n = 57$) was used to estimate NO^{-}_{2+3} concentration.

CHL concentration (μg L^{-1}) as an estimate of phytoplankton biomass was extracted from filters using hot or cold ethanol for Lake Simcoe, Kawartha Lakes, ELA, and urban ponds. Chlorophyll *a* extracts were measured as fluorescence, corrected for pheophytin *a* with the addition of 0.1 N HCl, and converted to concentration using standard extracts. For a subset of samples, pheophytin *a* corrected chlorophyll *a* concentration was not measured directly. Instead an internal correction curve was used ($\log(CHL_{Corrected}) = 1.023*\log(CHL_{Uncorrected}) - 0.266$; $R^2 = 0.93$, $n = 128$) to approximate the pheophytin corrected phytoplankton biomass. For the Great Lakes, chlorophyll *a* was extracted by grinding filters in 90% acetone buffered with 10% magnesium carbonate following the standard spectrophotometric method (APHA, 2006).

DOM light absorbance was measured from 800 to 230 nm using a Perkin Elmer Lambda 25 (Waltham, Massachusetts, USA) or Cary 50 spectrophotometer (Agilent Technologies, Mississauga, ON, Canada). UV-Visible excitation emission matrix (EEM) scans were conducted using the same Varian Cary Eclipse Fluorometer (Agilent Technologies). Scans were made from 600 to 270 nm emission at a 2 nm interval and 500 to 230 nm excitation at a 5 nm interval with 5 nm band widths at 0.25 nm s^{-1} with a daily Milli-Q blank scan (Murphy *et al.*, 2010; Williams *et al.*, 2010). EEMs were corrected following standard recommendations and converted to Raman Units (Cory *et al.*, 2010; Murphy *et al.*, 2010). The following DOM optical indices were calculated: (i) a spectral slope ratio used as a proxy for molecular size (S_R; Helms *et al.*, 2008), (ii) the specific UV absorbance at 254 nm used as a correlate for aromaticity (SUVA; Weishaar *et al.*, 2003), (iii) the modified fluorescence index used as an indicator of source (FI; McKnight *et al.*, 2001; Cory *et al.*, 2010), (iv) a modified humification index (HIX; Zsolnay *et al.*, 1999; Ohno, 2002), and (v) the freshness index used as an indicator of the extent of DOM degradation (β : α ratio; Parlanti *et al.*, 2000; Wilson & Xenopoulos, 2009).

EEMs were analyzed further through parallel factor analysis modeling (PARAFAC) using the DOMFluorv1_7 toolbox in Matlab 2007b and R2012b (Mathworks). EEMs from all freshwater ecosystems processed prior to fall 2010 were used to create our previously published seven component model (Williams *et al.*, 2013). EEMs analyzed after fall 2010 were fit to the existing model and residuals inspected to verify model fit (Fellman *et al.*, 2009). The seven component PARAFAC characterized the DOM pool as: C1, ubiquitous humic-like, 260/482 (Ex/Em; nm); C2, terrestrial humic-like, <250/420; C3, terrestrial humic-like, <250/440; C4, soil fulvic-like, 285/536; C5, microbial humic-like, 360/424; C6, anthropogenic, microbial humic-like, <250/386; C7, protein-like, 280/342 (Williams *et al.*, 2013 and references within). C6 was the least commonly observed PARAFAC component in the literature and, when found, tends to occur in urbanized aquatic ecosystems (Williams *et al.*, 2013; Hosen *et al.*, 2014; Parr *et al.*, 2015). C6 has a distinctive emission spectra with a rapidly increasing upslope, a clearly defined peak, and a gradually decreasing downslope that easily differentiates it from more commonly observed microbial, humic-like DOM, such as C5, with a more symmetrical and gradual sloped emission spectra. Components are presented as relative fluorescence (% total fluorescence). In addition, as a measure of specific fluorescence the sum of the PARAFAC F_{max} was divided by DOC (ΣF_{max}/DOC; Raman Units [RU] mg^{-1} C L) and used as an optical index.

Watershed delineation, land use, and human population density

Watershed size (km^2), catchment land use and cover, and human population density were determined using spatial analyst and hydrology toolboxes in ArcGIS (Esri) from digital elevation models, the southern Ontario Interim Landcover (SIL) Geospatial data set (OMNR, 2007), Ontario Provincial Landcover (PL; OMNR, 2000), the US National Land Cover Database (NLCD; Fry *et al.*, 2011) and US and CA census data. Watershed areas for large lakes and rivers were delineated for the direct drainage area to the nearshore location or upstream of the river sampling point. For ELA, urban ponds, and Kawartha Lakes, watersheds were delineated more broadly and included the hydrologic inputs around the lake for ELA and urban ponds or upstream and local hydrologic inputs for fluvial Kawartha Lakes chain. These watershed delineations were then used to calculate the relative land use and human population density associated with each sampling site. Due to lack of sewershed data, watershed size could not be calculated for 8 of 21 urban ponds.

Based on the land class definitions of SIL, PL, and NLCD, US and Canadian land use and cover data were combined into seven categories to allow comparison across ecosystems (Table S1). The land use classification used here were: developed, cultivated, rural, wooded, wetland, open water, and other. Developed lands comprise man-made vegetated areas with some construction to large metropolitan areas with high levels of impervious surfaces but do not include agricultural land uses. Cultivated lands include row crops and actively tilled land. Due to differences between US and Canada classifications, pasture land and hay fields are included in rural lands. All land use data were presented as the relative area of each land use type (% cover) within a watershed (Table S1). Human population counts were obtained for the study area from United States Census Bureau for 2010 (www.census.gov) and from Statistics Canada for the 2011 Census of Canada (http://www12.statcan.gc.ca). Human population density (persons km^{-2}) was calculated for each watershed using the population sum for all census tracts that overlapped with a watershed divided by the area of the overlapping census tracts.

Statistical analysis

All data analysis was performed in R (version 3.0.3) via RSTUDIO (version 0.98.501) using the 'vegan' and 'pls' libraries and R's built-in packages and functions (Mevik *et al.*, 2013; R Core Team, 2014; Oksanen *et al.*, 2015). R-script for statistical analyses is provided in SI. Water quality variables and population density were $\log_{10}$ transformed and land use and cover was square root transformed prior to analysis to allow linear fit and better meet the assumptions of equal variance among ecosystems. Box-whisker plots (SigmaPlot, Systat Software Inc., San Jose, CA, USA) were used to illustrate univariate patterns for each variable within and across freshwater systems (Fig. 2 and S2). Pearson's bivariate correlation was used to examine univariate relationships between variables (Table S2 and S3, Fig. S1), where the 95% confidence interval (CI; i.e., effect size) was estimated using bootstrap analysis. Prior to multivariate analysis, DOM and land use data were Z-score standardized for each measurement variable. Permutation MANOVA with resampling pairwise comparison was used to determine if DOM characteristics and land use differed by freshwater system. 10 000 permutations/resamplings were generated to estimate a significance probability and/or correlation CI.

Principle components analysis (PCA) was used to visualize statistical differences in DOM composition and land use. Partial least squares (PLS) linear models were used to examine multivariate correlations between DOM and land use and cover (LULC). The DOM and LULC scaled variable sets were the same for PLS and PCA. DOM included 13 measures of composition (S_R, $\beta : \alpha$, HIX, FI, SUVA, ΣF_{max}/DOC, and PARAFAC C1 to C7). LULC included 7 landscape features (developed lands, cultivated lands, rural lands, wooded lands, open water, wetlands, and other LULC). Components 1 and 2 were used for PCA to illustrate differences among ecosystems groups and for PLS the DOM vs. LULC model to examine covariation between these measures. PLS was used further to correlate DOM and land use multivariate data with univariate measures of watershed size, water quality, and population density. Finally, PLS was used to determine the main influences on C6 by relating C6 to normality transformed and scaled land use and cover, watershed size, and water quality. Spatial autocorrelation for DOM and land use data were assessed using simple and partial Mantel tests.

