MOLECULAR ECOLOGY

Molecular Ecology (2016) doi: 10.1111/mec.13784

Spatial and successional dynamics of microbial biofilm communities in a grassland stream ecosystem

ALLISON M. VEACH,*† JAMES C. STEGEN,‡ SHAWN P. BROWN,*§ WALTER K. DODDS* and ARI JUMPPONEN*

*Division of Biology, Kansas State University, Manhattan, KS 66502, USA, †Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA, ‡Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA, §Department of Plant Biology, University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA

Abstract

Biofilms represent a metabolically active and structurally complex component of freshwater ecosystems. Ephemeral prairie streams are hydrologically harsh and prone to frequent perturbation. Elucidating both functional and structural community changes over time within prairie streams provides a general understanding of microbial responses to environmental disturbance. We examined microbial succession of biofilm communities at three sites in a third-order stream at Konza Prairie over a 2- to 64-day period. Microbial abundance (bacterial abundance, chlorophyll *a* concentrations) increased and never plateaued during the experiment. Net primary productivity (net balance of oxygen consumption and production) of the developing biofilms did not differ statistically from zero until 64 days suggesting a balance of the use of autochthonous and allochthonous energy sources until late succession. Bacterial communities (MiSeq analyses of the V4 region of 16S rRNA) established quickly. Bacterial richness, diversity and evenness were high after 2 days and increased over time. Several dominant bacterial phyla (Beta-, Alphaproteobacteria, Bacteroidetes, Gemmatimonadetes, Acidobacteria, Chloroflexi) and genera (*Luteolibacter*, *Flavobacterium*, *Gemmatimonas*, *Hydrogenophaga*) differed in relative abundance over space and time. Bacterial community composition differed across both space and successional time. Pairwise comparisons of phylogenetic turnover in bacterial community composition indicated that early-stage succession (≤16 days) was driven by stochastic processes, whereas later stages were driven by deterministic selection regardless of site. Our data suggest that microbial biofilms predictably develop both functionally and structurally indicating distinct successional trajectories of bacterial communities in this ecosystem.

Keywords: algae, bacteria, community ecology, DNA barcoding, microbial biology

Received 19 November 2014; revision received 16 July 2016; accepted 18 July 2016

Introduction

Most bacteria exist attached to surfaces and survive in complex, multispecies communities. Biofilms form by initial adhesion of bacteria to substrata and subsequently grow into an interdependent, matrix-enclosed system (Davey & O'Toole 2000). In many stream ecosystems, which experience flow variability seasonally, biofilms are commonly exposed to disturbance events. Microbial cells, particulates and nutrients enter streams from the surrounding landscape, groundwater inputs and in subsurface flow (Veach *et al.* 2015; Battin *et al.* 2016), potentially interacting (both abiotic and biotic) with streambeds, thereby potentially influencing biofilm function and structure. In addition, stream biofilms have a greater biomass of respiring bacteria relative to flowing waters (Araya *et al.* 2003), and thus biofilm microbial communities represent a highly metabolically active component of freshwater ecosystems and provide

Correspondence: Allison M. Veach, Fax: (865) 574 6442; E-mail: veacham@ornl.gov

an ecologically relevant system to study microbial succession.

Biogeochemical cycling in freshwater is strongly influenced by biofilm communities and their developmental stage (Battin *et al.* 2003, 2016). Further, biogeochemical cycling controlled by stream biofilms can consequently influence nutrient transport (Mulholland *et al.* 2008) and be a significant source of nitrous oxide globally (Beaulieu *et al.* 2011). Grassland and wooded grassland streams drain approximately one-fourth of the world's land area and one-fifth of total continental run-off originate from them (Dodds 1997). Therefore, understanding the compositional and functional dynamics of microbial communities in general, and grassland streams specifically, is necessary to predict responses of nutrient cycles to global change (Wrona *et al.* 2006).

Successional ecology aims to characterize community assembly over time either after an initial colonization or following a disturbance and has provided mechanistic insights into community development across taxa, especially for plants and animals (e.g. Cowles 1899; Clements 1916; Gleason 1926; Connell & Slatyer 1977; Sousa 1979). Succession occurs as species abundances change over time via deterministic (i.e. niche-based) processes, such as selection through species interactions or environmental filtering, or stochastic (i.e. random) processes, such as ecological drift or random dispersal and immigration (Hubbell 2001; Chase & Myers 2011). Although successional ecology has been well characterized for macro-organisms, far less is known about temporal dynamics of bacterial community assembly (Fierer *et al.* 2010) especially in flowing waters (but see Jackson *et al.* 2001; Lyautey *et al.* 2005; Besemer *et al.* 2007; Li *et al.* 2015). Other work has found that bacterial communities exhibit successional dynamics highly influenced by deterministic processes such as physical conditions (Lyautey *et al.* 2005; Nemergut *et al.* 2007; Li *et al.* 2015), or biotic interactions (algal–bacterial associations; Besemer *et al.* 2007). Alternatively, bacterial communities may be controlled by stochasticity (i.e. random speciation or extinction events, dispersal or drift) during certain stages of succession (Zhou *et al.* 2014; Dini-Andreote *et al.* 2015). The relative importance of deterministic and stochastic processes during succession depends not only on species arrival and timing of substantial biofilm growth (biotic effects), but also on environmental fluctuations (e.g. nutrients, temperature) over time (Dini-Andreote *et al.* 2015).

