REFERENCES
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/10.2307/26932472?seq=1&cid=pdf-
reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

Ecosystem responses to channel restoration decline with stream size in urban river networks

Peter S. Levi,[1,3] and Peter B. McIntyre[2]

Center for Limnology, University of Wisconsin-Madison, Madison, Wisconsin 53706 USA

Abstract. Urban streams are often severely impaired due to channelization, high loads of nutrients and contaminants, and altered land cover in the watershed. Physical restoration of stream channels is widely used to offset the effects of urbanization on streams, with the goal of improving ecosystem structure and function. However, these efforts are rarely guided by strategic analysis of the factors that mediate the responsiveness of stream ecosystems to restoration. Given that ecological gradients from headwater streams to mainstem rivers are ubiquitous, we posited that location within a river network could mediate the benefits of channel restoration. We studied existing stream restorations in Milwaukee, Wisconsin, to determine (1) whether restorations improve ecosystem function (e.g., nutrient uptake, whole-stream metabolism) and (2) how ecosystem responses vary by position in the urban river network. We quantified a suite of physicochemical and biological metrics in six pairs of contiguous restored and concrete channel reaches, spanning gradients in baseflow discharge (19–196 L/s) and river network position (i.e., headwater to mainstem). Hydrology differed dramatically between the restored and concrete reaches; water velocity was reduced 2- to 13-fold while water residence time was 50–5,000% greater in adjacent restored reaches. Restored reaches had shorter nutrient uptake lengths for ammonium, nitrate, and phosphate, as well as higher whole-stream metabolism. Furthermore, the majority of reaches were autotrophic (i.e., gross primary production > ecosystem respiration), which is not common in stream ecosystems. The difference in ecosystem functioning between restored and unrestored reaches was generally largest in headwaters and declined to equivalence in mainstem restorations. Our results suggest that headwater sites offer higher return on investment compared to larger downstream channels, where ecosystem responsiveness is low. If this pattern proves to be general, the scaling of ecosystem responses with river size could be integrated into planning guidelines for urban stream restorations to enhance the societal and ecological benefits of these expensive interventions.

Key words: nutrient spiraling; nutrient uptake; river network; stream restoration; transient storage; urban ecology; whole-stream metabolism.

Introduction

In urban areas, streams suffer severe ecological impairment due to physical modification of the watershed, the riparian zone, and the stream channel itself. The collective effects of urbanization on stream ecosystems are consistent enough across biomes to be labeled as the urban stream syndrome (Walsh et al. 2005a). For example, urban streams are often funneled into linear concrete channels or buried underground (Kaushal and Belt 2012), greatly simplifying their physical structure and exchange with surrounding ecosystems (Beaulieu

et al. 2014). Channelization of urban streams can be reversed by engineering them back to a more sinuous geomorphology, but the degree to which physical restoration also returns hydrological and ecological processes (i.e., ecosystem function) to natural ranges remains uncertain. A key limitation is that, despite enormous investments in stream restoration over the last few decades, few projects are assessed or monitored after their completion (Bernhardt et al. 2005).

Stream restorations generally focus on improving the physical structure of stream channels, in hopes that subsequent responses of chemical and biological processes will yield improvements from both ecological and societal perspectives. The logic of this expected physical-chemical-biological cascade of responses is that hydrology is a key control on reaction rates, so increasing the residence time of water in a stream reach by removing pipes and concrete channels is requisite to improving ecosystem function. The return of a more naturalistic

flow regime would then spur shifts in a wide range of stream processes, including nutrient retention and whole-stream metabolism (Newcomer-Johnson et al. 2016). However, the evidence for ecological responses to restorations in urban streams is inconsistent. For instance, nutrient retention can be higher in a restored vs. unrestored reach (Kaushal et al. 2008); yet restoring natural channel geomorphology does not always change transient storage (Becker et al. 2013). Regarding ecosystem function, some studies document altered process rates following channel restoration, such as for whole-stream metabolism (Pennino et al. 2014), while others find comparable rates between restored and unrestored reaches (Sudduth et al. 2011). These contradictory patterns, along with the general dearth of post-restoration impact assessments, make it difficult to predict the ecological benefits of expensive projects to restore stream channels.

The location of a reach within a river network gives rise to many predictable shifts in ecosystem processes from headwater streams down to mainstem rivers (Vannote et al. 1980, Hall et al. 2013). In urban landscapes, predictable shifts in ecological processes may be apparent as well, but dictated by the heterogeneous physical alteration of streams throughout the river network. For example, headwater streams are often disrupted by pipes, leading to a disconnection between the landscape and the stream, whereas mainstem rivers are often constrained and channelized (Kaushal and Belt 2012). These gross physical modifications can strongly affect stream ecosystem processes in urban watersheds, such as stream corridors in Baltimore, Maryland, USA. There, lack of riparian shading results in high rates of gross primary production (GPP) that increase with distance downstream in the river network, leading in turn to higher rates of nitrogen uptake (Kaushal et al. 2014). Such longitudinal patterns in urban rivers suggest that network position could influence the responsiveness of stream ecosystems to restoration of the channel or surrounding watershed.

Channel restoration cost and expected societal benefits may also vary systematically based on location within an urban river network. The small size of headwater streams facilitates the logistics of restoring an impaired channel, enhances biogeochemical reaction rates (Peterson et al. 2001, Claessens et al. 2010), enhances macroinvertebrate biodiversity (Swan and Brown 2017), and could maximize the downstream extent of water quality gains. In contrast, the restoration of a low-gradient, mainstem channel may not have dramatic ecological benefits because water quality in larger streams reflects the influence of the entire upstream river network (Violin et al. 2011, Jarvie et al. 2018). In addition, the conditions in headwater streams, many of which are impaired in urban settings, have a greater impact on mainstem water quality than the riparian habitat along the mainstem itself (Dodds and Oakes 2006). Merging longitudinal perspectives on both

restoration cost and ecosystem responsiveness could foster cost-efficient restoration planning, potentially favoring headwater stream sites. However, mainstem restorations could still provide important ecosystem services, such as increasing greenspace and offering recreational opportunities, in portions of the urban landscape where population density may be highest. Indeed, return on investment depends on the outcomes being sought from both ecological and societal perspectives, underscoring the importance of selecting appropriate metrics to quantify responses to channel restoration.

Measuring the response of stream ecosystems to channel restoration is complicated by (1) a lack of accepted metrics to assess improvement and (2) disparities in the spatial footprint of response metrics relative to restored areas. Integrative metrics of ecosystem function are appealing for monitoring and evaluating restorations throughout urban river networks. For instance, whole-stream metabolism and nutrient cycling dynamics provide time-integrated measures of key ecosystem processes that reflect characteristics of the local reach, upstream channel, and broader watershed. Whole-stream metabolism has been used to assess stream ecosystem health (Fellows et al. 2006, Young et al. 2008), and is readily applicable to reach-scale restorations in urban watersheds. Furthermore, gross primary production (GPP) is sensitive to land use (Clapcott et al. 2010), suggesting its utility for assessing the improvements in riparian areas following restoration. However, the metabolism of a given reach reflects a longer length of channel and area of watershed than is modified in most restoration projects, creating doubt about whether restoring a short reach (<500 m long) can reverse watershed-wide impairment of urban streams (Bernhardt and Palmer 2011). A previous study comparing restored and unrestored streams in urban watersheds found no differences in ecosystem metabolism despite more habitat heterogeneity and longer water residence time in restored reaches (Sudduth et al. 2011). Yet comparisons among reach-scale restorations across a river network suggest that physical improvements of larger channels lead to greater increases in nitrogen retention and removal than in headwater streams (Filoso and Palmer 2011). The conflicting implications of these studies of metabolism and nutrient uptake illustrate the need to assess a suite of functional metrics simultaneously to identify robust and sensitive tools for quantifying improvements in stream ecosystem function arising from restoration projects.

Here, we evaluate whether restorations can improve stream ecosystem processes in an urban river network, and whether the magnitude of restoration benefits varies predictably with stream size. We compared a suite of metrics between contiguous unrestored and restored reaches of six streams in Milwaukee, Wisconsin (USA), where restoration involved replacing the concrete lining of the stream channel with heterogeneous substrates and more natural geomorphology (e.g., meanders, pools,

riffles). Replicate comparisons between paired reaches offer a powerful test of the relative responsiveness of physical, chemical, and biological metrics of stream ecosystem function to restoration. To synthesize across our paired comparisons, we tested for patterns of responsiveness with respect to stream size, spanning a gradient from headwater streams to a mainstem river. If urban restoration benefits scale with stream size or reach position within the river network, these factors could be readily incorporated into restoration planning to select sites that maximize local and downstream benefits for urban watersheds. Thus, our overarching aims were to inform expectations from urban stream restorations and elucidate strategies for selecting the most responsive restoration sites.

MATERIALS AND METHODS

Study streams and design

We studied six restored stream reaches in Milwaukee, Wisconsin, USA. The stream reaches were located in heavily urbanized watersheds, with the proportion of low, medium, and high intensity developed land cover ranging from 58.9% to 90.1% and impervious surface covering 30.3% to 54.5% of the watershed area (Homer et al. 2015; Table 1). All watersheds were ≥99.2% developed when we included urban open space (e.g., parks, lawns, golf courses) in our land cover analysis, with the exception of one (i.e., UWD; 91.0% developed; see Table 1 for watershed names). The primary goal of the channel restoration in these watersheds was to prevent flooding in adjacent streets and houses, though water quality, aesthetics, and recreational opportunities were also considered (U.S. Army Corps of Engineers 2016). These restorations have effectively reduced the frequency and magnitude of flooding in the immediate vicinity, but their outcomes have not been assessed using metrics of stream ecosystem structure or function.

We chose study reaches to represent a gradient in stream discharge, which differed among streams by one order of magnitude (baseflow discharge 19–196 L/s; Table 1). VLM was both the smallest stream and the most recent restoration project, completed in 2011 (~3 yr before our study). All other restorations were at least eight years old, and two were completed ~25 (WLP) and ~80 (HNY) years ago. Our observations suggested that ecosystem conditions within all reaches had recovered from the disturbance of restoration activities by the time of our study, despite differences in the time since channel re-naturalization. We did not formally assess the potential effects of time since restoration because imbalance in the age distribution precludes quantitative analysis, but stream ecosystem dynamics, particularly in urban watersheds, are sufficiently disturbance-driven that

TABLE 1. Watershed and stream characteristics and ambient nutrient concentrations (mean ± SE) of the six study streams in Milwaukee, Wisconsin (USA).