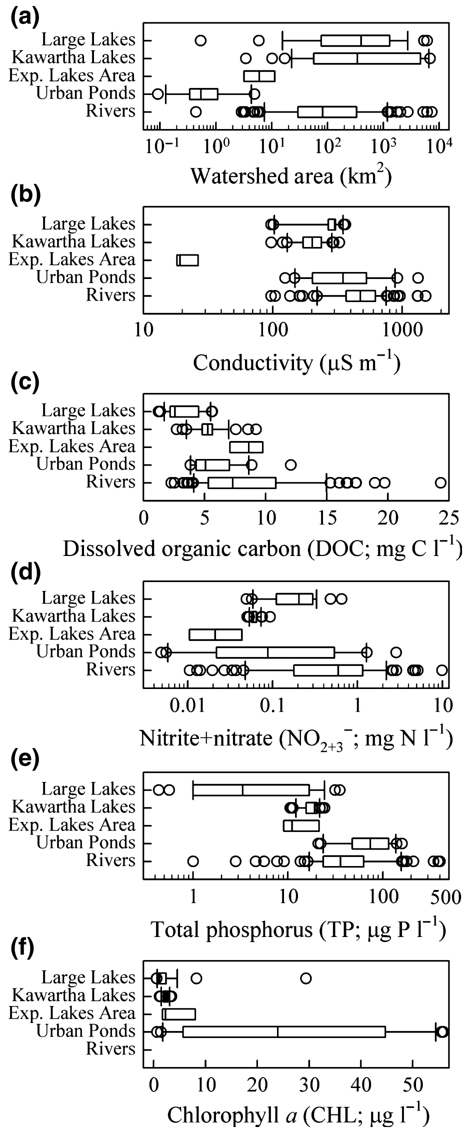

Fig. 2 Box and Whisker plot of water quality variables for each freshwater system. Boxes are 25th, 50th (median), and 75th intervals. Whiskers are 10th and 90th percentile. Dots are the remaining observations. Watershed size, Conductivity, Nitrite+Nitrate, and Total Phosphorus are set to a log scale. Chlorophyll *a* was not measured in rivers.

Results

Watershed size ranged nearly five orders of magnitude across sites from 0.1 to 7406 km^2 (Fig. 2a). Conductivity ranged from 12 to 1516 μS m^{-1} and was an order of magnitude lower at ELA than the other ecosystem types (Fig. 2b). Across sites, water quality varied from oligotrophic (TP = 0.4 μg L^{-1}) to eutrophic (TP = 401.1 μg L^{-1}) with low (1.2 mg L^{-1}) to moderate (24.3 mg L^{-1}) DOC concentrations (Fig. 2c–f). NO_{2+3}^{-}

concentration ranged from 0.005 to 9.9 mg L $^{-1}$ and was generally low across large lakes, ELA, and the Kawartha Lakes but overall more variable in urban ponds and rivers with at times high concentrations (Fig. F2c). In general, urban ponds and rivers had the widest range in water quality parameters, bracketing the ranges found in large lakes, ELA and the Kawartha Lakes. CHL concentrations suggested nonbloom conditions for all lake sites but one large lake observation (29.4 μg L $^{-1}$ near the inflow of Crane Creek, in the western basin of Lake Erie; Fig. 2f). CHL was variable among urban ponds, ranging from just detectible to bloom conditions (>50 μg L $^{-1}$). Across all ecosystems, positive correlations with weak to moderate effect sizes were observed between all pairs of water quality variables except $NO_{2+3}^{-}$and CHL, which did not correlate with each other (Table S2).

The relative distribution of land use and cover differed significantly across ecosystem type (MANOVA $F_{4,171} = 20.6$, $P = 0.0001$; Fig. 3a). ELA had the most wooded lands followed by some of the northern Kawartha Lakes sites whereas the most developed watersheds were found upstream of urban ponds. Southern Kawartha Lakes, large lakes, and river sites had a mixture of watershed land uses and cover (Fig. 3a; Table S1). Land use and cover across ecosystem type correlated with water quality variables (Fig. S3) with more human land use associated with higher conductivity, NO_{2+3}^{-}, TP, and CHL. While the connection between DOC and land use was weak ($r = 0.37$, CI = 0.23–0.50), we found higher DOC in sites with more cultivated, rural, and wetlands and less wooded lands and open water. These land use patterns across ecosystems were spatially correlated (Mantel $r = 0.32$) as observed by land use transitions from wooded and open water in the northern Canadian Shield region (ELA and northern watersheds of the Kawartha Lakes) to more cultivated and rural lands to the south (southern Ontario and northern USA sites). ELA, Kawartha Lakes, and urban ponds also had a clumped spatial distribution relative to large lakes and rivers (Fig. 1), but overall the spatial influence was weak, explaining 19% of the variation in land use across ecosystems, and sites oriented differently by land use and cover than spatial location (Fig. 1 and 3a).

DOM composition significantly differed by freshwater ecosystem (MANOVA $F_{4,179} = 66.4$, $P = 0.0001$; Fig. 3b). Pairwise comparison indicated that DOM composition was distinct for all freshwater ecosystem types ($P < 0.05$). DOM composition in large lakes, Kawartha Lakes, ELA, and rivers covered a gradient from lower molecular weight (S_R), recently produced ($\beta : \alpha$), protein-like (C7) DOM toward more aromatic (SUVA), humic-like (C1, C2, C4, C5, and HIX) DOM (Fig. 3b and S2). Urban pond DOM composition was distinguished from the other ecosystems by more recently produced ($\beta : \alpha$) and anthropogenic/microbial-like (C6) DOM signatures with relatively fluorescent DOM (i.e., higher ΣF_{max}/DOC; Fig. 3b and S2).

DOM composition across freshwater ecosystems correlated strongly with land use and cover (Fig. 4; Table S4). As with land use and cover, patterns in DOM composition were spatially correlated (Mantel $r = 0.19$). However, the relationships between DOM and land use (Mantel $r = 0.24$) remained significant when spatial location was accounted for (Partial Mantel $r = 0.12$). Along the DOM and LULC PLS component 1 comparison, DOM composition transitioned across the river to large lake and ELA gradient as land use and cover shifted from cultivated, rural, and other lands toward wooded and open water ($r = 0.50$, CI 0.40–0.61; Fig. 4a; Table S4). Visually, the overall trend was split

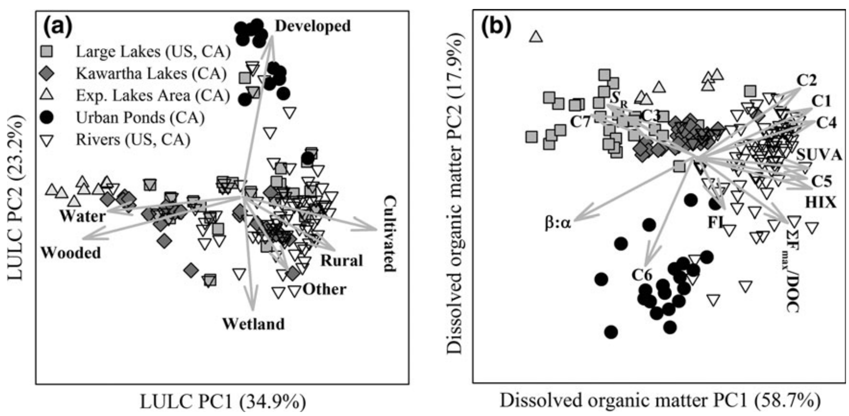

Fig. 3 Principle components analysis ordination of component 1 (PC1) and 2 (PC2) with vector loadings for land use and cover (LULC; a) and dissolved organic matter (DOM; b) variables. Points are grouped by freshwater ecosystem. All variables used in each PCA are presented as labeled vectors.

between two parallel relationships with large lakes and southern Kawartha Lakes shifted toward DOM composition with higher S_R, β : α, C3, and C7 than river, urban pond, northern Kawartha Lakes, and ELA sites with similar land use and cover. The DOM and LULC PLS component 2 comparison revealed a consistent developed to wooded and wetland gradient across sites that corresponded to a shift in DOM from higher C6 and β : α to higher C1, C2, and C4 (r = 0.77, CI = 0.70–0.84; Fig. 4b; Table S4). Hence, DOM composition across sites was sensitive to anthropogenic land use, although developed and agricultural land use produced different DOM composition.