In this study, we characterized primary successional dynamics of biofilm-associated microbial communities in a native prairie stream by measuring both compositional (algal and bacterial biomass, bacterial taxa relative abundance, community composition) and functional (biofilm net primary productivity (NPP)) aspects of microbial succession. Microbial communities subjected to physical and chemical disturbances in this ecosystem rapidly acclimate functional attributes to match local environments (<6 days; Kemp & Dodds 2002). Ephemeral prairie streams, which experience long periods of little to no flow, but very flashy, high discharge flood events on an annual basis (Dodds *et al.* 2004; Costigan *et al.* 2015) may have microbiota adapted to disturbance and be particularly important ecosystems to study microbial community successional dynamics. First, we hypothesized that microbial abundance (algae, bacteria) would increase over time, but peak ca. 30 days (indicative of late-stage succession) as seen in previous metabolic studies in prairie streams (Dodds *et al.* 1996). Second, biofilm NPP would initially be net heterotrophic (greater dissolved oxygen (DO) consumption than production) because of low abundances and biomass of autotrophs relative to heterotrophs, but would shift to net autotrophy as a result of increasing autotrophic biomass. Third, we hypothesized that early-stage bacterial communities would be composed of a few pioneer species capable of colonizing substrata, and these communities would be amended with additional species resulting in a gradual increase in species diversity over time. Lastly, as other studies have found that specific bacterial groups may be adapted for early biofilm formation and succession (e.g. Proteobacteria; Dang & Lovell 2000; Lyautey *et al.* 2005), we hypothesized that although spatial differences will be apparent, biofilm bacterial communities will be primarily influenced by stochastic community assembly processes during colonization and early biofilm formation, and selection will be the predominant driver of later microbial community succession due to environmental filtering and/or species interactions. Community composition across sites will be variable during early succession, but over time this site variability will decrease. During late succession, communities will converge across sites, due to deterministic processes, resulting in similar community compositions.

Materials and methods

Study location

We used the main reach of Kings Creek at Konza Prairie Biological Station located in the Flint Hills of northeastern Kansas in this study. Unglazed ceramic tiles ($N = 200$ per site to allow for random sampling; 4.8 × 2.8 cm) were autoclaved and then adhered to large bricks (40.6 × 40.6 cm) using aquarium silicon and submerged in three separate pools (Fig. S1, Supporting information). Tiles were placed within each of three stream locations on 5 April 2013 and removed 2,

4, 8, 16, 35 and 64 days later. These pools were disconnected from one another at the beginning of the experiment (zero surface flow), but became connected after a 38.4 mm rainfall event at 13 day post-tile placement (Fig. S2, Supporting information).

Physicochemical variables

Water chemistry (NO_3–N, NH_4–N), water turbidity (NTU) and water temperature data were collected at each site during each sampling period. Water was collected in each pool and filtered (GF/F, 2 μm size, GE Healthcare Companies) into acid-washed bottles and kept at −20 °C until thawed for inorganic N chemical analysis. Water column NO_3–N and NH_4–N analyses were performed by segmented flow analysis on an OI Analytical Flow Solution IV instrument using colorimetric techniques. Water turbidity was measured *in situ* using a YSI sonde (YSI 6136, Yellow Springs, OH, USA) which was calibrated each sampling time prior to measurement. Water temperature data were also collected using this instrument. Due to probe malfunction, turbidity data were not collected on day 4 (Table S1, Supporting information).

Microbial abundance and biofilm metabolism

We randomly sampled tiles to estimate algal biomass, bacterial cell abundance and biofilm metabolism (NPP). From each of the three pools and at each time (2, 4, 8, 16, 35, 64 days postplacement), (i) three tiles were collected for algal biomass, placed in Whirl-Pak bags (Nasco International, Fort Atkinson, WI, USA) and stored at −20 °C; (ii) three tiles were collected for bacterial cell abundance, placed in a nuclease-free 50-mL centrifuge tube, preserved in a 3% formalin solution and stored at 4 °C; and (iii) two tiles were collected into a 50-mL centrifuge tube containing stream water from each respective sampling pool and kept upright (to prevent biofilm disturbance) to measure NPP upon arrival in the laboratory. Day 2 samples did not have adequate biomass to obtain NPP measurements.

Algal biomass tiles analysed for chlorophyll *a* were placed in a 95% ethanol: H_2O solution, heated at 78 °C for 5 min and kept at 4 °C for ~12 h (Sartory & Grobbelaar 1984). Extract solution was analysed using a spectrophotometer (Hitachi High Technologies America, Inc., Schaumburg, IL, USA) according to standard methods (APHA 1995) and corrected for tile surface area. Preserved biofilm subsamples were incubated with 4′,6-diamidino-2-phenylindole (DAPI) nucleic acid stain (5 mg/mL) for 5 min and filtered (Whatman Nuclepore, 0.2 μm membrane size; GE Healthcare Companies). For each sample, the number of bacteria, as determined by cellular morphology (fungal hyphae and algal cells were noted, but not included in this count), on the membranes was estimated by counting 10–15 optical fields under an epifluorescent microscope (Nikon Labophot-2; Nikon Corporation, Tokyo, Japan). We acknowledge this may exclude larger cells (e.g. cyanobacteria) in our estimates, but is still likely a good representation of changes in bacterial abundance. Replicates for both bacterial abundance and chlorophyll *a* were averaged to obtain one value for each site across time points for statistical analyses.