Stream	Latitude (°N)	Longitude (°W)	Watershed area (km²)	Baseflow discharge (L/s)	Year restoration completed	Impervious surface (%)	Developed total (%)	Developed open (%)	Developed low (%)	Developed medium (%)	Developed high (%)
Villa Mann (VLM)	42.965	87.944	3.0	19	2010	42.0	100.0	20.5	42.6	28.3	8.6
Southbranch (SBH)	43.166	87.978	4.6	20	2001	42.1	100.0	18.0	51.9	18.1	12.0
Honey (HNY)	43.042	88.012	27.3	57	1934	47.4	99.9	9.8	47.3	33.7	9.1
Wilson Park (WLP)	42.962	87.920	17.0	95	1988	54.5	100.0	14.9	27.3	31.2	26.6
Underwood (UWD)	43.048	88.047	48.0	147	2010	30.3	91.0	32.1	37.8	14.3	6.8
Kinnickinnic (KKR)	42.996	87.919	55.4	191	2011	51.7	99.2	11.9	31.5	39.3	16.5

Notes: Stream abbreviations are used throughout the manuscript. Streams are ordered by baseflow discharge from July and August. Developed land cover was determined using data from the NLCD 2011 database.

short-term (daily to seasonal) variation in weather and discharge are expected to exceed any differences arising from years since restoration (Larsen and Harvey 2017). Our analyses focus on stream size and river network position, the factors that motivated our selection of sites, but we recognize that our cross-site comparisons could be confounded with decisions by municipalities and private land owners. For instance, land use surrounding the study reaches ranged from a greenway park (HNY) to single-family residential properties (VLM, SBH, WLP) to a municipal floodplain (UWD) to a dense urban residential neighborhood (KKR). These complications temper inferences about controls on responses to restoration, but pairing contiguous restored and unrestored reaches at each site nonetheless allows strong interpretation of the general effects of stream restoration on ecosystem function.

In each stream, we paired a restored reach with a contiguous channelized concrete reach. Both reaches were the same length within each stream, but varied in length between streams from 150 m in the smallest stream (VLM) to 300 m in the largest (UWD and KKR; Table 2) due to differences in the length of restored channel. The concrete reach was immediately upstream of the restored reach in four streams (HNY, KKR, UWD, KKR), and downstream in two (SBH, VLM) because the area above the restoration was not accessible. In every case, an abrupt shift in channel geomorphology marked the division between the paired reaches, and we measured reach-scale functional metrics at the boundary between contiguous study reaches. We collected data during July, August, and September 2014, and measured each response metric on the same date for both reaches in each stream. We recorded all metrics at baseflow discharge, and treated metrics from the paired reaches as independent observations for statistical analysis.

To determine the effect of restoration, we calculated the log response ratio (Hedges et al. 1999) as

$$L = \log_{10} \frac{\overline{x}_{\text{RESTORATION}}}{\overline{x}_{\text{CONCRETE}}}$$

where $\overline{x}_{\text{RESTORATION}}$ and $\overline{x}_{\text{CONCRETE}}$ are the mean values for the restored and concrete reaches, respectively. Response ratios offer a concise way to summarize the disparity between the contiguous reaches with a single quantity. A positive or negative L indicates the metric was higher or lower, respectively, in the restored reach by a factor of 10^L, such that $L = 0$ signifies no differences between the reaches, and $L = 1$ indicates a 10-fold higher response in the restored reach.

Physicochemical and biological variables

We quantified the following physicochemical variables on multiple dates ($n = 5$–12) during the 3-month study: discharge, width, depth, velocity, and dissolved inorganic nitrogen and phosphorus. We determined the channel geomorphology by averaging the width of 10 transects in each reach and five equidistant depth measurements per transect. To measure water velocity through the reach, we released a salt or rhodamine-WT (RWT) dye pulse and recorded the timing of downstream changes in conductivity or RWT, respectively. We used the average width (W) and depth (D) as well as the reach velocity (v) to calculate discharge ($Q = W \times D \times v$). To determine the concentrations of nutrients, we collected water samples on each sampling date at the top and bottom of each reach. In the field, we filtered water samples into acid-washed centrifuge tubes using a syringe and ashed glass-fiber filters (Whatman GFF) and froze them until laboratory analysis. We analyzed ammonium (NH_4^+), nitrate (NO_3^-), and soluble reactive phosphorus (SRP) using the phenol-hypochlorite, cadmium-reduction, and ascorbic acid methods, respectively, on an Astoria-Pacific II segmented flow autoanalyzer (Clackamas, Oregon, USA; APHA 2005). Background nutrient concentrations were similar at all points in the paired reaches on each sampling date (Table 2).

We measured static physicochemical variables once at each site (i.e., sinuosity, slope, canopy cover). To calculate the sinuosity of each reach, we divided the length of the thalweg by the straight-line distance between the top and bottom of the reach (Gordon et al. 2004). We determined the slope of each reach using a digital clinometer and stadia rod (Bisson et al. 2017). We measured canopy cover across all streams over two days during peak foliage using a spherical densiometer at three points in each reach (Hill 2017).

We quantified benthic resources (e.g., chlorophyll *a*, fine particulate organic matter), sediment size, and habitat cover within two to three days of conducting the functional measurements at each stream. We scrubbed biofilm off of benthic surfaces (e.g., concrete channel, gravel, rocks), measured the volume of the slurry, and homogenized and subsampled it for chlorophyll *a* and organic matter analysis (Steinman et al. 2017). When larger rock substrates were absent, we collected sand samples to 1 cm depth using a small sediment core (area = 4.9 cm^2). We analyzed chlorophyll *a* using a fluorometer following ethanol extraction, and quantified organic matter content (i.e., OM) by drying, weighing, ashing, and re-weighing the sediments to calculate the percent OM (Steinman et al. 2017). To estimate the size distribution of benthic substrates, we measured the median diameter of 100 randomly selected sediments in restored reaches using a gravelometer for sediments between 2 and 256 mm and a meter stick for sediments >256 mm. We also estimated the proportional coverage of the different benthic habitats (i.e., sand, gravel, rock) using a quadrat across each width–depth transect ($n = 10$ transects per reach). To determine the habitat-weighted chlorophyll *a* and organic matter content in the reach, we multiplied the areal concentration of each habitat by the proportional coverage of that habitat.

TABLE 2. Metrics of ecosystem structure in the restored reaches and concrete channels of the six study streams.

Stream and reach type	Reach length (m)	Width (m)†	Depth (cm)†	Canopy cover (%)	Slope (%)	NH_4^+ (µg N/L)‡	NO_3^- (µg N/L)‡	SRP (µg P/L)‡	Chl a (mg/m²)§	OM (g/m²)§
VLM										
Restored	150	2.66 (0.15)	11.9 (1.0)	5.2	0.2	21 (4)	169 (99)	9 (3)	82 (21)	110 (23)
Concrete	150	0.61 (0.04)	3.9 (0.3)	2.1	1.5	16 (3)	142 (91)	8 (3)	431 (73)	58 (8)
SBH										
Restored	200	2.34 (0.18)	10.2 (0.9)	64.9	0.2	81 (9)	214 (72)	148 (20)	319 (61)	97 (9)
Concrete	200	1.20 (0.05)	4.4 (0.4)	80.2	0.2	69 (10)	206 (71)	151 (23)	200 (39)	34 (3)
HNY										
Restored	200	5.78 (0.38)	6.6 (0.7)	64.0	0.6	26 (3)	405 (94)	57 (13)	296 (32)	84 (4)
Concrete	200	2.73 (0.18)	3.7 (0.3)	82.1	0.2	28 (4)	449 (92)	61 (14)	155 (23)	22 (2)
WLP										
Restored	250	6.28 (0.10)	21.5 (0.7)	56.4	0.2	97 (40)	195 (62)	123 (49)	245 (18)	161 (10)
Concrete	250	4.56 (0.95)	7.5 (0.7)	3.0	0.2	90 (31)	207 (57)	120 (57)	392 (57)	127 (14)
UWD										
Restored	300	3.89 (0.19)	24.6 (2.4)	0.0	0.2	23 (2)	61 (32)	11 (4)	384 (27)	187 (8)
Concrete	300	7.79 (0.53)	7.0 (0.8)	15.9	0.2	24 (2)	71 (36)	11 (3)	227 (25)	207 (17)
KKR										
Restored	300	7.65 (0.28)	19.6 (1.3)	6.6	0.2	38 (6)	206 (41)	27 (6)	386 (52)	118 (7)
Concrete	300	6.26 (0.46)	8.1 (0.7)	4.0	0.3	20 (2)	139 (45)	24 (6)	411 (85)	93 (14)

Note: Values are means with SE in parentheses. SRP, soluble reactive phosphorus; chl a, chlorophyll a; OM, organic matter.
†Mean width and depth at baseflow discharge from July and August.
‡Nutrient concentrations represent mean over 3-month study period at the downstream sampling point for each reach.
§Reach-weighted mean values.

Transient storage metrics

As water flows downstream in lotic ecosystems, it can be temporarily retained or slowed in areas of the stream channel such as pools or gravel bars, which are collectively referred to as transient storage zones. To determine the transient storage of the stream reaches, we conducted short-term releases of RWT. We injected RWT into a well-mixed portion of the stream at a rate of 200 mL/min and measured concentrations downstream every 15 s using a calibrated RWT sensor (Turner Designs and Precision Measurement Engineering, Vista, California, USA). In order to ensure saturation through the reach, we continued injections until a consistent plateau concentration was achieved (target plateau concentration = 20 μg/L); therefore, the duration of the injections varied across reaches and streams. Rhodamine is not entirely conservative in stream ecosystems (Runkel 2015), but given the short-term releases we conducted, our data provide a sound and consistent assessment of transient storage dynamics for comparison among reaches and streams.

We modeled the transient storage dynamics using the RWT break-through curves and the one-storage zone transient storage model (Runkel 1998). The model output included optimized estimates and standard error of stream water dispersion (D, m^2/s) and velocity (u, m/s) as well as the exchange coefficient with the storage zone (α, s^{-1}) and mean water residence time (WRT, h). From these values, and our measurements of discharge (Q, L/s) and channel width (w, m), we determined the area of the storage zone (A_S) and calculated the ratio of storage zone to active channel area (A_S/A). We also calculated the fraction of median transport time due to storage using the metric F_{MED}^{200}, which allowed us to compare the relative importance of storage between releases measured in streams of different sizes (Runkel 2002).