Across freshwater ecosystems, DOM composition co-varied with water quality measures and correlated weakly with watershed size (r = 0.17, CI = 0.05–0.30; Fig. 5). These relationships were evident across ecosystem type but were less clear within each ecosystem type. The lack of within ecosystem group response is likely a result of how the data were processed and oriented within the multivariate data space, which maximized the amount of explainable variation across ecosystems not within. Clearer within ecosystems relationships between DOM and water quality were observed when these ecosystem groups were analyzed independently (Williams *et al*., 2010, 2013; Larson *et al*., 2014). DOC and TP correlations with DOM composition were more consistent across and within ecosystem gradients in freshwater DOM quality, DOM quantity, and ecosystem trophic state.

Shifts in human population density associated with higher amounts of developed land (i.e., impervious and pervious lands disturbed by humans for nonagricultural uses; Fig. 6a,b) correlated with DOM composition (Fig. 6c) and, in particular, the dominant signature of urban pond DOM (C6; Fig. 6d,e). Across all ecosystems types, these changes in land use and DOM composition were not clearly associated with changes in DOC concentration (Fig. 6f). Increases in human population density across sites altered the DOM pool away from more common humic- and protein-like features of DOM toward compositions driven by C6 and to a lesser extent β : α and ΣF_{max}/DOC.

The nonlinear nature to the relationships presented in Fig. 6, suggested additional factors influence the relative abundance of the urban/anthropogenic DOM marker (C6; Fig. 7 and S1; Table S5). Overall, C6 was best described by landscape features (developed lands, human population density, watershed size, wetlands, and wooded lands) with weaker influences from TP, conductivity, open water, and other LULC (r = 0.78, CI = 0.70–0.86; Fig. 7a; Table S5). When CHL was included in the PLS model for lentic sites only a more complex picture of C6 emerged (Fig. 7b; Table S5). For lentic sites, C6 positively related to developed lands, human population density, conductivity, TP, and CHL and negatively related to watershed size, open water, wetlands, and wooded area (r = 0.86, CI = 0.80–0.92). These changes in C6 were weakly but positively correlated to DOC (Table S5; Fig. S1). For lotic sites, C6 again positively correlated with developed lands, human population density, and conductivity and negatively correlated with wetland area (r = 0.82, CI = 0.73–0.91; Fig. 7c; Table S5). In addition for lotic

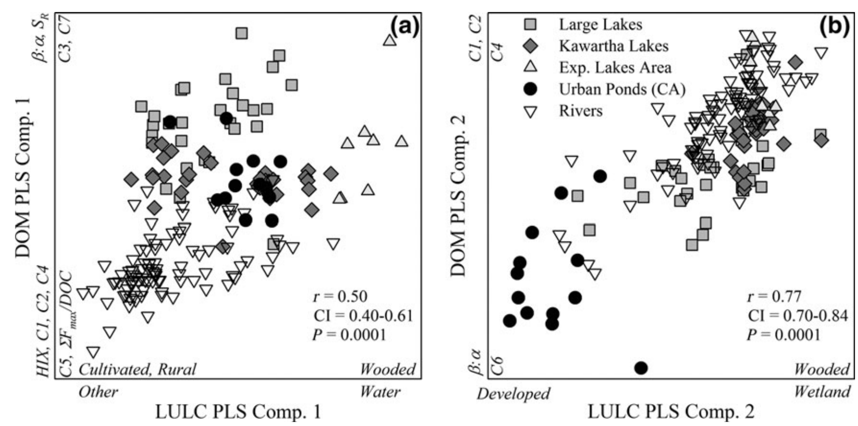

Fig. 4 Partial least squares (PLS) linear model correlation (r ± 95% effect size) between 13 dissolved organic matter (DOM) Y-variables and 7 land use and cover (LULC) X-variables (DOM ~ LULC). Model input DOM variables were S_R, β : α, HIX, FI, SUVA, ΣF_{max}/DOC, C1, C2, C3, C4, C5, C6, and C7. Model input LULC variables were developed lands, cultivated lands, rural lands, open water, wooded, wetland, and other LULC. PLS components 1 (a) and components 2 (b) are plotted. Points are grouped by freshwater ecosystem. To facilitate PLS interpretation, univariate labels are provided for variables with $r \geq 0.50$ with the PLS components. Variable labels are oriented on the figure based on the direction of their influence. PLS loading scores for each component and variable set are found in Table S4.