We estimated NPP by placing two tiles into a closed, circulating chamber containing stream water from the location where the tiles were collected. The chamber was constructed using clear, acrylic plastic (US Plastics, Lima, OH, USA) which allows for ~92% transmittance of photosynthetically available radiation. A logging membrane oxygen probe (YSI 600-XLM, Yellow Springs, OH, USA) was placed horizontally in the chamber. A fluorescent, full-spectrum light (20-watt mini-compact bulb; Central Aquatics, Franklin, WI, USA) was placed over the chamber to mimic daylight conditions (~140 μmol quanta/m^2/s). Both temperature and DO were measured every 5 min for 20–25 min to obtain NPP rates. Subsequently, the chamber was placed in the dark to measure community respiration (CR) rates as above. The slope of DO concentration change over time in light and dark incubations was used to calculate metabolic rates (NPP, CR), whereas gross primary productivity (GPP) is NPP-CR, as in Bott (1996).

DNA extractions and Illumina MiSeq analysis

Total genomic DNA for three tiles from each site at each sampling time was extracted using a Qiagen DNeasy Plant Maxi kit (Qiagen, Venlo, the Netherlands) using the manufacturer's protocol with the following modifications: tiles were sonicated (FS20; Fisher Scientific, Pittsburgh, PA, USA) in the cell lysis solution until biomass was removed from the tiles. Two tiles did not yield enough DNA extract (Day 16 at two sites) and were omitted ($N = 52$). Extracted DNA was quantified using a NanoDrop 1000 spectrophotometer (NanoDrop Products, Wilmington, DE, USA). Template DNA was aliquoted into a 96-well plate at a concentration of 2 ng/μL.

We used a two-step PCR approach (see Berry *et al.* 2011) to avoid a 3′-end amplification bias generated with DNA tags. This approach was used to prevent potential barcode-specific PCR biases (Berry *et al.* 2011). In the first PCR step, the 16S rRNA gene V4 region was amplified using the 515F and 806R primers (Caporaso *et al.* 2012). Each sample was amplified in three independent 50 μL PCRs. Each reaction consisted of 2 μM of

forward and reverse primers, 10 ng of template DNA, 25 μL Phusion High-Fidelity Master Mix (New England Biolabs, Inc., Ipswich, MA, USA) and 10 μL of molecular grade water. Thermal cycler parameters (Eppendorf, Hamburg, Germany) included 5 min denaturation at 94 °C, followed by 25 cycles with denaturation at 94 °C for 1 min, annealing for 30 s. at 50 °C, extension for 1 min at 72 °C, with final extension for 10 min. Negative controls were included in PCRs to detect contamination and all controls remained contaminant free. For the secondary PCR, 10 μL of primary PCR products was amplified as above with the exception of using only five cycles and the inclusion of a reverse primer joined with 12-bp unique molecular identifier tags (MID-806R; Caporaso *et al.* 2012; Table S2, Supporting information). All technical replicates for both primary and secondary PCRs were visualized on a 1.5% agarose (w/v) gel to check for amplification. After secondary PCR visualization, the remaining PCR volume was pooled per experimental unit and cleaned with Agencourt AMPure (Beckman Coulter Inc., Pasadena, CA, USA) as per manufacturer's instructions except that we used a 1:1 ratio of AMPure bead solution to PCR volume to further discriminate against short PCR fragments. Each experimental unit was quantified for DNA yield and pooled at equal molarity (115 ng per sample). Amplicons were submitted to the Integrated Genomics Facility at Kansas State University (Manhattan, KS, USA), where Illumina-specific primers and adapters were ligated into amplicons using a NEBNext® DNA MasterMix for Illumina kit (Protocol E6040; New England Biolabs Inc.) and sequenced using a MiSeq Reagent Kit v2 (Illumina, San Diego, CA, USA) with 500 (2 × 250) cycles.

Bioinformatics

Sequences (.fastq) were processed using mothur (version 1.32.1; Schloss *et al.* 2009). Paired-end.fastq files were contiged and sequences with ambiguous bases, with >2 mismatches to the primers, 1 mismatch to the MID and homopolymeric regions >8 were removed. Remaining sequences were aligned against a mothur-implemented SILVA reference. Likely sequencer-generated errors were screened using a pseudo-single linkage clustering (pre.cluster with diff = 2; see Huse *et al.* 2010). Remaining sequences were screened for chimeric properties with the mothur-implemented UCHIME algorithm (Edgar 2010). After these steps, the median read length was at 253 bp. Sequences were assigned to taxonomic affinities using the naïve Bayesian classifier (Wang *et al.* 2007) with a bootstrap threshold of 80% against the RDP training set, version 9. Sequences not assigned to Domain Bacteria (including Archaea, mitochondria and chloroplasts) were omitted. A pairwise sequence distance matrix was calculated (extended gaps not penalized) and sequences were clustered to operational taxonomic units (OTUs) at a 97% similarity threshold using an average neighbour-joining method. Rare OTUs (abundance < 10 across all experimental units) were removed and taxonomic affinities were assigned to clustered OTUs. We rarefied at 18 000 sequences per sample to minimize loss of samples. The final rarefied data set with no 'rare' OTUs had 837 657 sequences and 3835 OTUs in total.