Nutrient uptake dynamics

In stream ecosystems, dissolved nutrients spiral between the water column and biota during unidirectional transport downstream (Newbold et al. 1981). We calculated the uptake length (i.e., distance nutrients travel before removal from the water column; S_W, m), uptake velocity (i.e., relative demand; v_f, mm/s), and areal uptake rates (U, mg·m^{-2}·h^{-1}) of stream nutrients by conducting short-term additions of ammonium (NH$_4^+$ as NH$_4$Cl) or of nitrate (NO$_3^-$ as NaNO$_3$) plus soluble reactive phosphorus (PO$_4^{3-}$ as KH$_2$PO$_4$; see Stream Solute Workshop 1990 for calculations). We included sodium bromide (NaBr) as a conservative tracer to account for hydrological dilution. We combined NO$_3^-$ and PO$_4^{3-}$ releases to complete all nutrient releases on the same day. It is possible that combined releases could increase uptake rates of one nutrient by alleviating limitation of the other, but this effect would be minimal during the short releases (i.e., 30–

45 minutes) used in this study. Any such bias would be comparable between paired reaches and would have little impact on between-stream comparisons given the differences in background nutrient availability and ratios. We targeted levels of enrichment to minimize the increase in concentrations while still enabling detectability of nutrient uptake (20–40 μg N/L for NH$_4^+$, 60–150 μg N/L for NO$_3^-$, and 15–40 μg P/L for SRP). Even our temporary augmentation of nutrient concentrations may have enhanced dissimilatory nutrient transformations that contribute to nutrient removal from the water column; hence, our results were likely to overestimate nutrient uptake rates (Mulholland et al. 2002). However, such changes in ambient stream water nutrient concentrations occur regularly with storm-driven discharge fluctuations in urban streams (Filoso and Palmer 2011), and our intent in measuring nutrient dynamics at baseflow conditions was simply to compare potential uptake across the 12 study reaches.

Whole-stream metabolism

Whole-stream metabolism represents the balance of energy production by autotrophs against energy consumption by heterotrophs. We calculated gross primary production (GPP) and ecosystem respiration (ER) from diel fluctuations in dissolved oxygen (DO) and temperature logged every 10 minutes (HOBO U-26; Onset Computer Corporation, Bourne, Massachusetts, USA). Photosynthetically active radiation (PAR) was recorded in the riparian zone (Odyssey light meter; Dataflow Systems, Christchurch, New Zealand). To account for seasonal variation during the study, we maintained one set of sensors at KKR, the largest stream, for the duration of our study (i.e., July to September). We rotated the other DO and PAR sensors among the five streams approximately every 7–14 d. For analysis, we pooled one week of metabolism data at baseflow conditions from each reach from each month (July, August, and September; total n = 18–21; Data S1: Metabolism output).

To estimate GPP, ER, and the reaeration coefficient (K_{600}; gas exchange with the atmosphere), we used a modified Bayesian metabolic model fitted to diel DO, temperature, and PAR data (Holtgrieve et al. 2010). We executed metabolism models using the JAGS package in the R statistical environment, including posterior distributions for model parameters based on a Markov Chain Monte Carlo analysis (MCMC; R Core Team 2013). We ran three MCMC chains with 50,000 iterations and a 10,000 iteration burn, which were thinned by 5 to yield 10,000 draws. We evaluated model convergence by testing the MCMC chains for autocorrelation, visually examining chains to evaluate mixing, and using R-hat statistics (Gelman and Hill 2007). The model fits of the dates we included in our analysis were all strong (R^2 < 0.9 for only 3 of 245 models), and the upper and lower bounds of both GPP and ER were consistently within 0.05–0.8 g O$_2$·m^{-2}·d^{-1} of the estimated value

(median difference = 0.37 g $O_2 \cdot m^{-2} \cdot d^{-1}$; Data S1: Metabolism output). Given the short length of our study reaches, the oxygen footprint represented by our fitted GPP and ER estimates usually extended well beyond the reach boundaries, sometimes by more than 100-fold. Nonetheless, our one-station, whole-stream metabolism data provide a robust comparative perspective on ecosystem function because of the stark contrast in hydrology and geomorphology between the concrete and restored reaches. To synthesize metabolism results, we also estimate net ecosystem production (NEP) as GPP minus ER.

Comparative statistical analyses

We used analysis of covariance (ANCOVA) to test ecosystem response metrics for effects of reach classification (concrete vs. restored) and discharge across the six streams. We also evaluated organic material in the stream benthos (i.e., biofilm chlorophyll a and percent OM) using ANCOVA to compare three substrate types: rock and sand substrates in the restored reach and concrete substrate in the unrestored reach. In addition, we used a one-sample t test to investigate whether the L values differed from zero, followed by Pearson's correlations to determine whether L values were associated with discharge. The ANCOVAs allowed us to compare the absolute magnitude of each metric across all reaches. The t tests and correlations on the L ratios allowed us to focus on the consistency of responses to restoration. Both sets of statistics also enabled us to test for the influence of site position within the river network.

Finally, we evaluated the strength of literature-based predictors to explain patterns in the hydrological metrics (i.e., dispersion, A_S/A, WRT, F_{MED}^{200}), nutrient uptake dynamics (i.e., S_W, v_f, U), and whole-stream metabolism (i.e., GPP, ER, NEP) across these urban streams. We examined models for each response metric with one or two factors that had proven important in previous studies and identified the most parsimonious model based on the corrected Akaike information criterion (AIC$_c$; Burnham and Anderson 2002). For example, we used characteristics of channel geomorphology to examine variation in hydrological metrics, while models for nutrient uptake and whole-stream metabolism included combinations of physical, chemical, and biological predictors. The goal of these models was to identify environmental factors that might serve as coarse proxies for responsiveness of urban stream ecosystems to restoration efforts.

RESULTS

Physical, chemical, and biological response of urban streams to restoration

The hydrology of the study reaches differed starkly between contiguous restored and concrete channels. Stream water velocity was more than four times higher in concrete reaches than restored reaches (0.32 ± 0.07 vs. 0.07 ± 0.01 m/s, respectively [mean $\pm$ SE]; Fig. 1A; ANCOVA, $P_{REACH} < 0.001$) and dispersion tended to be higher as well (0.43 ± 0.16 vs. 0.15 ± 0.04 m²/s, respectively; Fig. 1B; ANCOVA, $P_{REACH} = 0.09$). Correspondingly, the metrics of transient storage were higher in the restored reaches relative to the concrete reaches (Fig. 1C–E; ANCOVA, $P_{REACH} < 0.05$). For example, the relative area of the storage zone and mean residence time were two and five times greater, respectively, in the restored vs. concrete reaches. However, stream velocity, dispersion, and the transient storage metrics did not vary by stream size (ANCOVA, $P_{DISCHARGE} \geq 0.15$). Regarding the L ratios, stream velocity ratios were consistently negative (i.e., restored reach < concrete reach; t test, $P < 0.003$), and those for dispersion tended to be below zero as well (Fig. 1F, G, respectively), whereas L ratios for the transient storage metrics were overwhelmingly positive (Fig. 1H–J; t test, $P \leq 0.05$). None of the L ratios of the hydrological metrics were correlated with discharge, though the difference in water velocity between the concrete and restored reaches was higher in the headwater streams (Fig. 1F).

The structure of the stream benthos varied by habitat and stream. Benthic chlorophyll-a concentrations indicated higher algal biomass in the concrete reach and on rocks in the restored reach (307 ± 25 and 290 ± 23 mg/m², respectively) relative to the sand in the restored reach (44 ± 5 mg/m²; ANCOVA, $P_{HABITAT} < 0.001$). In addition, chlorophyll a concentrations across all three habitats tended to increase with stream size (Fig. 2A; $P_{DISCHARGE} = 0.001$). As with chlorophyll a, the percentage of organic matter in the benthic biofilms was higher in the concrete channel and rock habitats relative to sand ($29.4 \pm 1.4\%$, $27.4 \pm 1.1\%$, and $1.9 \pm 0.2\%$, respectively; ANCOVA $P_{HABITAT} < 0.001$). While percent OM also varied by stream size ($P_{DISCHARGE} < 0.001$), the variation depended on habitat ($P_{HABITAT \times DISCHARGE} < 0.001$). The percent OM on the channel and rocks decreased as stream size increased, whereas percent OM of sand did not change with stream size (Fig. 2B). Given the variation in habitat types between the restored and concrete reaches, we did not calculate and assess L ratios for the chlorophyll a or percent OM data.