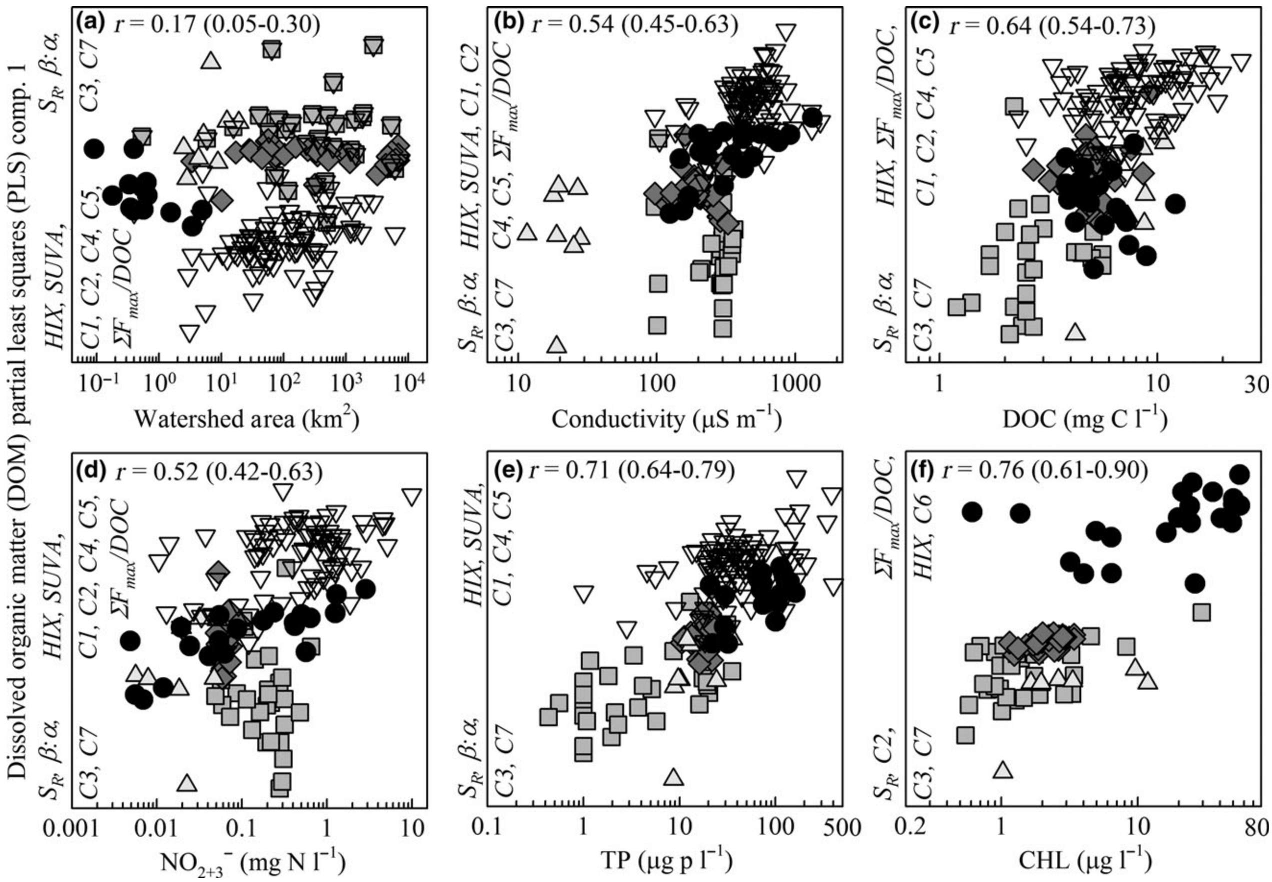

Fig. 5 Partial least squares (PLS) linear model correlation (r ± 95% effect size) between the dissolved organic matter (DOM) Y-variables and water quality X-variables. Model input DOM variables were S_R, β : α, HIX, FI, SUVA, ΣF_{max}/DOC, C1, C2, C3, C4, C5, C6, and C7. Points are grouped by freshwater ecosystem: large lakes (gray squares), Kawartha Lakes (gray diamonds), Experimental Lakes Area (light gray upturned triangles), urban ponds (black circles) and rivers (open downturned triangles). To facilitate PLS interpretation, univariate DOM labels are provided for variables with $r \geq 0.50$ with the PLS components. Variable labels are oriented on the figure based on the direction of their influence. X scale is set to log 10.

sites, C6 negatively correlated with DOC and positively correlated with TP and NO_{2+3}^{-} (Table S5; Fig. S1). Cultivated lands were never a useful predictor of C6 suggesting that this aspect of the DOM is enhanced by ecosystem alterations caused by areas developed for nonagricultural human uses and with larger populations.

Discussion

Our study is the first to link DOM quality to human population density. We found strong correlations between human population density and DOM, which identified anthropogenic influenced signatures within the DOM pool. The multiple comparisons reported in our study suggest that urban and nonagricultural human disturbed lands (i.e., developed lands), areas of higher human population density, and the altered processes that these landscapes invoke yield a DOM composition that is distinctive from DOM compositions observed in aquatic ecosystems influenced by natural land covers and agriculture land use. The ecosystem controls on the main anthropogenic marker, C6, slightly differed between lotic and lentic systems, but at their core revolved around differences in watershed human developed lands, population density, wetlands, and conductivity with no predictable association with cultivated lands. Sites with higher levels of cultivated lands were associated with more humic-like DOM composition, including C5 (a common microbial, humic-like marker), than those with more human developed lands. Moreover, our study suggests that nutrient and DOC impacts on DOM composition depend on watershed land use and cover, human population density, and the ecosystem type and its size.

We expected DOM composition to vary among freshwater ecosystems based on differences in land use and cover, nutrient concentrations, and the level of anthro-

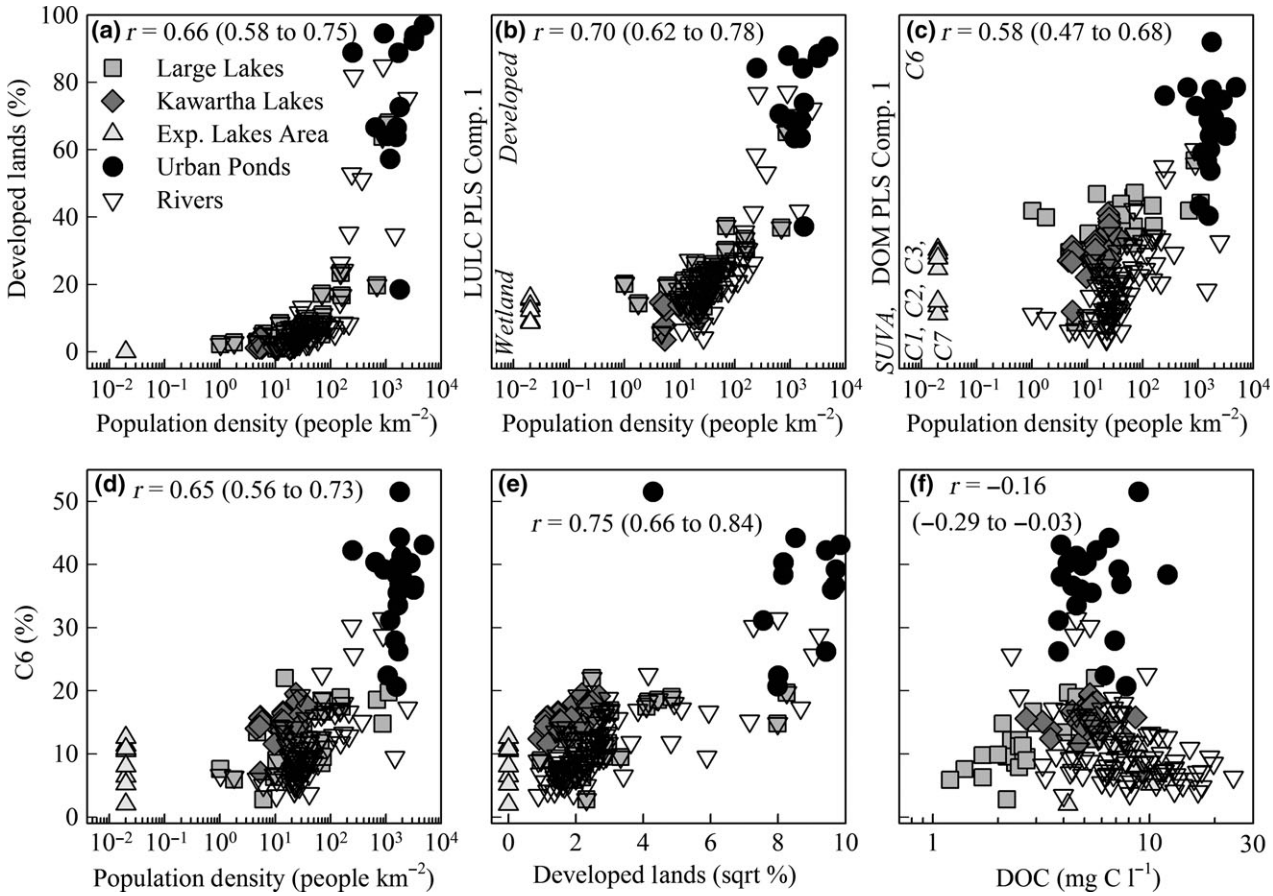

Fig. 6 Correlation ($r \pm 95\%$ effect size) between human population density and developed land use (a), land use and cover (LULC) partial least squares (PLS) component 1(b), and dissolved organic matter (DOM) PLS component 1 (c). Modeled LULC and DOM variables were developed lands, cultivated lands, rural lands, open water, wooded, wetland, and other LULC and S_R, $\beta : \alpha$, HIX, FI, SUVA, ΣF_{max}/DOC, C1, C2, C3, C4, C5, C6, and C7, respectively. In addition, correlations ($r \pm 95\%$ effect size) between PARAFAC C6 (%) and human population density (d), developed land use (e), and dissolved organic carbon (DOC; f) are highlight to illustrate the dominant human impact on DOM composition. To facilitate PLS interpretation, univariate labels are provided for variables with $r \geq 0.50$ with the PLS component. Variable labels are oriented on the figure based on the direction of their influence. X scale is set to log 10 or square root for developed lands.