Statistical analyses

Multiple regression models were used to determine whether microbial abundance and biofilm metabolism (NPP, GPP, CR) differed over time and across sites. We visualized data distributions over time to assess curvature, as it is likely over successional time that biomass, production and respiration, or diversity and relative abundance of taxa may plateau and not necessarily follow linear trends over time. Data were analysed using two regression models: one with time and site as predictor variables and another with an additional quadratic term for time ($time^2$). The model with significantly lower sum of square errors was chosen based on an ANOVA table of both regression models. Microbial abundance data and time were $\log_{10}$-transformed for microbial abundance models to meet model normality assumptions. Biofilm metabolism regression models had the y intercept at zero as it is logical to assume metabolism rates equal to zero at the onset of the experiment. An outlier was detected (high Cook's D value; Cook 1979), for a 4-day measurement taken for GPP – likely indicating probe malfunction – and was removed.

For bacterial communities, Good's coverage (the complement of the ratio of local OTU singletons to total number of sequences) was calculated to determine how well each sample represented the resident bacterial communities. Observed OTU richness (S_{obs}), the complement of Simpson's diversity ($1 - D$: $1 - \Sigma p_i^2$) and Simpson's evenness (E_D: $1/\Sigma p_i^2/S$), with p_i representing frequency of each OTU within a sample, were also calculated for each site across each sampling time after iteratively subsampling at a depth of 18 000 sequences per experimental unit to reduce biases when comparing diversity estimates (general term throughout remaining text referring to richness, diversity and evenness) across samples which vary in library size (Gihring *et al.* 2012). A multiple linear regression model was used to determine whether observed bacterial OTU richness, diversity and evenness differed over time and across sites.

Multiple regression models were also used to determine whether the relative abundance of dominant

bacterial phyla and genera (represented ≥ 1% sequences) differed over time and across sites. Compositional differences among bacterial communities across time and sites were determined by computing Bray–Curtis dissimilarities and visualized using nonmetric multidimensional scaling (NMDS). A multiple linear regression model was then used to test whether NMDS axis scores differed over time and across sites. Further, a permutational multivariate ANOVA (PERMANOVA; Anderson 2001) with 1000 permutations was used to determine sources of variation in community composition across time and site after converting sequence abundance data to a Bray–Curtis distance matrix.

A stepwise regression model selection by Akaike's information criterion minimization approach was used to determine which physicochemical variables were correlated with microbial abundance, biofilm metabolism, diversity estimates and relative abundance of dominant phyla and genera. We also used environmental vector fitting of these physicochemical variables with bacterial community composition after 999 permutations. These analyses were used to understand whether these response variables varied with environmental conditions measured spatiotemporally.

Phylogenetic turnover in community composition framework

Community composition varies spatiotemporally due to stochastic or deterministic processes or their combination (Chase & Myers 2011). Based on a recent synthesis (Vellend 2010), we can categorize four main forces in community assembly: selection, drift, speciation and dispersal. Selection is a deterministic process which results from environmental filtering and/or biotic interactions, whereas the other processes are stochastic operating by themselves or in conjunction with selection. Here, we use a statistical framework (Stegen *et al.* 2013) which quantifies the influence of selection (deterministic process) vs. stochastic processes on bacterial community assembly over primary succession. We calculated β-mean nearest taxon distance (βMNTD; Fine & Kembel 2011; Webb *et al.* 2011) as the phylogenetic distance between each OTU in a community (k) and its closest relative in a second community (m):

$$\beta\text{MNTD} = 0.5\left[\sum_{i_k=1}^{n_k} f_{i_k} \min\left(\Delta_{i_k j_m}\right) + \sum_{i_m=1}^{n_m} f_{i_m} \min\left(\Delta_{i_m j_k}\right)\right],$$