Nutrient uptake dynamics of the urban streams varied by nutrient form and restoration, but the L ratios were not consistently positive or negative. S_W was shortest for SRP and longest for NO_3^- (206 ± 46 m and 457 ± 105 m, respectively; ANCOVA, $P_{NUTRIENT} = 0.03$). Among all three nutrient forms, S_W was shorter in restored vs. concrete reaches (ANCOVA, $P_{REACH} = 0.04$), and generally longer in large streams compared to headwaters (Table 3; $P_{DISCHARGE} = 0.07$). The L ratios for S_W were negative in five of the six streams for NH_4^+ and NO_3^- (Fig. 3A, B, respectively; t test, $P = 0.1$ and 0.06, respectively), and all six streams for SRP (Fig. 3C; t test, $P = 0.01$). The L ratios
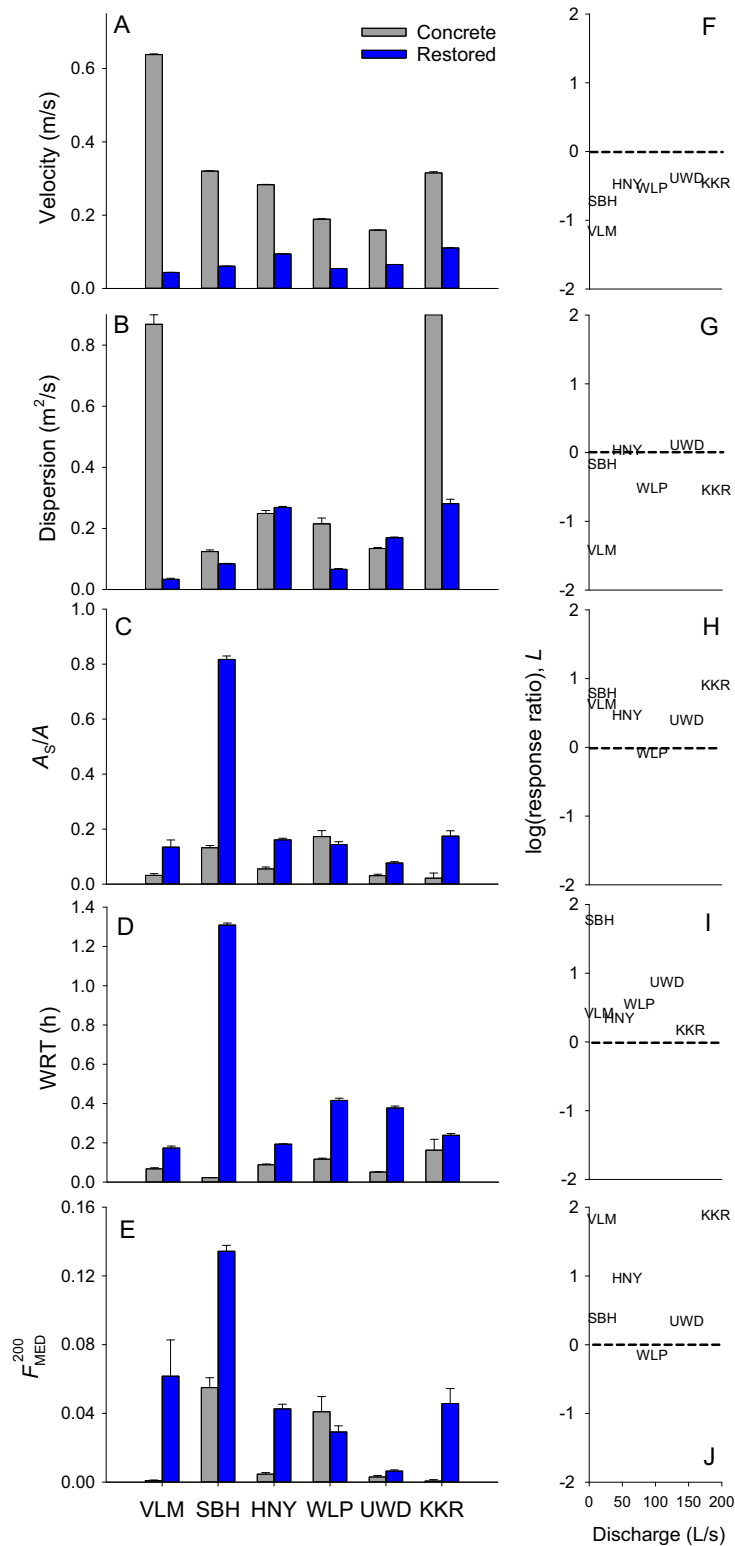
Fig. 1. Stream water (A) velocity and (B) dispersion and transient storage metrics, including (C) the ratio of storage zone to active channel area (A_S/A), (D) mean water residence time (WRT), and (E) the fraction of median transport time due to storage (F_{MED}^{200}; Runkel 2002), of the concrete and restored reaches ordered in terms of increasing stream size along the x-axis. Error bars represent SE generated during the modeling of these parameters. L ratios represent the effect of the restoration on (F) velocity, (G) dispersion, (H) A_S/A, (I) WRT, and (J) F_{MED}^{200}. See Table 1 for stream names.
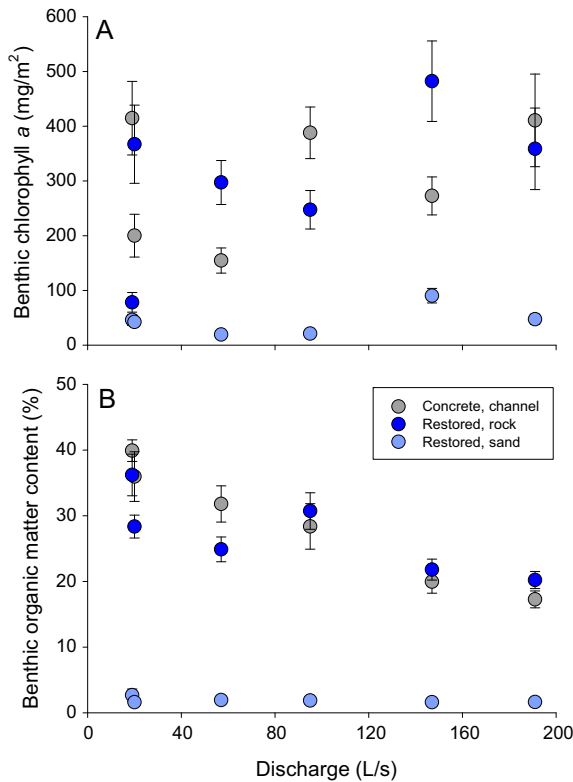
FIG. 2. Patterns of stream size (i.e., discharge) and biofilm (A) chlorophyll a and (B) organic matter content in the benthic habitats of the restored (rocks/gravel and sand) and concrete reaches of the six study streams.

generally increased with discharge, but the correlations were not statistically significant for any nutrient form.

The uptake velocity (i.e., v_f) for NH_4^+, NO_3^-, and SRP was 0.093 ± 0.017), 0.049 ± 0.009), and 0.10 ± 0.015 mm/s, respectively. As with S_W, v_f (e.g., relative demand) for SRP was significantly higher than for NO_3^-, while NH_4^+ was intermediate (Table 3; ANCOVA, $P_{NUTRIENT} = 0.02$). The variation in v_f among the 12 reaches was not associated with reach or discharge. In fact, the three highest v_f measurements for each nutrient were recorded in reaches of both types (i.e., restored, concrete) and spanned the discharge gradient (Table 3). Similarly, the L ratios were a mix of both positive and negative values for each nutrient (Fig. 3D–F) indicating that there was no consistent effect of the restoration on relative nutrient demand in the streams. The v_f L ratios were not correlated with discharge.

Nutrient uptake rates (U) were highly variable among nutrient compounds, reaches, and streams. The mean uptake rates for NH_4^+, NO_3^-, and SRP were 0.35 ± 0.12, 0.90 ± 0.30, and 0.35 ± 0.13 g·m^{-2}·d^{-1}, respectively (Table 3). Given the large variation in rates for each nutrient, uptake did not differ significantly between nutrient forms or reach. However, uptake rates were generally higher in the headwater and middle-sized streams relative to the largest streams (Table 3; ANCOVA,

$P_{DISCHARGE} = 0.07$). As we observed with v_f, the L ratios were scattered above and below zero demonstrating that U did not vary consistently across the paired stream reaches (Fig. 3G–I). However, the L ratios for NH_4^+ uptake were positively correlated with discharge (Fig. 3G; $r = 0.92$, $P = 0.01$), indicating that restored reaches removed more NH_4^+ relative to their paired concrete reach as stream size increased. A similar positive trend was apparent for NO_3^- and SRP uptake, but these correlations were not statistically significant.

Whole-stream metabolism exhibited clear differences by reach type and stream size. Gross primary production (GPP) was consistently higher in the restored reaches, and was higher in larger streams (ANCOVA, P_{REACH} and $P_{DISCHARGE} < 0.001$; Fig. 4A). Mean GPP rates ranged an order of magnitude across the 12 study reaches, from ~1.4 ± 0.1 g O_2·m^{-2}·d^{-1} in concrete channels of headwater streams (VLM and SBH) to ~17.4 ± 0.7 g O_2·m^{-2}·d^{-1} in restored reaches of large streams (UWD and KKR). L ratios for GPP were consistently positive (Fig. 4D; t test, $P < 0.001$), and the difference in GPP between the restored and concrete reaches declined as stream size increased (Pearson's $r = -0.44$, $P < 0.001$).

Ecosystem respiration (ER) was also higher in the restored reaches and in larger streams (Fig. 4B; ANCOVA, P_{REACH} and $P_{DISCHARGE} < 0.001$). Variation in ER among the reaches followed a similar pattern to that observed in GPP: concrete channels in headwater streams had the lowest rates while restored reaches of large streams had the highest (i.e., $-1.6 ± 0.1$ vs. $-11.0 ± 0.5$ g O_2·m^{-2}·d^{-1}, respectively). The L ratios of ER were all positive (t test, $P < 0.001$) and, like GPP, decreased with stream size (Fig. 4E; Pearson's $r = -0.28$, $P = 0.002$), demonstrating that ER varied the most between the two reaches in the headwater streams and less so in the larger streams.

The majority of reaches were autotrophic during the study (i.e., NEP > 0). NEP did not vary consistently with reach, but was higher in larger streams (Fig. 4C; ANCOVA, $P_{REACH} = 0.3$, $P_{DISCHARGE} < 0.001$). However, the effect of stream size was dependent on reach type (ANCOVA, $P_{REACH×DISCHARGE} < 0.001$). In the headwaters (e.g., VLM, SBH), NEP was negative (i.e., GPP < ER) and the concrete reaches tended to have higher NEP than the restored reaches. As stream size increased, NEP became ever more positive and tended to be higher in restored reaches than contiguous concrete reaches. As with GPP and ER, the difference in NEP between restored and concrete reaches was most stark in the headwater streams, then decreased as stream size increased (Fig. 4F; Pearson's $r = -0.48$, $P < 0.001$).

Interrelationships among physical, chemical, and biological metrics

The variation in hydrological metrics was often predicted by other physical characteristics of stream reaches.

TABLE 3. Nutrient uptake dynamics of the restored and concrete reaches.