pogenic influence in each respective watershed. Across these gradients in our study, water samples collected from different types of freshwater ecosystems could be broadly identified from their optical DOM composition. Large lakes had DOM composition with more recently produced, more protein-like, less humic-like and less C6 DOM characteristics. The Kawartha Lakes had an even DOM composition relative to the other ecosystems with more balanced levels of humic-like, protein-like, microbial-like DOM characteristics and average amounts of recently produced DOM. The ELA had DOM composition with relatively more protein-like, lower FI, less C6, more C2, and more degraded DOM characteristics. Urban ponds had DOM composition with relatively higher C6, lower C2, lower C1, higher F_{max}/DOC, and more recently produced DOM characteristics. Rivers, the final ecosystem type considered in our study, had DOM composition with relatively more aromatic, more humic-like, higher C5, and lower C7 DOM characteristics. Overall, these distinctions clearly separated lotic from lentic ecosystems along DOM PLS component 1 and sites with greater human populations and developed lands on DOM PLS component 2 (Fig. 3b). These ecosystem type distinctions in DOM composition were controlled strongly by watershed land use and cover and water quality and weakly by spatial location. Hence, a waterbody's DOM composition seems to share more similarity with other waterbodies of the same type than with their geographically proximate neighbors.

Aquatic DOM composition has shown distinct differences based on salinity, water residence time, soil

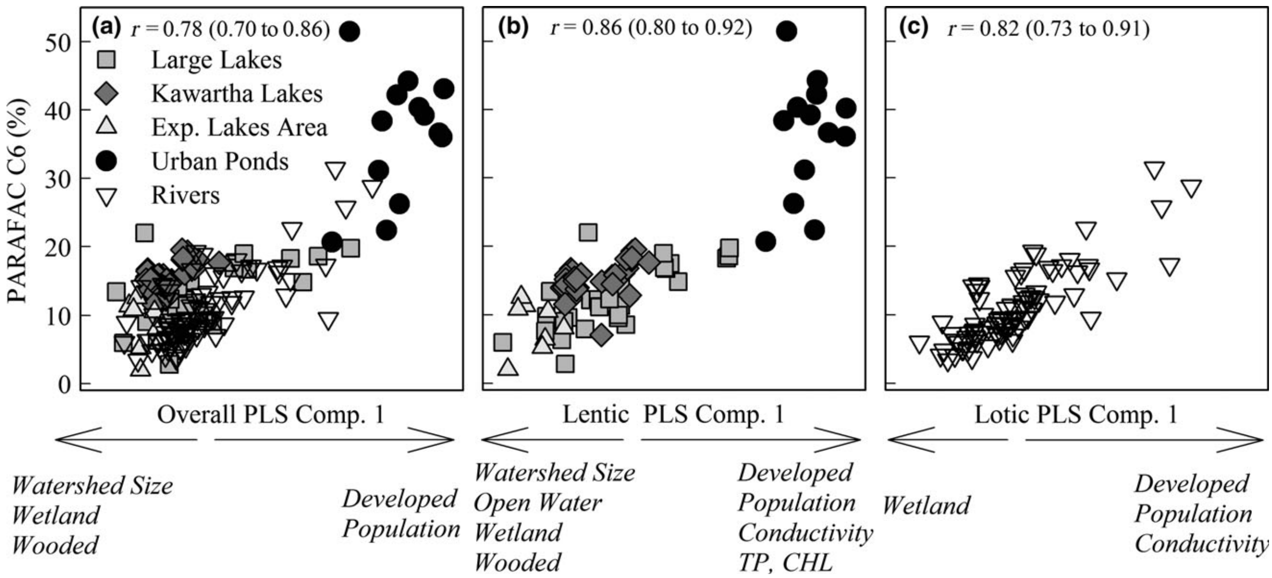

Fig. 7 Partial least squares (PLS) linear model correlation ($r \pm 95\%$ effect size) between PARAFAC C6 (%) Y-variable and 14 land use and cover (LULC) and water quality X-variables. X-variables included in the model were watershed size, human population density, developed lands, cultivated lands, rural lands, open water, wooded, wetland, other LULC, conductivity, total phosphorus, nitrate+nitrite, dissolved organic carbon, and chlorophyll a (lentic only). PLS models were run on the overall dataset (a), lotic sites (b), and lentic sites (c). Prior to use in the model, X-variables were transformed (log10 or square root) and scaled. To facilitate PLS interpretation, univariate labels are provided for variables with $r \geq 0.50$ with the PLS component and the direction of each variables influence provided. PLS loading scores for component 1 are found in Table S5.

formations, watershed land use, riparian cover, and between the inlet and outlet of open basin lakes (Larson *et al.*, 2007a; Jaffé *et al.*, 2008; Miller *et al.*, 2009; Stephens & Minor, 2010; Fellman *et al.*, 2011; Kothawala *et al.*, 2014). Across ecosystems, similar changes in the DOM composition to what we observed here have been observed along river to lake or ocean transitions, across lakes and urban ponds of different nutrient status, and along salinity gradients in coastal areas (Jaffé *et al.*, 2004; Larson *et al.*, 2007a; Webster *et al.*, 2008; Stephens & Minor, 2010; Stets *et al.*, 2010; Fellman *et al.*, 2011; Yamashita *et al.*, 2011b; McEnroe *et al.*, 2013). These DOM changes are often tied to alterations in photo- and biodegradation, water residence time, precipitation, temperature, soil-derived inputs and more generally landscape connectivity (Larson *et al.*, 2007b; Koehler *et al.*, 2012; Kothawala *et al.*, 2012, 2014; Kellerman *et al.*, 2014). Here, DOM composition across ecosystems of the Great Lakes region was influenced by interactions between human land uses and natural land covers, watershed size, eutrophication, and DOC concentration.

In our study, land use and cover remained significant drivers of DOM composition and nutrient levels across broad spatial scales when spatial location was accounted for. In Swedish lakes, DOM composition shifted in a similar manner to what we observed along a water and wetland cover to agricultural land use and nutrient enrichment gradient, but their patterns were strongly auto-correlated with latitudinal shifts in light, precipitation, and temperature (Kellerman *et al.*, 2014; Kothawala *et al.*, 2014). Across lentic and lotic ecosystems, nutrient loads, which can trigger phytoplankton blooms and enhance microbial processes, are often linked with allochthonous carbon inputs with more aromatic, larger sized, and more complex DOM signatures than that produced *in situ* (Webster *et al.*, 2008; Zhang *et al.*, 2010; Massicotte & Frenette, 2011; Solomon *et al.*, 2013). Across lentic ecosystems, allochthonous DOM and TP loading, which increased background respiration rates and gross primary production, was strongly tied to respiration in oligotrophic and humic lakes but weaker in eutrophic lakes, suggesting autochthonous DOM drives heterotrophic respiration mainly in eutrophic lakes (Solomon *et al.*, 2013; Sadro *et al.*, 2014). Although we did not quantify ecosystem metabolism in the present study, our observations support these broad scale, across ecosystem links between DOM composition and ecosystem trophic state. In our study, however, the across ecosystem DOM gradients diverged by a cultivated and a human developed land use impact, which had different interactions with DOM composition, ecosystem nutrient status, and ecosystem type.