where f_{i_k} is the relative abundance of each OTU i in the community k, n_k is the number of OTUs in community k and $\min(\Delta_{i_k j_m})$ is the minimum phylogenetic distance between i in k and all OTUs j in community m. βMNTD values can be less than, equal to or greater than the amount of phylogenetic turnover expected when selection does not affect community composition. Values less than expected phylogenetic turnover occur due to selection of species via similar environmental conditions or biotic interactions resulting in community composition to be similar to each other, whereas estimates greater than the expected turnover are due to differing environments constraining communities to be distinct from one another (species-sorting mechanisms, Chase & Myers 2011). Assessment of the degree of deviation from a null expectation was calculated by a randomization test that shuffled species identity and relative abundance across tips of the phylogeny. After this procedure, a null value βMNTD was calculated and repeated 999 times. Null values of βMNTD were generated using a regional species pool derived from all OTUs present over time and sites. The difference between observed βMNTD and the mean of the null distribution (after 999 randomizations) was measured in standard deviation units. This estimate is termed the β-nearest taxon index (βNTI). Values <-2 or >2 indicate that observed βMNTD differs by more than 2 standard deviations from a null expectation and indicate that communities are driven by selection, not stochastic processes. To assess the influence of stochasticity vs. determinism in influencing microbial community succession, we compared βNTI of the earliest successional stage communities (day 2) with all other time points within a site. Subsequently, a multiple linear regression was used to determine whether βNTI comparisons with day 2 communities differed over time and sites. This allowed the detection of both temporal changes in phylogenetic turnover and the degree of divergence in successional trajectories across sites. These analyses assume (i) there is no dispersal limitation among communities across space and time (all communities are assembled from the same regional species pool) and (ii) that phylogenetic distance accurately reflects ecological distance between OTUs (i.e. phylogenetic signal, Stegen *et al.* 2012, 2013). The first assumption is reasonable for communities sampled in this ecosystem due to each site being at least marginally connected and within 500 m of each other in flowing waters (site 1 and 2 are downstream of site 3, Fig. S1, Supporting information). The second assumption could not be effectively tested with this data set as we did not sample across an environmental gradient, and thus could not calculate differences in environmental optima among OTUs. However, other studies have found significant phylogenetic signal in bacterial OTU ecological niches across multiple ecosystems (subsurface, stream biofilms, salt marsh sediments, lake water, sediment and soil; Wang *et al.* 2013; Dini-Andreote *et al.* 2015) so we assume these communities follow a similar trend. The phylogeny was estimated using QIIME (Caporaso *et al.*

2010) implemented in FASTTREE program (Price *et al.* 2009) and all null model analyses were performed in R (R Development Core Team 2015).

All richness and diversity estimates were calculated in mothur (version 1.32.1, Schloss *et al.* 2009). All regression analyses were performed using the STATS package, βMNTD calculations in the PICANTE package (Kembel *et al.* 2010), stepwise model selection in the MASS package (Venables & Ripley 2002) and permutational multivariate ANOVA and NMDS using the VEGAN package (Oksanen *et al.* 2016) in R (version 2.13.1, R Development Core Team 2015).

Results

Microbial abundance and biofilm metabolism

No physicochemical variable differed over time or across sites (Table S3, Supporting information). Chlorophyll *a* increased linearly over time (Table 1, Fig. 1), whereas bacterial abundance increased linearly until late-stage succession where it began to slow (Table 1, Fig. 1). Early succession (2 days of incubation) was characterized by low microbial abundance (chlorophyll: 0.25 μg/cm^2; bacteria: 7.5×10^5 cells/cm^2), whereas later during succession (64 days), abundance was more than an order of magnitude greater (chlorophyll: 4.77 μg/cm^2; bacteria: 2.71×10^6 cells/cm^2). Chlorophyll concentrations differed across sites and had a time and site interaction (Table 1, Fig. 1). Chlorophyll was positively correlated with temperature and NO_3–N (T = 6.32, 2.44), and negatively correlated with NH_4–N and turbidity (T = −2.60, −5.31, $F_{4,9}$ = 28.48, Adj. R^2 = 0.89, $P < 0.01$). Bacterial abundance was also positively correlated with temperature (T = 2.50) and negatively correlated with turbidity (T = −4.63, $F_{2,11}$ = 18.37, Adj. R^2 = 0.73, $P < 0.01$).

All biofilm metabolism rates increased over time except CR. GPP and CR rates differed across sites (Table 1, Fig. 2). CR rates were greatest at 35 day (average: -1.92×10^{-4} mg DO/cm^2/min), whereas GPP rates increased linearly over time and late succession (64 days, average: 3.71×10^{-4} mg DO/cm^2/min) having rates more than twice those of early stages (4 days, average: 1.6×10^{-4} mg DO/cm^2/min). A one-sample *t*-test for each time point indicated that NPP did not differ from zero until 64 days (T = 7.5, d.f. = 2, P = 0.017). These data indicate that during primary succession, relative rates of DO production and consumption are similar until the late stage when communities become net autotrophic. Both GPP and NPP were positively correlated with temperature (T = 7.28, 5.32) and turbidity (T = 4.66, 2.76, $F_{2,8}$ = 26.82, 14.14, Adj. R^2 = 0.84, 0.72, respectively, $P < 0.01$).

Bacterial community diversity and composition

Bacterial communities established rapidly on submerged tiles. Bacterial richness, diversity and evenness increased linearly over time ($P < 0.04$, Table 2, Fig. 3). These response variables also differed between sites (Table 2, Fig. 3). Good's estimate of coverage did differ between sites (Site 2, T = −4.27, $P < 0.01$), but was, on average, 0.88 ± 0.02 meaning communities are undersampled even at 18 000 subsampled sequences per experimental unit. Rarefaction curves confirmed this: no plateau was reached at 18 000 sequences per sample (Fig. S3, Supporting information). Therefore, our conclusions on bacterial diversity for this study may be limited. No diversity estimate, or Good's coverage, was correlated with any physicochemical variable ($P \geq 0.70$).