Stream and reach type	Uptake length (S_W; m)			Uptake velocity (v_f; mm/s)			Uptake rate (U; mg·m^{-2}·d^{-1})		
	NH_4^+	NO_3^-	SRP	NH_4^+	NO_3^-	SRP	NH_4^+	NO_3^-	SRP
VLM									
Restored	64	354	109	0.094	0.017	0.055	920	718	30
Concrete	134	487	178	0.192	0.053	0.144	1,400	2,160	93
SBH									
Restored	382	479	73	0.019	0.017	0.108	83	81	1,340
Concrete	556	1,460	161	0.026	0.010	0.093	117	98	1,110
HNY									
Restored	111	85	57	0.065	0.068	0.101	212	3,410	505
Concrete	245	325	228	0.061	0.033	0.047	187	1,680	276
WLP									
Restored	75	165	58	0.164	0.069	0.195	314	153	367
Concrete	86	269	115	0.196	0.055	0.128	334	229	220
UWD									
Restored	474	268	197	0.076	0.127	0.172	171	1,220	74
Concrete	352	255	276	0.042	0.066	0.061	92	555	37
KKR									
Restored	254	635	452	0.093	0.037	0.052	256	319	37
Concrete	326	708	569	0.088	0.041	0.051	149	141	49

Among our measures of discharge, reach type (i.e., restored vs. concrete), and stream geomorphology (i.e., width : depth, slope, sinuosity), velocity, and dispersion were best explained by models containing slope and reach type (R^2 = 0.88 and 0.48, respectively; P < 0.001 and 0.02, respectively). Both velocity and dispersion increased as channel slope increased. The two most parsimonious models for the ratio of storage zone to active channel area (i.e., A_S/A) included discharge and reach type with (R^2 = 0.52, P = 0.03; Table 4) or without (R^2 = 0.42, P = 0.03) slope included; A_S/A decreased with discharge and slope. Water residence time (i.e., WRT) was predicted by several metrics of geomorphology and reach type (all R^2 = 0.50–0.53, P < 0.02), including positive correlations with dispersion, discharge, and the width-to-depth ratio and a negative relationship with slope and velocity (Table 4). The two predictive models for F_{MED}^{200} included dispersion and reach type (R^2 = 0.49, P = 0.02) or slope, discharge, and reach type (R^2 = 0.57, P = 0.02), where F_{MED}^{200} increased as dispersion, slope, and/or discharge decreased (Table 4).

We examined the patterns among S_W, v_f, and U using AIC_c models that contained physical, chemical, and biological variables. SRP S_W was best predicted by a model with discharge and reach type (R^2 = 0.61, P = 0.006), with higher S_W in streams with higher discharge. For SRP, U was predicted by A_S/A and reach type (R^2 = 0.41, P = 0.04), where uptake was higher in reaches with a larger A_S/A. Interestingly, no models predicted the variation in any nutrient dynamics for NH_4^+ and NO_3^- nor v_f for SRP.

Variation in whole-stream metabolism across study reaches was associated with only a few metrics among the suite of chemical and biological variables that we evaluated. The best model for GPP and NEP included positive effects of discharge and reach restoration (R^2 = 0.62 and 73, respectively, P ≤ 0.004; Table 4). ER was predicted by habitat-weighted OM and reach type (R^2 = 0.60, P = 0.007); ER increased with benthic OM content (Table 4).

DISCUSSION

Our study of paired restored and concrete reaches across a gradient in stream size demonstrates that channel restoration improves some key ecosystem metrics while leaving others unchanged. The strongest response was in stream hydrology, where the restoration of sinuous reaches to include pool-riffle sequences dramatically decreased velocity and increased dispersion and transient storage. In addition, nutrients traveled shorter distances in restored reaches as a result of the decreased water velocity and increased residence time. The shorter S_W may reduce downstream eutrophication by removing nutrients from circulation through biophysical uptake (Kaushal and Belt 2012, McMillan and Noe 2017). Ecosystem autotrophy and high rates of whole-stream metabolism in both concrete channels and restored reaches suggest continued impairment despite restoration efforts (Young et al. 2008). Other metrics, such as chlorophyll a and nutrient demand (i.e., v_f), did not differ consistently between concrete and restored reaches. Overall, the variability in responses to restoration across our diverse suite of metrics underscores the challenge of achieving effective and beneficial stream restoration in urban landscapes, yet also suggests some simple guidelines for restoration planning to maximize ecological and societal benefits of these costly interventions (Table 5).
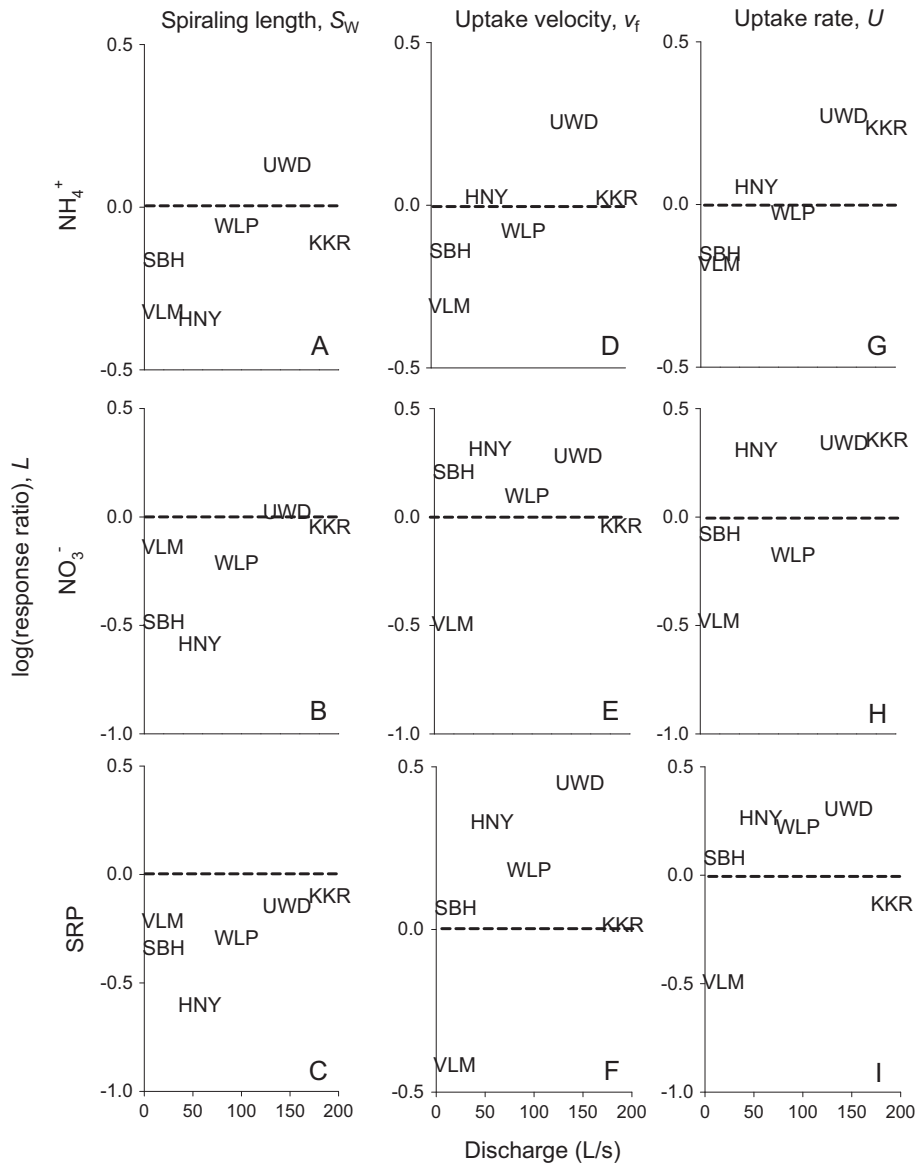
Fig. 3.　L ratios, which represent the magnitude of the restoration effect, of nutrient dynamics, including (A–C) uptake length (S_W), (D–F) uptake velocity (v_f), and (G–I) uptake rates (U) for ammonium (NH_4^+), nitrate (NO_3^-), and soluble reactive phosphorus (SRP). See Table 1 for stream names.

Scale-dependence of ecosystem responses to restoration

Our study demonstrates that restorations of small streams can substantially affect ecosystem processes, whereas restoration of mainstem rivers has more modest effects. For many of the metrics we quantified, the relative difference between the restored and concrete reaches was starkest in the headwater streams and dampened in the larger streams. In other words, the L ratios of one-half of the response metrics were farthest from zero in headwater streams and many approached zero as stream size increased. For example, water velocity and residence time were drastically different between reach types in headwaters, but differed only slightly in the mainstem

rivers. The effect of stream size may owe to geomorphic variation related to position in the river network; the restorations increased stream width several-fold in the headwaters, but the width of the restored reach was often comparable or lower to that of the concrete channel in the larger streams. Consequently, restorations have a more profound effect on channel residence time in headwaters, which can exert strong control on nutrient retention in urban river networks by governing the contact time between the water column and stream benthos (Kaushal and Belt 2012, Newcomer-Johnson et al. 2016). Furthermore, headwater streams showed the strongest response to restorations for certain functional measurements as well (e.g., S_W, GPP). The greater
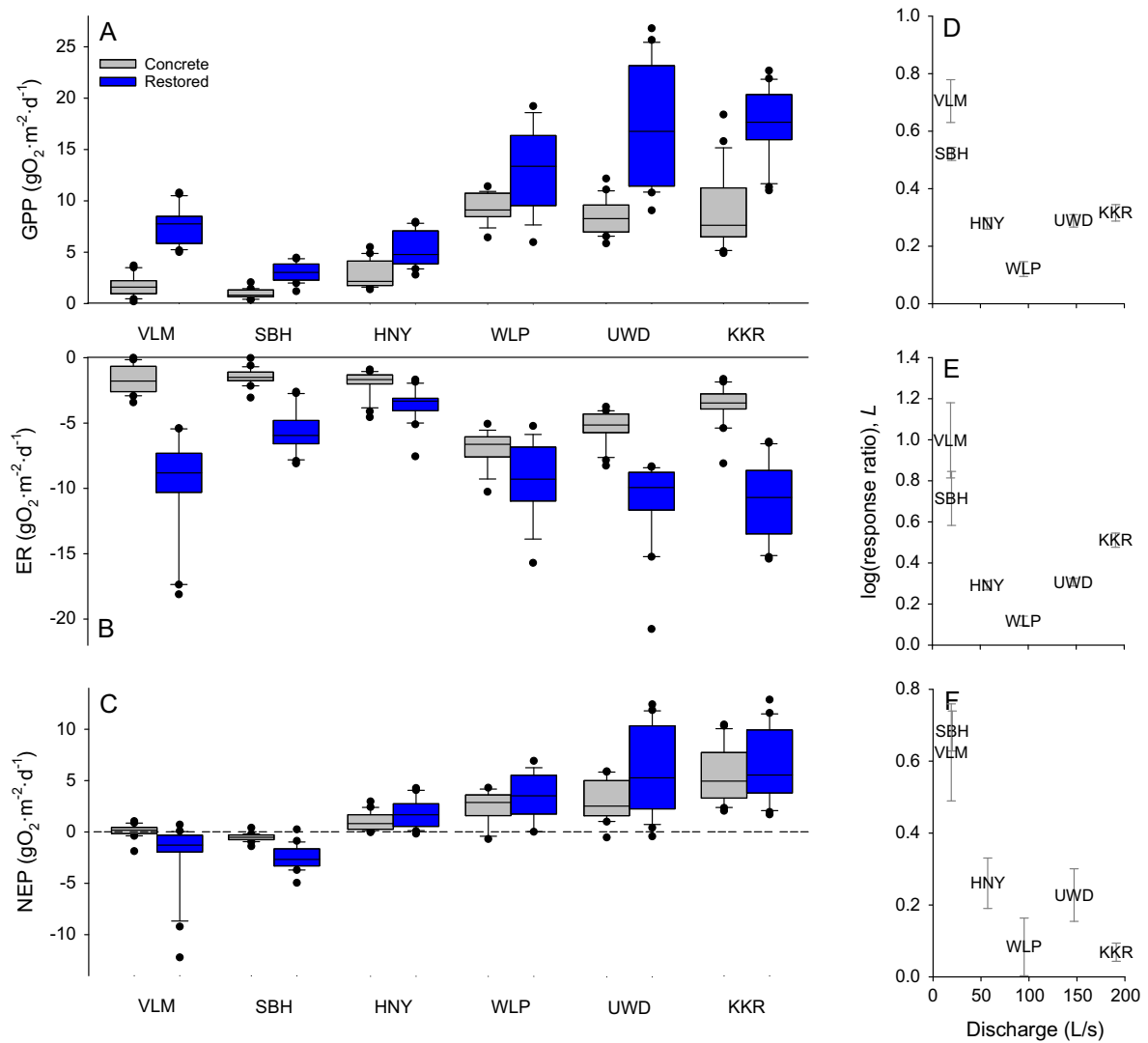
Fig. 4. Whole-stream metabolism metrics of the concrete and restored reaches, including (A) gross primary production (GPP), (B) ecosystem respiration (ER), and net (C) ecosystem production (NEP) during the study period (n = 18–21 d per reach; Data S1: Metabolism output). ER data presented as negative values to represent the consumption of oxygen during respiration. L ratios represent the effect of the restoration on (D) GPP, (E) ER, and (F) NEP.