Changes in soil organic matter inputs associated with land use disturbance influence aquatic DOM composition but do not fully account for the distinctive anthropogenic DOM composition observed in our study. For example, rivers with greater human population density and land use disturbance had proportionally older DOC (Butman *et al.*, 2015). We did not measure DOC age in our study, but this global pattern was hypothesized to correspond to altered inputs of soil organic matter in agricultural and human developed ecosystems. High $\Sigma F_{max}/DOC$ in waterbodies with increased development suggest that soil organic matter or fresh DOM inputs influenced DOM composition, but our observed patterns in DOM composition with human population densities were not uniform across agricultural and human landscapes (as was the case for aged-DOC riverine patterns [e.g., Butman *et al.*, 2015)]. Hence, other mechanisms than soil disturbance must be at play across ecosystems that forge different DOM composition between agriculturally impacted aquatic ecosystems and aquatic ecosystems with higher population densities and more developed lands.

Agriculture alters the amount and rate of nutrient delivery to the landscape, alters how water moves through the landscape, and destabilizes and exposes soil (Carpenter *et al.*, 2011; Rabotyagov *et al.*, 2014; Butman *et al.*, 2015). This agricultural disturbance influenced aquatic DOM signatures. Across ecosystems, higher levels of agricultural influence in a watershed were associated with more allochthonous, humic-like and microbial, humic-like (C5) characteristics (Fig. 4a). Similar gradients with agriculture have been found in multiple ecosystems (Fellman *et al.*, 2011; Petrone *et al.*, 2011; Kothawala *et al.*, 2014). The main difference here is that at broad spatial scales protein-like characteristics seem to be replaced by terrestrial DOM markers when ecosystem nutrient status was higher. This pattern was likely influenced by the relatively higher protein-like DOM signatures found in lentic ecosystems as compared to lotic ecosystems, which tended to have higher nutrient and DOC concentrations. When examined at finer spatial scales, within ecosystems of the same type, increases in watershed agriculture land use are associated with increases in protein-like DOM as well as microbial, humic-like markers (Wilson & Xenopoulos, 2009; Williams *et al.*, 2010; Petrone *et al.*, 2011). Overall, these patterns indicate that aquatic DOM composition in agricultural watersheds reflect terrestrial signatures, while at the same time generates microbially-derived aliphatic and protein-like components *in situ*. Hence, *in situ* processing of DOM does not seem to outpace the co-delivery of nutrients and organic matter from the watershed, though the overall balance of DOM shifts from terrestrial derived signatures in lotic ecosystems toward microbial derived signatures in lentic ecosystems.

Developed lands with higher human population densities can be associated with similar landscape disturbances as agriculture. These high density developed areas, however, tend to have more impervious surfaces, roadways, and lawns, which will modify and transport nutrients and pollutants differently than cultivated lands. In our study, we identified nonagricultural anthropogenic influences on DOM composition. Relationships between developed land use and human population density echoed that of PARAFAC C6 (Ex < 250 nm, Em = 386 nm) and human population density. This aspect of the DOM appears to be present at background levels in freshwater ecosystems with lower levels of human land uses or population density but increases markedly with developed land use and human population density (Yamashita *et al.*, 2010; Hosen *et al.*, 2014). In combination with higher specific fluorescence (i.e., higher $\Sigma F_{max}/DOC$) and $\beta:\alpha$ values in developed watersheds, C6, in particular, seems to represent a developed land use influenced DOM marker above 15% relative abundance, which resembles a different type of microbial, humic-like DOM than what is generally associated with cultivated lands and DOM very recently produced or not processed extensively *in situ*. The terrestrial DOM signatures identified in agricultural and wetland watersheds seemed relatively diminished in watersheds more influenced by high human populations and developed lands, suggesting these aspects of the DOM were processed extensively *in situ* or failed to enter the aquatic sphere.

The relative abundance of C6 showed similar but not identical patterns between lotic and lentic ecosystems (Fig. 7 and S1; Table S5). In lotic systems, C6 negatively correlated with DOC, which is typically associated with terrestrial DOM inputs in lotic ecosystem (Williams *et al.*, 2010; Petrone *et al.*, 2011), suggesting in our study relatively less terrestrial derived inputs in rivers when human population densities were higher. In contrast, C6 increased with DOC in lentic ecosystems when terrestrial and aromatic DOM markers decreased, suggesting *in situ* production. In support of our *in situ* production hypothesis, C6 also increased in lentic ecosystems with increased CHL, TP, and conductivity in higher population developed areas. In lentic systems, C6 was higher in waterbodies with smaller watersheds (mostly urban ponds), suggesting that these small but numerous aquatic ecosystems have a strong impact on local and regional aquatic DOM composition (Downing, 2010; Williams *et al.*, 2013). Nutrient and watershed size associations were weaker across lotic ecosystems, which could indicate that C6 is being transported through these systems with less *in situ* processing. The presence of C6 in large lakes above the 15% background levels suggests further that anthropogenic DOM is not immediately processed out of freshwater ecosystems as the connection with the human landscape is diluted and human altered biochemical cycles are maintained in coastal areas of the large lakes. Hence, this anthropogenic DOM composition should become dominant in many aquatic ecosystems with climate change and further land use modification to support urbanization population trends, which would alter aquatic carbon cycles differently than what is predicted for boreal and tundra areas.

Our results and those of other studies suggest small lentic ecosystems act as rapid biogeochemical reactors in anthropogenically modified landscapes (Tranvik *et al.*, 2009; Downing, 2010; Williams *et al.*, 2013). In areas with developed and urban lands with higher population densities, the DOM composition is likely controlled by *in situ* algal/cyanobacterial production of DOM and the destruction of terrestrial DOM in the presence of the aquatic pollutants associated with these landscapes. Hence, the conditions associated with aquatic ecosystems in areas with greater human population densities, promote the photo- and bio-degradation of terrestrial, humic-like DOM, production and cycling of protein-like (C7) and the more typically observed microbial, humic-like (C5) DOM markers, and production of a less reactive microbial, humic-like DOM signature (C6). Urban ponds are designed to funnel and store terrestrial inputs, the dominance of C6 in these ecosystems and other high human population watersheds, suggests a fundamental shift in terrestrial DOM processing and/or production with human disturbance (McEnroe *et al.*, 2013; Williams *et al.*, 2013). Similarly, in headwater streams, urbanization was associated with the replacement of terrestrial, humic-like DOM with more biologically reactive *in situ* produced forms (Parr *et al.*, 2015). Across temperate regions, anthropogenic and climate change disturbances are predicted to cause DOM in lentic ecosystems to experience higher *in situ* primary production, more photo- and biodegradation of DOC, and proportional increases in more stable DOM compounds (Tranvik *et al.*, 2009). Although we are not able to identify the true chemical composition of C6, it appears fairly photo-resistant and might be relatively stable in aquatic ecosystems (McEnroe *et al.*, 2013).

An additional factor that we have not yet touched on in temperate freshwater ecosystems is the positive correlation between conductivity and C6. In temperate regions, winter salt additions to roadways cause a buildup of chloride and other salts in freshwater ecosystems, which might influence DOM composition. For example, within Lake Simcoe, its tributaries, and urban ponds located within its watershed, chloride correlated positively with the relative abundance of C6 ($r = 0.54$, CI = 0.44–0.65; C.J. Williams, unpublished data). In temperate regions, salt addition should increase with human population and road density and would be much higher in developed areas than in cultivated areas. Higher nitrates in these ecosystems might also contribute to the relationship between C6 and conductivity. Perhaps, chloride pollution among other ions influences water quality and freshwater biogeochemical cycles in a way that promotes the production of C6, through compositional changes in the microbial community and/or abiotic and biotic DOM reactivities of over terrestrial DOM. More detailed and finer scaled investigations in different regions are need to clarify the mechanisms that allows C6 to become the dominant DOM signature in areas with high human population densities. This type of information will help resolve how representative this relationship is outside the Great Lakes region and determine if elevated levels of C6 are fueled by direct responses to human population densities or an indirect response through altered aquatic productivity, phytoplankton biomass, and/or contaminant loads.