Across all sampling times, dominant phyla (or classes of Proteobacteria) (≥1% of total abundance) were Betaproteobacteria (22.2% sequences), Bacteroidetes (21.8%), Verrucomicrobia (17.7%), Alphaproteobacteria (7.2%), Gammaproteobacteria (6.1%), Gemmatimonadetes (3.0%), Deltaproteobacteria (2.9%), Acidobacteria (1.1%) and Chloroflexi (1.0%). Dominant families consisted of the Verrucomicrobiaceae (14.1% sequences), Comamonadaceae (12.9%), Flavobacteriaceae (8.4%), Xanthomonadaceae (3.1%), Rhodobacteraceae (3.1%), Gemmatimonadaceae (3.0%), Chitinophagaceae (2.0%), Opitutaceae (1.5%), Sphingomonadaceae (1.5%), Cytophagaceae (1.3%) and Rhodocyclaceae (1.1%). *Luteolibacter* (11.6%), *Flavobacterium* (8.3%), *Gemmatimonas* (3.0%), *Hydrogenophaga* (2.1%) and *Opitutus* (1.2%) were the dominant genera.

Bacteroidetes relative abundance declined over time initially, but remained stable ~16–35 days ($P < 0.01$), whereas Alphaproteobacteria and Gemmatimonadetes increased initially and plateaued ~35 days and remained stable thereafter ($P < 0.01$, Table S4, Supporting information). Chloroflexi relative abundance remained stable until 35 days and then declined. Betaproteobacteria relative abundance decreased, whereas Acidobacteria increased linearly over time ($P < 0.01$, Table S4, Supporting information). All dominant phyla's relative abundance differed across sites and/or had a site-by-time interaction except Alphaproteobacteria and Verrucomicrobia ($P < 0.05$, Table S4, Supporting information). *Flavobacterium* and *Hydrogenopha* linearly declined in relative abundance over time ($P < 0.01$, Table 3). *Gemmatimonas* and *Luteolibacter* increased over time, until day 35 when their relative abundance plateaued ($P < 0.01$). *Flavobacterium*, *Gemmatimonas* and *Hydrogenophaga* had a time and site interaction, whereas *Luteolibacter* only differed across sites ($P < 0.05$, Table 3). Both Betaproteobacteria and Gammaproteobacteria were negatively correlated with water temperature (T = −3.01, −2.56, $F_{1,12}$ = 9.07, 6.54, Adj. R^2 = 0.38, 0.30,

Table 1 Multiple regression model statistics for microbial abundance and biofilm metabolism data. Model results indicate all response variables change over successional time except CR. Note that chlorophyll *a*, bacterial abundance and time were $\log_{10}$-transformed for microbial abundance models. Estimate coefficients represent regression slope values unless found in the intercept component of each model

				Full model statistics			
Response variable	Estimate	*T*-value	*P*-value	*F*-statistic	d.f.	Adj. R^2	*P*-value
Chlorophyll *a*				45.46	12	0.93	<0.01
Intercept	−0.71	−4.86	<0.01				
Time	0.71	5.68	<0.01				
Site 2	−0.91	−4.38	<0.01				
Site 3	0.15	0.72	0.49				
Time × Site 2	0.67	3.84	<0.01				
Time × Site 3	0.12	0.70	0.50				
Bacteria				44.71	11	0.94	<0.01
Intercept	4.86	16.88	<0.01				
Time	3.39	6.43	<0.01				
Time^2	−0.88	−3.73	<0.01				
Site 2	−0.15	−0.51	0.62				
Site 3	0.37	1.26	0.24				
Time × Site 2	0.26	1.04	0.32				
Time × Site 3	−0.08	−0.32	0.76				
GPP				53.59	7	0.96	<0.01
Time	-2×10^6	0.70	0.51				
Time^2	9×10^8	2.49	0.04				
Site 1	1×10^4	2.55	0.04				
Site 2	2×10^4	4.31	<0.01				
Site 3	2×10^4	5.88	<0.01				
Time × Site 2	-6×10^7	−0.40	0.70				
Time × Site 3	-6×10^7	−0.48	0.64				
CR				51.50	8	0.96	<0.01
Time	-6×10^7	−0.72	0.48				
Site 1	-2×10^4	−4.68	<0.01				
Site 2	-1×10^4	−5.73	<0.01				
Site 3	-2×10^4	−7.31	<0.01				
Time × Site 2	-4×10^7	−0.35	0.74				
Time × Site 3	-8×10^7	0.76	0.47				
NPP				6.96	7	0.75	0.01
Time	-6×10^6	−1.76	0.12				
Time^2	1×10^7	3.18	0.02				
Site 1	1×10^5	0.20	0.85				
Site 2	4×10^5	0.98	0.36				
Site 3	6×10^5	1.41	0.20				
Time × Site 2	-7×10^7	−0.45	0.66				
Time × Site 3	4×10^7	0.24	0.81				

$P = 0.01$, 0.03, respectively). Verrucomicrobia was negatively correlated with water turbidity ($T = -2.62$, $F_{1,12} = 6.85$, Adj. $R^2 = 0.31$, $P = 0.03$). Concerning bacterial genera relative abundance, *Luteolibacter* was negatively correlated with turbidity ($T = -2.49$, $F_{1,12} = 6.15$, Adj. $R^2 = 0.28$, $P = 0.03$) and *Flavobacterium* was negatively correlated with temperature ($T = -4.52$, $F_{1,12} = 20.44$, Adj. $R^2 = 0.60$, $P < 0.01$).