impact in headwaters is likely due to network position, because the restored reach represents a greater proportion of stream length in the watershed than a similarly sized restoration of a mainstem river (Dodds and Oakes 2006). Improving the reach hydrology also led to shorter nutrient uptake lengths and increased rates of whole-stream metabolism, which suggests that these responses are predictable outcomes of restoring geomorphology to slow water passage. In contrast, L ratios of nutrient demand (v_f) and uptake (i.e., U) were unrelated to stream size, perhaps because high variation in nutrient concentrations and biological activity between the streams muted any network-scale patterns when comparing the L ratios (Hoellein et al. 2011). We did not observe many correlations between the L ratios and

baseflow discharge, demonstrating that the magnitude of the ecological responses to restoration did not scale directly with stream size.

The strong ecological responses to restorations that we observed may only be apparent in severely impaired stream ecosystems. Urban infrastructure, such as the concrete channels in our study, can alter the ecosystem function of streams in diverse ways. For example, comparisons between open and buried headwater streams have shown lower rates of GPP and ER and higher export of nutrients in buried channels where the urban infrastructure limits light availability and channel residence time (Beaulieu et al. 2014, Pennino et al. 2014). However, previous studies on restored streams that were less severely impaired prior to the restoration (e.g.,

TABLE 4. Mean (minimum, maximum)† whole-stream metabolism parameters estimated from the modeling exercise (GPP, ER, NEP, and K_{600}) and model fits (R^2).

Stream and reach type	n	Metabolism parameters (g $O_2 \cdot m^{-2} \cdot d^{-1}$)			K_{600} (d^{-1})	Model R^2
		GPP	ER	NEP		
VLM						
Restored	21	7.5 (5.0, 10.8)	−9.7 (−18.1, −5.4)	−2.1 (−12.2, 0.7)	2.6 (1.5, 5.7)	0.98 (0.94, 0.99)
Concrete	20	1.8 (0.2, 3.7)	−1.7 (−3.4, −0.01)	0.1 (−1.9, 1.0)	2.8 (0.5, 14.4)	0.97 (0.89, 0.99)
SBH						
Restored	21	3.1 (1.2, 4.5)	−5.6 (−1.9, −7.6)	−2.5 (−4.9, 0.3)	1.1 (0.6, 1.6)	0.95 (0.69, 0.99)
Concrete	21	1.0 (0.4, 2.1)	−1.5 (−3.0, −0.01)	−0.5 (−1.4, 0.4)	0.4 (0.2, 0.7)	0.98 (0.91, 0.99)
HNY						
Restored	21	5.4 (2.9, 8.0)	−3.6 (−7.6, −1.7)	1.8 (−0.1, 4.3)	2.9 (1.7, 5.5)	0.95 (0.88, 0.98)
Concrete	21	3.0 (1.4, 5.6)	−1.9 (−4.6, −0.9)	1.0 (0.02, 3.0)	1.2 (0.6, 4.0)	0.95 (0.91, 0.99)
WLP						
Restored	18	12.9 (6.0, 19.3)	−9.3 (−15.7, −5.3)	3.5 (0.1, 7.0)	2.7 (1.6, 5.8)	0.98 (0.91, 0.99)
Concrete	18	9.4 (6.5, 11.5)	−6.9 (−10.3, −5.1)	2.4 (−0.6, 4.4)	1.6 (1.0, 2.5)	0.96 (0.80, 0.99)
UWD						
Restored	21	17.1 (9.1, 26.8)	−10.9 (−20.7, −8.3)	6.2 (−0.3, 12.5)	3.9 (2.3, 7.8)	0.98 (0.96, 0.99)
Concrete	21	8.5 (5.9, 12.2)	−5.3 (−8.3, −3.7)	3.1 (−0.4, 6.0)	2.1 (1.5, 3.6)	0.97 (0.91, 0.99)
KKR						
Restored	21	17.6 (11.0, 22.7)	−11.0 (−15.4, −6.4)	6.6 (1.7, 12.9)	4.0 (3.0, 6.4)	0.99 (0.96, 0.99)
Concrete	21	8.9 (4.9, 18.4)	−3.5 (−8.1, −1.6)	5.4 (2.1, 10.5)	1.8 (1.1, 3.5)	0.98 (0.93, 0.99)

Note: Values are means with SE in parentheses.

†See Data S1: Metabolism output for all whole-stream metabolism data, including estimated upper and lower bounds for gross primary production (GPP), ecosystem respiration (ER), net ecosystem production (NEP), and the reaeration coefficient (K_{600}) on each day.

channelized, but not concrete lined) found no effect of the restoration on transient storage metrics or nutrient uptake dynamics (Bukaveckas 2007, Becker et al. 2013). Regarding whole-stream metabolism, prior work suggested no differences in GPP and ER between restored and non-restored urban streams, even when compared to nearby forested streams (Sudduth et al. 2011). However, the non-restored streams in those studies were not as severely degraded as concrete channels or buried streams. The lack of a response when comparing a set of restored streams vs. a set of unrestored streams reflects the difficulty of detecting patterns across inherently variable stream ecosystems. A strength of our study design was sensitivity to differences in transient storage and whole-stream metabolism between contiguous unrestored and restored reaches in the same stream, but we also recognize that the pre-restoration condition was severely impaired channels.

Generally, stream ecosystem characteristics and processes are intimately linked with watershed land use and vary longitudinally through a river network. An urban stream continuum concept has been proposed to establish a framework for identifying the common stressors and impairments within urban river networks (Kaushal and Belt 2012). In turn, the continuum concept can be used to generalize how and why patterns and processes change from headwaters to mainstem rivers in urban landscapes. Some of the longitudinal patterns we observed as stream size increased followed expectations (e.g., longer S_W; Hall et al. 2013) and previous findings (i.e., GPP; Vannote et al. 1980, Kaushal et al. 2014).

However, others failed to show longitudinal patterns reported elsewhere (i.e., N uptake; Claessens et al. 2010, Filoso and Palmer 2011). In our study, stream size (i.e., discharge) and geomorphology were the strongest predictors for nine of the response metrics among the six streams. However, no model of physical, chemical, and biological metrics predicted the other eight responses. Given the landscape heterogeneity and high connectedness of urban watersheds via impervious surfaces and storm drain networks (Lookingbill et al. 2009), the metrics we quantified that did show longitudinal patterns may be robust tools to evaluate stream ecosystems and further refine an urban continuum framework (e.g., GPP, S_W; Clapcott et al. 2010, Kaushal and Belt 2012). Stream size and geomorphology represent simple predictors to estimate the potential responsiveness of urban streams to restoration efforts.

Our study was conducted during low-flow conditions in the summer months to allow for strong intra- and inter-stream comparisons. However, the hyper-connectivity of the surrounding watershed to the streams via impervious surface and networks of storm sewers can lead to rapid changes in the hydrograph during precipitation events (Walsh et al. 2005a). Storm-flow events can have a significant impact on stream ecosystem function, changing downstream nutrient transport and depressing whole-stream metabolism (Filoso and Palmer 2011, Beaulieu et al. 2013). Therefore, the differences we report between the restored and concrete reaches would be less evident during storm-flow events when water velocity and volume are higher. In addition, the six restorations

TABLE 5. Predictors of ecosystem structure and function among the 12 study reaches.