Overall, we observed two noticeable forces that shaped DOM composition: (i) the land cover and use of the watershed with links to DOC and nutrient concentrations and (ii) the density of humans and developed land use in the watershed. The first observation suggests that terrestrial-like aspects of the DOM are comparatively higher across freshwater ecosystems with increased terrestrial nutrient input and increased cultivated and rural land uses. Higher protein-like DOM features were observed across ecosystems with more open water and wooded area in the watershed. The second observation indicates a novel human land use influenced DOM composition. Hence, the Anthropocene epoch may be further distinguished by the introduction of a distinctive freshwater DOM composition not typically observed in watersheds dominated by natural land cover. These findings suggest human activities and land use change have altered DOM composition in manner as distinctive as the ecosystem effects of cultural eutrophication. There is a need to expand these broad scale comparisons to include intensively managed landscapes with degraded soils, which were underrepresented in our study and are underrepresented at broad scales in the literature. Our findings also indicate human activities in and around freshwater resources influence carbon quality in predictable ways, which will have consequences toward freshwater carbon cycles and their contribution to global carbon cycles.

Acknowledgements

We thank A. Scott, J. Veldboom, B. Knights, R. Kennedy, A. Milde, our undergraduate research assistants, and the Kawartha Lakes Stewards Association for field and laboratory assistance. Funding was provided by Canada's Natural Sciences and Engineering Research Council (NSERC) Strategic programs, Environment Canada, and Lake Simcoe Clean-Up Fund to PCF and MAX. MAX acknowledges additional support from the NSERC Discovery program, an NSERC University Faculty Award, and an Early Researcher Award from the Ontario Ministry of Economic Development and Innovation. CJW acknowledges support from a postdoctoral fellowship from the Ontario Ministry of Economic Development and Innovation. JHL was supported by the Great Lakes Restoration Initiative (Project #82). Finally, we thank three anonymous reviewers for their high quality reviews of our manuscript, which lead to a better and more insightful manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

APHA (2006) Standard Methods for the Examination of Water and Wastewater.

Battin TJ, Kaplan LA, Findlay S *et al*. (2008) Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience*, **1**, 95–100.

Benner R (2003) Molecular indicators of the bioavailability of dissolved organic matter. In: *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter* (eds Findlay SEG, Sinsabaugh RL), pp. 121–138. Academic Press, New York.

Butman DE, Wilson HF, Barnes RT, Xenopoulos MA, Raymond PA (2015) Increased mobilization of aged carbon to rivers by human disturbance. *Nature Geoscience*, **8**, 112–116.

Carpenter SR, Stanley EH, Vander Zanden MJ (2011) State of the world's freshwater ecosystems: physical, chemical, and biological changes. *Annual Review of Environment and Resources*, **36**, 75–99.

Cole JJ, Prairie YT, Caraco NF *et al*. (2007) Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems*, **10**, 172–185.

Cory RM, Miller MP, Mcknight DM, Guerard JJ, Miller PL (2010) Effect of instrument-specific response on the analysis of fulvic acid fluorescence spectra. *Limnology and Oceanography: Methods*, **8**, 67–78.

Crumpton WG, Isenhart TM, Mitchell PD (1992) Nitrate and organic N analyses with second-derivative. *Limnology and Oceanography*, **37**, 907–913.

Dillon PJ, Molot L (1997) Effect of landscape form on export of dissolved organic carbon, iron, and phosphorus from forested stream catchments. *Water Resources Research*, **33**, 2591–2600.

Downing JA (2010) Emerging Global Role of Small Lakes and Ponds?: Little Things Mean a Lot. 29.

Fellman JB, Miller MP, Cory RM, D'Amore DV, White D (2009) Characterizing dissolved organic matter using PARAFAC modeling of fluorescence spectroscopy: a comparison of two models. *Environmental Science & Technology*, **43**, 6228–6234.

Fellman JB, Petrone KC, Grierson PF (2011) Source, biogeochemical cycling, and fluorescence characteristics of dissolved organic matter in an agro-urban estuary. *Limnology and Oceanography*, **56**, 243–256.

Frost PC, Larson JH, Johnston CA, Young KC, Maurice PA, Lamberti GA, Bridgham SD (2006) Landscape predictors of stream dissolved organic matter concentration and physicochemistry in a Lake Superior river watershed. *Aquatic Sciences*, **68**, 40–51.

Fry J, Xian G, Jin S *et al*. (2011) Completion of the 2006 national land cover database for the conterminous United States. *Photogrammetric Engineering & Remote Sensing*, **77**, 858–864.

Grimm NB, Foster D, Groffman P *et al*. (2008) The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. *Frontiers in Ecology and the Environment*, **6**, 264–272.

Helms JR, Stubbins A, Ritchie JD, Minor EC, Kieber DJ, Mopper K (2008) Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnology and Oceanography*, **53**, 955–969.

Hosen JD, McDonough OT, Febria CM, Palmer MA (2014) Dissolved organic matter quality and bioavailability changes across an urbanization gradient in headwater streams. *Environmental Science & Technology*, **48**, 7817–7824.

Jaffé R, Boyer J, Lu X, Maie N, Yang C, Scully N, Mock S (2004) Source characterization of dissolved organic matter in a subtropical mangrove-dominated estuary by fluorescence analysis. *Marine Chemistry*, **84**, 195–210.

Jaffé R, McKnight D, Maie N, Cory R, McDowell WH, Campbell JL (2008) Spatial and temporal variations in DOM composition in ecosystems: the importance of long-term monitoring of optical properties. *Journal of Geophysical Research-Biogeosciences*, **113**, G04032.

Kellerman AM, Dittmar T, Kothawala DN, Tranvik LJ (2014) Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nature communications*, **5**, 3804.

Koehler B, von Wachenfeldt E, Kothawala D, Tranvik LJ (2012) Reactivity continuum of dissolved organic carbon decomposition in lake water. *Journal of Geophysical Research: Biogeosciences*, **117**, G01024.

Kothawala DN, von Wachenfeldt E, Koehler B, Tranvik LJ (2012) Selective loss and preservation of lake water dissolved organic matter fluorescence during long-term dark incubations. *The Science of the total environment*, **433**, 238–246.

Kothawala DN, Stedmon CA, Müller RA, Weyhenmeyer GA, Köhler SJ, Tranvik LJ (2014) Controls of dissolved organic matter quality: evidence from a large-scale boreal lake survey. *Global change biology*, **20**, 1101–1114.

Larsen S, Andersen T, Hessen DO (2011) Climate change predicted to cause severe increase of organic carbon in lakes. *Global Change Biology*, **17**, 1186–1192.

Larson JH, Frost PC, Zheng Z, Johnston CA, Bridgham SD, Lodge DM, Lamberti GA (2007a) Effects of upstream lakes on dissolved organic matter in streams. *Limnology and Oceanography*, **52**, 60–69.

Larson JH, Frost PC, Lodge DM, Lamberti GA (2007b) Photodegradation of dissolved organic matter in forested streams of the northern Great Lakes region. *Journal of the North American Benthological Society*, **26**, 416–425.

Larson JH, Frost PC, Xenopoulos MA *et al*. (2014) Relationships between land cover and dissolved organic matter change along the river to lake transition. *Ecosystems*, **17**, 1413–1425.

Massicotte P, Frenette J-J (2011) Spatial connectivity in a large river system: resolving the sources and fate of dissolved organic matter. *Ecological Applications*, **21**, 2600–2617.

McDowell WH, Likens GE (1988) Origin, composition, and flux of dissolved organic carbon in the Hubbard Brook Valley. *Ecological Monographs*, **58**, 177–195.