Bacterial communities changed compositionally over both time and site (NMDS, 2D Stress = 0.13; Fig. 4). Temporal trends in composition were most apparent visually across NMDS Axis 2. Based on linear regression, NMDS Axis 2 scores increased over time ($T = 6.16$, $P < 0.01$) and differed between sites (Site 3 $T = 2.40$, $P = 0.03$; Model Adj. $R^2 = 0.72$). NMDS Axis 1 scores also differed across sites (Site 2 and 3 $T \geq 3.53$, $P < 0.01$, Model Adj. $R^2 = 0.74$). PERMANOVA of bacterial community data indicated that community composition differed over both time ($R^2 = 0.28$, $P < 0.01$) and site ($R^2 = 0.30$, $P < 0.01$), with a significant time and site interaction ($R^2 = 0.12$, $P < 0.01$). Some variation remained unrepresented by our experimental design (Residual $R^2 = 0.29$). No physicochemical variables were correlated with community composition (*envfit* function; $P \geq 0.10$). However, based on phylogenetic turnover analyses, bacterial communities significantly increased in βNTI over time in pairwise comparisons between other 'later' successional communities and the earliest

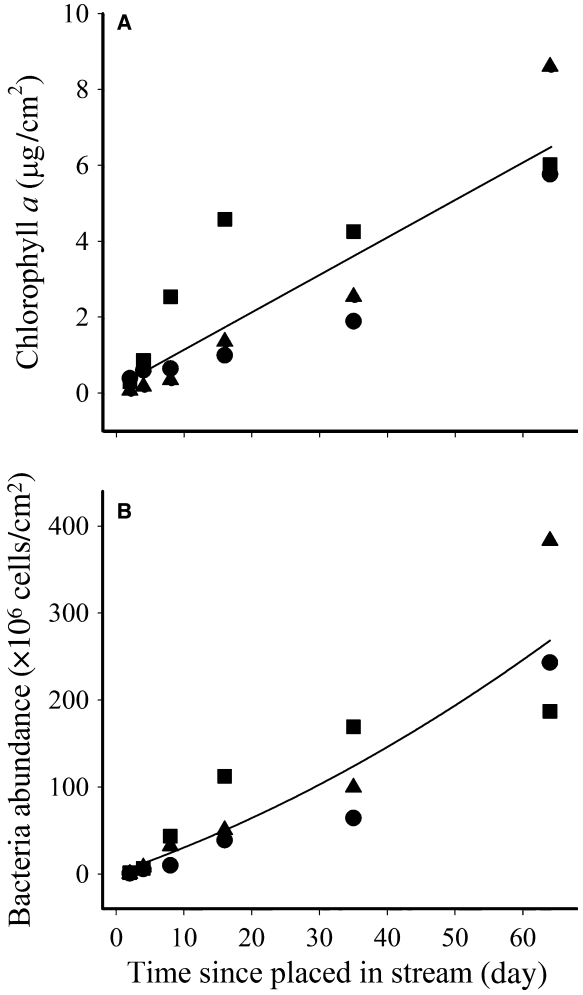

Fig. 1 Algal biomass (chlorophyll *a;* Panel A) and bacterial cell abundance (Panel B) over time (2–64 days). Both measures of microbial abundance increased over time ($P < 0.01$, Adj. $R^2 \geq 0.93$), whereas chlorophyll *a* also differed across sites. Both panels are displaying raw data. Different shapes represent the three site samples (circles = site 1, triangle = site 2, square = site 3). Regression statistics are listed in Table 1.

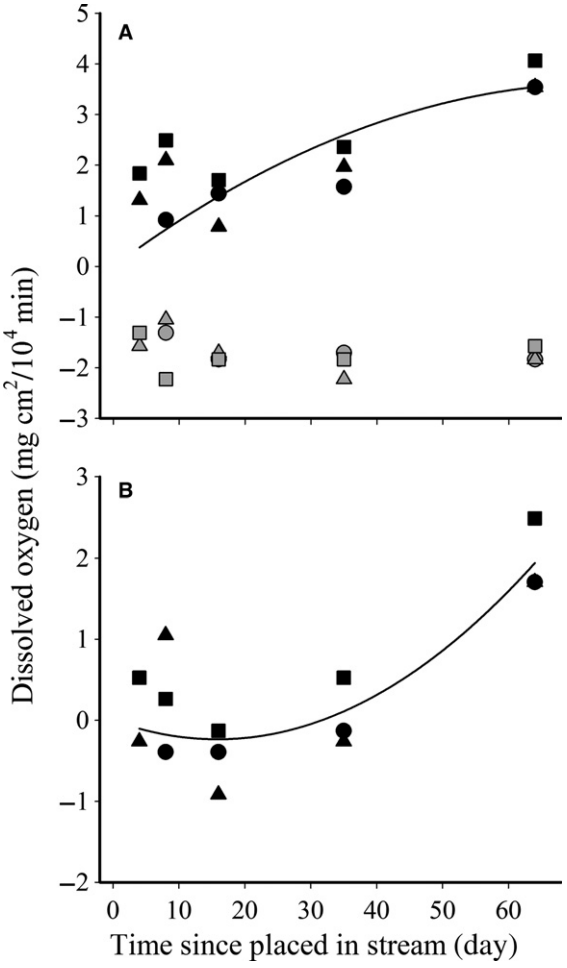

Fig. 2 Gross primary productivity (GPP), community respiration (CR; Panel A) and net primary productivity (NPP; Panel B) measured on tiles across time (4–64 days). GPP and NPP increased over time ($P \leq 0.01$, Adj. $R^2 \geq 0.75$), whereas GPP also differed across sites. Note that respiration signifies oxygen consumption, and therefore all values are negative. Different shapes represent the three site samples (circles = site 1, triangle = site 2, square = site 3). Regression statistics are listed in Table 1.