Predictors and predictive models	R^2	P
Velocity		
$\ln(y) = 0.45(\ln(\text{Slope})) - 1.4(\text{Reach}) - 0.65$	0.88	<0.001
Dispersion		
$y = 0.80(\ln(\text{Slope})) - 0.83(\text{Reach}) - 0.13$	0.48	0.02
A_S/A		
$y = -0.54(\ln(\text{Slope})) - 0.0069(Q) + 1.1$ (Reach) $- 3.0$	0.52	0.03
$y = -0.0056(Q) + 1.2(\text{Reach}) - 2.4$	0.42	0.03
WRT		
$y = 0.20(\ln(\text{Dispersion})) + 1.8(\text{Reach}) - 2.4$	0.53	0.01
$y = 0.0018(Q) + 1.6(\text{Reach}) - 2.8$	0.51	0.02
$y = -0.21(\ln(\text{Velocity})) + 1.3(\text{Reach}) - 2.9$	0.50	0.02
$y = -0.12(\ln(\text{Slope})) + 1.6(\text{Reach}) - 2.8$	0.50	0.02
$y = 0.027(W/D) + 1.6(\text{Reach}) - 2.8$	0.50	0.02
F_{MED}^{200}		
$y = -1.0(\ln(\text{Dispersion})) + 1.0(\text{Reach}) - 6.5$	0.49	0.02
$y = -1.2(\ln(\text{Slope})) - 0.013(Q) + 1.8$ (Reach) $- 5.8$	0.57	0.02
SRP S_W		
$y = 1.8(Q) - 97(\text{Reach}) + 93$	0.61	0.006
SRP U		
$y = 1.2(\ln(A_S/A)) - 1.4(\text{Reach}) + 8.6$	0.41	0.04
$y = 1.1 \quad (\ln(A_S/A)) - 0.16(\text{ER}) - 0.48$ (Reach) $+ 8.8$	0.48	0.04
GPP		
$y = 0.061(Q) + 5.5(\text{Reach}) - 0.36$	0.62	0.004
$y = 0.049(Q) - 1.2(\ln(\text{SRP})) + 5.4$ (Reach) $+ 4.4$	0.63	0.01
ER		
$y = 0.035(\text{Reach-weighted} \quad \text{OM}) + 3.9$ (Reach) $+ 0.10$	0.60	0.007
NEP		
$y = 0.039(Q) + 0.79(\text{Reach}) - 1.6$	0.73	0.001

Notes: Reach was included in all models because the drastic differences in reach type was a major factor in our study design. All models for each parameter are consider equally likely using the AIC_c values (<2; Akaike information criterion corrected for sample size; Burnham and Anderson 2002). Parameters are A_S/A, ratio of storage zone to active channel area; WRT, water residence time; F_{MED}^{200}, the fraction of median transport time due to storage; SRP, soluble reactive phosphorus; S_W, uptake length; U, uptake velocity; W, width; D, depth; Q, discharge; GPP, gross primary production; ER, ecosystem respiration; NEP, net ecosystem production.

we studied varied in their age since implementation; three of the six streams were restored ~3 to 4 yr prior to our study, whereas two were restored several decades earlier. All of these restorations shared the same goal: to alter the channel geomorphology from a highly constrained or concrete channel to a more natural meandering state (U.S. Army Corps of Engineers 2016). Although we could not quantify the influence of time since restoration, the disturbance-driven dynamics of stream ecosystem processes and biological communities (Arango et al. 2015, Larsen and Harvey 2017) are expected to help streams achieve a new dynamic

equilibrium fairly soon after restoration. Overall, we consider it improbable that the variation among our study streams in age of restoration, order of restored and concrete reaches, or reach length contributed substantially to the patterns documented in this study. However, we cannot exclude that possibility for any of these factors, and we recognize that differences in the recovery of riparian vegetation, fish communities, and other important features following restoration could mediate the magnitude of responses that we documented. To minimize that risk, we measured ecological functions that are known to respond strongly to stream conditions (e.g., nutrient uptake, metabolism), and we evaluated the consistency of our results with earlier findings. Thus, the sharp biophysical differences between contiguous reaches coupled with the consistent effects of restoration across our six study streams strongly suggest that our inferences are robust.

Benchmarks for successful restoration of urban streams

The degraded state of urban watersheds presents major challenges for restoring the ecological structure and function of stream ecosystems. Our strongest finding is that restoring a mere 150–300 m of stream channel can alter many aspects of ecosystem functioning, yet it is equally clear that responses of other key metrics are inconsistent at best (see also Violin et al. 2011). Thus, judging whether stream ecosystems have been substantially "improved" by channel restoration is difficult. Our approach was rooted in objective comparisons of a broad suite of metrics in contiguous concrete and restored reaches, but interpreting changes in structure and function as a qualitative improvement in the ecosystem remains subjective. Nonetheless, such judgments play an important role in motivating expensive restoration efforts, so they merit careful consideration.

Our hydrological metrics present the clearest case of ecosystem improvement: concrete channels are a highly unnatural geomorphology that were designed specifically to hasten water conveyance, and we have shown that restoration lowers stream velocity and enhances dispersion and transient storage to produce more naturalistic hydrology. We regard this transformation as an unambiguous improvement in stream ecosystem functioning. These metrics have an additional virtue: articulating the benefits of slower water velocity and higher residence time to non-scientists is straightforward when discussing restoration goals. Thus, we conclude that hydrological metrics offer a compelling means of assessing the benefits of channel restoration in urban streams.

Among nutrient spiraling metrics, uptake length (S_W) was consistently shorter in the restored reaches than contiguous concrete reaches, which we interpret as improved ecosystem function resulting from achieving more naturalistic hydrology and geomorphology (see also McMillan and Noe 2017). Shorter nutrient uptake lengths in cross-stream comparisons can also reflect

differences in nutrient concentrations or other factors (Stream Solute Workshop 1990), and the variation in S_W among our study streams suggests the influence of such contingencies. However, contiguous restored and concrete reaches within each stream showed similar nutrient concentrations, suggesting that reach-scale differences in nutrient spiraling indeed indicate improved functioning of the stream ecosystem. Also, like hydrological metrics, the benefits of decreased transport distances for nutrient retention can be explained intuitively to practitioners.

Stream metabolism data present a more complicated context for interpreting restoration outcomes. Metabolism measurements integrate stream and watershed characteristics over a larger spatial scale than the restored reach alone, providing a holistic assessment of the condition of urban streams (Roley et al. 2014). Rates of GPP, ER, and NEP can be used as benchmarks for managers to assess the efficacy of restoration projects through before-after comparisons as well as abundant comparison data from the literature (Bernot et al. 2010, Marcarelli et al. 2011). We observed higher whole-stream metabolic rates than expected from previous reports (Marcarelli et al. 2011; Fig. 5); autotrophy (i.e., GPP > ER) was more common in the highly urbanized networks around Milwaukee than urban, agricultural, or reference reaches elsewhere (Bernot et al. 2010; Fig. 5). Moreover, restoration consistently enhanced both GPP and ER (Fig. 4). High GPP, especially when it exceeds ER, is often viewed as an indicator of stream, riparian, or watershed impairment (e.g., high nutrient loading, low canopy cover; Dodds and Oakes 2006, Young et al. 2008). We interpret increased metabolic rates in restored reaches as evidence of ecological impairment in this study; the more naturalistic substrates and hydrology

allowed the accumulation of algae and benthic OM. We believe this interpretation is reasonable given that whole-stream metabolism reflects the state of the watershed more so than a single stream reach, but differences in top-down control and other factors could also be involved. In addition to inferential challenges, quantifying ecosystem metabolism requires substantial technical expertise that may present a barrier to widespread application of these metrics to categorize stream health and restoration responses (Young et al. 2008). However, both instrument accessibility and understanding of ecosystem metabolism in urban streams are continuing to grow (Beaulieu et al. 2014, Pennino et al. 2014), so we are optimistic that standardized guidelines for interpretation will emerge in the context of both urban streams and responses to channel restoration. If so, metabolic data will offer a powerful tool for assessing responses to restoration in individual reaches and throughout the watershed.

Taken together, the diverse suite of responses we quantified elucidates both the strengths and challenges of using multiple metrics to assess restorations in urban watersheds. Functional metrics are desirable for assessing restorations because they summarize ecosystem dynamics through time (Feio et al. 2010, Roley et al. 2014), and are more responsive to urbanization than to other anthropogenic land uses (e.g., agriculture; Yates et al. 2014). Yet we found that some metrics were largely insensitive to restoration (e.g., v_f), while others shifted dramatically (e.g., water residence time, S_W; see also Clapcott et al. 2010), suggesting that an efficient assessment protocol could focus on only the most sensitive subset of metrics. Our findings also suggest that restoration efforts that extend beyond the boundaries of the focal reach may be essential for improving stream health (see also Walsh et al. 2005b, Bernhardt and Palmer 2011). In urban landscapes, interventions such as retention ponds, pervious pavement, rain barrels, and rain gardens may complement channel restoration efforts to yield synergistic responses of stream ecosystem functioning. This argues for systematic restoration planning, combining interventions of different sorts at multiple locations in order to maximize outcomes such as nutrient and sediment retention.

Our comparisons of sites in urban river networks also underscore the potential benefits of developing simple guidelines for strategic selection of sites for future restorations. First, headwater streams appear to respond more strongly than larger mainstem rivers (Swan and Brown 2017). Second, stream size and geomorphology predicted some hydrological and biological responses, providing proxies for setting expectations for restoration outcomes. Finally, our study demonstrates that a suite of low-cost functional metrics can provide a robust means of quantitatively evaluating the ecosystem-scale improvements resulting from urban stream restorations. Even if streams in densely populated areas are not ecologically salvageable due to constraints on restoration
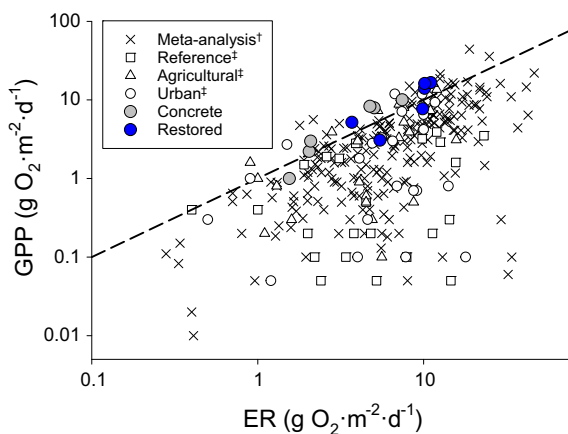

FIG. 5. Whole-stream metabolism data from restored and concrete reaches in urban streams from this study relative to results from syntheses of global streams († Marcarelli et al. 2011) and headwater streams in different land use types across the major biomes of North America (‡ Bernot et al. 2010). The dashed one-to-one line represents streams where GPP equals ER.

cost and spatial scale, a restoration designed around both societal benefits (e.g., minimizing flood risk, increasing recreational and aesthetic value) and ecological benefits (e.g., increasing transient storage, decreasing nutrient export) can encourage support and investment in a reach-scale restoration project (Bernhardt and Palmer 2007). Our results support calls to think beyond the stream channel to successfully restore the structure and function of urban stream ecosystems (Smith et al. 2016), but also demonstrates that restoring the geomorphology of stream channels, especially in headwater streams, is a worthwhile initial step.

Literature Cited

APHA. 2005. Standard methods for the examination of water and wastewater. American Public Health Association, Washington, D.C., USA.

Arango, C. P., P. W. James, and K. B. Hatch. 2015. Rapid ecosystem response to restoration in an urban stream. Hydrobiologia 749:197–211.

Beaulieu, J. J., C. P. Arango, D. A. Balz, and W. D. Shuster. 2013. Continuous monitoring reveals multiple controls on ecosystem metabolism in a suburban stream. Freshwater Biology 58:918–937.