McEnroe NA, Williams CJ, Xenopoulos MA, Porcal P, Frost PC (2013) Distinct optical chemistry of dissolved organic matter in urban pond ecosystems. *PLoS One*, **8**, e80334.

McKnight DM, Boyer EW, Westerhoff PK, Doran PT, Kulbe T, Andersen DT (2001) Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnology and Oceanography*, **46**, 38–48.

Mevik B-H, Wehrens R, Liland KH (2013) PLS: Partial Least Squares and Principal Component regression.

Miller MP, McKnight DM, Chapra SC, Williams MW (2009) A model of degradation and production of three pools of dissolved organic matter in an alpine lake. *Limnology and Oceanography*, **54**, 2213–2227.

Murphy KR, Butler KD, Spencer RGM, Stedmon CA, Boehme JR, Aiken GR (2010) Measurement of dissolved organic matter fluorescence in aquatic environments: an interlaboratory comparison. *Environmental science & technology*, **44**, 9405–9412.

Ohno T (2002) Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. *Environmental Science & Technology*, **36**, 742–746.

Oksanen J, Blanchet FG, Kindt R *et al*. (2015) Vegan: Community Ecology Package.

OMNR TOM of NR (2000) 28 Class Ontario Provincial Land Cover Raster.

OMNR TOM of NR (2007) Southern Ontario Interim Landcover.

Opsahl S, Benner R (1997) Distribution and cycling of terrigenous dissolved organic matter in the ocean. *Nature*, **386**, 480–482.

Parlanti E, Wörz K, Geoffroy L, Lamotte M (2000) Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. *Organic Geochemistry*, **31**, 1765–1781.

Parr TB, Cronan CS, Ohno T, Findlay SEG, Smith SMC, Simon KS (2015) Urbanization changes the composition and bioavailability of dissolved organic matter in headwater streams. *Limnology and Oceanography*, **60**, 885–900.

Paul MJ, Meyer JL (2001) Streams in the urban landscape. *Annual Review of Ecology and Systematics*, **32**, 333–365.

Petrone KC, Fellman JB, Hood E, Donn MJ, Grierson PF (2011) The origin and function of dissolved organic matter in agro-urban coastal streams. *Journal of Geophysical Research*, **116**, G01028.

R Core Team (2014) R: A Language and Environment for Statistical Computing.

Rabotyagov SS, Campbell TD, White M *et al*. (2014) Cost-effective targeting of conservation investments to reduce the northern Gulf of Mexico hypoxic zone. *Proceedings of the National Academy of Sciences*, **111**, 201405837.

Sadro S, Holtgrieve GW, Solomon CT, Koch GR (2014) Widespread variability in overnight patterns of ecosystem respiration linked to gradients in dissolved organic matter, residence time, and productivity in a global set of lakes. *Limnology and Oceanography*, **59**, 1666–1678.

Solomon CT, Bruesewitz DA, Richardson DC *et al.* (2013) Ecosystem respiration : drivers of daily variability and background respiration in lakes around the globe. *Limnology and Oceanography*, **58**, 849–866.

Stanley EH, Powers SM, Lottig NR, Buffam I, Crawford JT (2012) Contemporary changes in dissolved organic carbon (DOC) in human-dominated rivers: is there a role for DOC management? *Freshwater Biology*, **57**, 26–42.

Stephens BM, Minor EC (2010) DOM characteristics along the continuum from river to receiving basin: a comparison of freshwater and saline transects. *Aquatic Sciences*, **72**, 403–417.

Stets EG, Striegl RG, Aiken GR (2010) Dissolved organic carbon export and internal cycling in small, headwater lakes. *Global Biogeochemical Cycles*, **24**, GB4008, doi:10.1029/2010GB003815.

Tranvik LJ, Downing JA, Cotner JB *et al.* (2009) Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, **54**, 2298–2314.

Webster KE, Soranno PA, Cheruvelil KS *et al.* (2008) An empirical evaluation of the nutrient-color paradigm for lakes. *Limnology and Oceanography*, **53**, 1137–1148.

Weishaar JL, Aiken GR, Bergamaschi BA, Fram MS, Fujii R, Mopper K (2003) Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environmental Science & Technology*, **37**, 4702–4708.

Williams CJ, Yamashita Y, Wilson HF, Jaffé R, Xenopoulos MA (2010) Unraveling the role of land use and microbial activity in shaping dissolved organic matter characteristics in stream ecosystems. *Limnology and Oceanography*, **55**, 1159–1171.

Williams CJ, Scott AB, Wilson HF, Xenopoulos MA (2012) Effects of land use on water column bacterial activity and enzyme stoichiometry in stream ecosystems. *Aquatic Sciences*, **74**, 483–494.

Williams CJ, Frost PC, Xenopoulos MA (2013) Beyond best management practices: pelagic biogeochemical dynamics in urban stormwater ponds. *Ecological Applications*, **23**, 1384–1395.

Williamson CE, Morris DP, Pace ML, Olson OG (1999) Dissolved organic carbon and nutrients as regulators of lake ecosystems: resurrection of a more integrated paradigm. *Limnology and Oceanography*, **44**, 795–803.

Wilson HF, Xenopoulos MA (2008) Ecosystem and seasonal control of stream dissolved organic carbon along a gradient of land use. *Ecosystems*, **11**, 555–568.

Wilson HF, Xenopoulos MA (2009) Effects of agricultural land use on the composition of fluvial dissolved organic matter. *Nature Geoscience*, **2**, 37–41.

Xenopoulos MA, Lodge DM, Frentress J, Kreps TA, Bridgham SD, Grossman E, Jackson CJ (2003) Regional comparisons of watershed determinants of dissolved organic carbon in temperate lakes from the Upper Great Lakes region and selected regions globally. *Limnology and Oceanography*, **48**, 2321–2334.

Yamashita Y, Scinto LJ, Maie N, Jaffé R (2010) Dissolved organic matter characteristics across a subtropical Wetland's landscape: application of optical properties in the assessment of environmental dynamics. *Ecosystems*, **13**, 1006–1019.

Yamashita Y, Kloeppel BD, Knoepp J, Zausen GL, Jaffé R (2011a) Effects of watershed history on dissolved organic matter characteristics in headwater streams. *Ecosystems*, **14**, 1110–1122.

Yamashita Y, Panton A, Mahaffey C, Jaffé R (2011b) Assessing the spatial and temporal variability of dissolved organic matter in Liverpool Bay using excitation–emission matrix fluorescence and parallel factor analysis. *Ocean Dynamics*, **61**, 569–579.

Zhang Y, Zhang E, Yin Y *et al.* (2010) Characteristics and sources of chromophoric dissolved organic matter in lakes of the Yungui Plateau, China, differing in trophic state and altitude. *Limnology and Oceanography*, **55**, 2645–2659.

Zsolnay A, Baigar E, Jimenez M, Steinweg B, Saccomandi F (1999) Differentiating with fluorescecne spectroscopy the sources of dissolved organic matter in soils subjected to drying. *Chemosphere*, **38**, 45–50.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Seven category land use and cover reclassification.
Table S2. Pearson's correlation matrix for water quality variables across freshwater systems.
Table S3. Pearson's correlation matrix for optical dissolved organic matter (DOM) variables.
Table S4. PLS components 1 and 2 loadings for DOM as a function of LULC (Fig. 4).
Table S5. PLS component 1 loadings for the x variables included in the predictive model results presented in Fig. 7 for PARAFAC C6.
Figure S1. Pearson's correlation between C6 and land use and cover and water quality variables.
Figure S2. Box and Whisker plot of dissolved organic matter optical indices and PARAFAC components.
Figure S3. Partial least squares (PLS) linear model correlation between land use and cover and water quality variables.
Figure S4. Bivariate correlations between PARAFAC C6 and water quality variables.