Beaulieu, J. J., et al. 2014. Effects of urban stream burial on organic matter dynamics and reach scale nitrate retention. Biogeochemistry 121:107–126.

Becker, J. F., T. A. Endreny, and J. D. Robinson. 2013. Natural channel design impacts on reach-scale transient storage. Ecological Engineering 57:380–392.

Bernhardt, E. S., and M. A. Palmer. 2007. Restoring streams in an urbanizing world. Freshwater Biology 52:738–751.

Bernhardt, E. S., and M. A. Palmer. 2011. River restoration: the fuzzy logic of repairing reaches to reverse catchment scale degradation. Ecological Applications 21:1926–1931.

Bernhardt, E. S., et al. 2005. Synthesizing U.S. river restoration efforts. Science 308:636–637.

Bernot, M. J., et al. 2010. Inter-regional comparison of land-use effects on stream metabolism. Freshwater Biology 55:1874–1890.

Bisson, P. A., D. R. Montgomery, and J. M. Buffington. 2017. Valley segments, stream reaches, and channel units. Pages 21–47 in F. R. Hauer and G. A. Lamberti, editors. Methods in stream ecology. Academic Press, London, UK.

Bukaveckas, P. A.. 2007. Effects of channel restoration on water velocity, transient storage, and nutrient uptake in a channelized stream. Environmental Science & Technology 41:1570–1576.

Burnham, K. P., and D. R. Anderson 2002. Model selection and multi-model inference: a practical information-theoretical approach. Springer, New York, New York, USA.

Claessens, L., C. L. Tague, P. M. Groffman, and J. M. Melack. 2010. Longitudinal and seasonal variation of stream N uptake in an urbanizing watershed: effect of organic matter, stream size, transient storage and debris dams. Biogeochemistry 98:45–62.

Clapcott, J. E., R. G. Young, E. O. Goodwin, and J. R. Leathwick. 2010. Exploring the response of functional indicators of stream health to land-use gradients. Freshwater Biology 55:2181–2199.

Dodds, W. K., and R. M. Oakes. 2006. Controls on nutrients across a prairie stream watershed: land use and riparian cover effects. Environmental Management 37:634–646.

Feio, M. J., T. Alves, M. Boavida, A. Medeiros, and M. A. S. Graca. 2010. Functional indicators of stream health: a river-basin approach. Freshwater Biology 55:1050–1065.

Fellows, C. S., J. E. Clapcott, J. W. Udy, S. E. Bunn, B. D. Harch, M. J. Smith, and P. M. Davies. 2006. Benthic metabolism as an indicator of stream ecosystem health. Hydrobiologia 572:71–87.

Filoso, S., and M. A. Palmer. 2011. Assessing stream restoration effectiveness at reducing nitrogen export to downstream waters. Ecological Applications 21:1989–2006.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge, UK.

Gordon, N. D., T. A. McMahon, B. L. Finlayson, C. J. Gippel, and R. J. Nathan 2004. Stream hydrology: an introduction for ecologists. John Wiley and Sons Ltd, West Sussex, UK.

Hall, R. O., M. A. Baker, E. J. Rosi-Marshall, J. L. Tank, and J. D. Newbold. 2013. Solute-specific scaling of inorganic nitrogen and phosphorus uptake in streams. Biogeosciences 10:7323–7331.

Hedges, L. V., J. Gurevitch, and P. S. Curtis. 1999. The meta-analysis of response ratios in experimental ecology. Ecology 80:1150–1156.

Hill, W. A. 2017. Light. Pages 121–127 in F. R. Hauer and G. A. Lamberti, editors. Methods in stream ecology. Academic Press, London, UK.

Hoellein, T. J., C. P. Arango, and Y. Zak. 2011. Spatial variability in nutrient concentration and biofilm nutrient limitation in an urban watershed. Biogeochemistry 106:265–280.

Holtgrieve, G. W., D. E. Schindler, T. A. Branch, and Z. T.A'mar. 2010. Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a Bayesian statistical model of oxygen dynamics. Limnology and Oceanography 55:1047–1063.

Homer, C. G., J. A. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. D. Herold, J. D. Wickhma, and K. Megown. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States representing a decade of land cover change information. Photogrammetric Engineering and Remote Sensing 81:345–354.

Jarvie, H. P., D. R. Smith, L. R. Norton, F. K. Edwards, M. J. Bowes, S. M. King, P. Scarlett, S. Davies, R. M. Dils, and N. Bachiller-Jareno. 2018. Phosphorus and nitrogen limitation and impairment of headwater streams relative to rivers in Great Britain: a national perspective on eutrophication. Science of the Total Environment 621:849–862.

Kaushal, S. S., and K. Belt. 2012. The urban watershed continuum: evolving spatial and temporal dimensions. Urban Ecosystems 15:409–435.

Kaushal, S. S., P. M. Groffman, P. M. Mayer, E. Striz, and A. J. Gold. 2008. Effects of stream restoration on denitrification in

an urbanizing watershed. Ecological Applications 18:789–804.

Kaushal, S. S., K. Delaney-Newcomb, S. E. G. Findlay, T. A. Newcomer, S. Duan, M. J. Pennino, G. M. Sivirichi, A. M. Sides-Raley, M. R. Walbridge, and K. T. Belt. 2014. Longitudinal patterns in carbon and nitrogen fluxes and stream metabolism along an urban watershed continuum. Biogeochemistry 121:23–44.

Larsen, L. G., and J. W. Harvey. 2017. Disrupted carbon cycling in restored and unrestored urban streams: critical timescales and controls. Limnology and Oceanography 62:160–182.

Lookingbill, T. R., S. S. Kaushal, A. J. Elmore, R. Gardner, K. N. Eshleman, R. H. Hilderbrand, R. P. Morgan, W. R. Boynton, M. A. Palmer, and W. C. Dennison. 2009. Altered ecological flows blur boundaries in urbanizing watersheds. Ecology and Society 14.

Marcarelli, A. M., C. V. Baxter, M. M. Mineau, and R. O. Hall. 2011. Quantity and quality: unifying food web and ecosystem perspectives on the role of resource subsidies in freshwaters. Ecology 92:1215–1225.

McMillan, S. K., and G. B. Noe. 2017. Increasing floodplain connectivity through urban stream restoration increases nutrient and sediment retention. Ecological Engineering 108:284–295.

Mulholland, P. J., et al. 2002. Can uptake length in streams be determined by nutrient addition experiments? Results from an interbiome comparison study. Journal of the North American Benthological Society 21:544–560.

Newbold, J. D., J. W. Elwood, R. V. O'Neill, and W. V. Winkle. 1981. Measuring nutrient spiralling in streams. Canadian Journal of Fisheries and Aquatic Sciences 38:860–863.

Newcomer-Johnson, T., S. Kaushal, P. Mayer, R. Smith, and G. Sivirichi. 2016. Nutrient retention in restored streams and rivers: a global review and synthesis. Water 8:116.

Pennino, M. J., S. S. Kaushal, J. J. Beaulieu, P. M. Mayer, and C. P. Arango. 2014. Effects of urban stream burial on nitrogen uptake and ecosystem metabolism: implications for watershed nitrogen and carbon fluxes. Biogeochemistry 121:247–269.

Peterson, B. J., et al. 2001. Control of nitrogen export from watersheds by headwater streams. Science 292:86–90.

R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org

Roley, S. S., J. R. Griffiths, P. S. Levi, C. J. Patrick, S. Sadro, and J. P. Zarnetske 2014. State of the ecosystem: progress in quantifying ecosystem health in aquatic ecosystems. Pages 101–118 in Eco-DAS X Symposium Proceedings. Association for the Sciences of Limnology and Oceanography, Waco, TX.

Runkel, R. L. 1998. One dimensional transport with inflow and storage (OTIS): a solute transport model for streams and rivers. Water Resources Investigations Report. U.S. Geological Survey, Reston, Virginia, USA.

Runkel, R. L. 2002. A new metric for determining the importance of transient storage. Journal of the North American Benthological Society 21:529.

Runkel, R. L. 2015. On the use of rhodamine WT for the characterization of stream hydrodynamics and transient storage. Water Resources Research 51:6125–6142.

Smith, R. F., et al. 2016. Urban stream renovation: incorporating societal objectives to achieve ecological improvements. Freshwater Science 35:364–379.

Steinman, A. D., G. A. Lamberti, P. R. Leavitt, and D. G. Uzarski 2017. Biomass and pigments of benthic algae. Pages 223–242 in F. R. Hauer and G. A. Lamberti, editors. Methods in stream ecology. Academic Press, London, UK.

Stream Solute Workshop. 1990. Concepts and methods for assessing solute dynamics in stream ecosystems. Journal of the North American Benthological Society 9:95–119.

Sudduth, E. B., B. A. Hassett, P. Cada, and E. S. Bernhardt. 2011. Testing the field of dreams hypothesis: functional responses to urbanization and restoration in stream ecosystems. Ecological Applications 21:1972–1988.

Swan, C. M., and B. L. Brown. 2017. Metacommunity theory meets restoration: isolation may mediate how ecological communities respond to stream restoration. Ecological Applications 27:2209–2219.

U.S. Army Corps of Engineers. 2016. Corps of engineers. Underwood Creek Great Lakes fishery and ecosystem restoration project.Detroit District, U.S. Army Corps of Engineers, Wauwatosa, Wisconsin, USA.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130–137.

Violin, C. R., P. Cada, E. B. Sudduth, B. A. Hassett, D. L. Penrose, and E. S. Bernhardt. 2011. Effects of urbanization and urban stream restoration on the physical and biological structure of stream ecosystems. Ecological Applications 21:1932–1949.

Walsh, C. J., A. H. Roy, J. W. Feminella, P. D. Cottingham, P. M. Groffman, and R. P. Morgan. 2005a. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24:706.

Walsh, C. J., T. D. Fletcher, and A. R. Ladson. 2005b. Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream. Journal of the North American Benthological Society 24:690–705.

Yates, A. G., R. B. Brua, J. M. Culp, P. A. Chambers, and L. I. Wassenaar. 2014. Sensitivity of structural and functional indicators depends on type and resolution of anthropogenic activities. Ecological Indicators 45:274–284.

Young, R. G., C. D. Matthaei, and C. R. Townsend. 2008. Organic matter breakdown and ecosystem metabolism: functional indicators for assessing river ecosystem health. Journal of the North American Benthological Society 27:605–625.

SUPPORTING INFORMATION

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/eap.2107/full