stage of succession (day 2) within each site (Adj. $R^2 = 0.77$, $P < 0.01$; Fig. 5). Further, βNTI did not differ across sites, suggesting biofilm-associated bacterial community assembly is consistently driven by stochastic processes (e.g. drift, dispersal) during early succession and influenced by selection later (≥35 days) (Fig. 5).

Discussion

In this study, our goal was to understand how stream biofilm microbial communities establish and develop over time, in terms of how microbial groups (algae, bacteria) change in biomass and metabolism (auto- and heterotrophy) temporally, as well as pinpoint what mechanisms influence bacterial community succession. These results serve as a first step towards understanding temporal dynamics of microbial community structure and function in streams, but also confirm that bacterial communities are likely influenced by a stochastic arrival of propagules leading eventually to a convergence towards similar communities as a result of selection-based processes.

Development of biofilm microbial abundance and metabolic rates

Contrary to our hypothesis, algal biomass and bacterial cell abundance increased over time but did not appear to plateau (Fig. 1). Our results contrast similar

Table 2 Multiple regression model statistics for bacterial OTU richness (S_{obs}), the complement of Simpon's diversity and Simpson's evenness with successional time, site and their interaction. Estimate coefficients represent regression slope values unless found in the intercept component of each model

				Full model statistics			
Response variable	Estimate	*T*-value	*P*-value	*F*-statistic	d.f.	Adj. R^2	*P*-value
S_{obs}				11.05	12	0.75	<0.01
Intercept	2634.11	19.43	<0.01				
Time	11.50	2.61	0.02				
Site 2	1152.76	6.01	<0.01				
Site 3	239.88	1.25	0.23				
Time × Site 2	−17.12	−2.75	0.02				
Time × Site 3	−13.46	−2.16	0.05				
Diversity				18.65	12	0.84	<0.01
Intercept	0.97	538.97	<0.01				
Time	2×10^4	3.41	<0.01				
Site 2	0.02	6.57	<0.01				
Site 3	-5×10^4	−0.21	0.83				
Time × Site 2	-2×10^4	−2.17	0.05				
Time × Site 3	-2×10^5	−0.27	0.79				
Evenness				23.69	12	0.87	<0.01
Intercept	0.01	12.42	<0.01				
Time	1×10^4	3.27	<0.01				
Site 2	0.008	6.06	<0.01				
Site 3	−0.001	−0.79	0.45				
Time × Site 2	-3×10^5	−0.74	0.47				
Time × Site 3	3×10^5	0.72	0.48				

studies that evaluate the return time of benthic algae in prairie streams after a flood event (Fisher *et al.* 1982; Dodds *et al.* 1996). Those studies concluded that algal biofilm communities reach predisturbance levels within 2 weeks (Dodds *et al.* 1996). These studies examined secondary successional dynamics after a flooding event instead of a primary successional sequence. Habitat differences (e.g. geochemistry, discharge) may partially explain discrepancies. Primary succession and colonization of biofilm microorganisms may be delayed or require conditioning (i.e. polymeric substance present from the overlying water) before biofilm construction occurs. In addition, nutrient availability, environmental temperature, hydrophobicity of substrata or ionic interactions with bacteria and substrata may affect the timing and rate of biofilm development (Melo & Bott 1997; Siboni *et al.* 2007). Although temperature was likely not low enough to slow biofilm growth (Tables S1 and S3, Supporting information), substrata may have required conditioning of organic materials (Siboni *et al.* 2007) before biofilm communities establish. Further, grazing scars were noticeable on biofilms across sampling times and sites (A. Veach, personal observation), and therefore continual removal of biomass by consumers may also have limited the retention of microbial biofilm biomass. Minnows, whose diet is largely composed of algae (*Phoxinus erythrogaster* and *Campostoma anomalum*), are the most abundant fish consumers in the study reach (Franssen *et al.* 2006). In fact, after a simulated flooding event, the presence of *P. erythrogaster* decreased algal biomass, albeit weakly, compared to no fish treatments (Murdock *et al.* 2011). Due to low baseflow before a precipitation event (prior to 16 days samples), fish densities were likely high and grazing effects and biomass removal may have been amplified during our early observations. Despite the lag in microbial biomass, bacterial diversity was high within 2 days, suggesting that it is unlikely that colonization was delayed (at least for the heterotrophic component of biofilms), but that the growth and establishment of microbial communities, particularly algae, within biofilms may be stunted because of unfavourable habitat conditions or grazing.

In agreement with our hypothesis, NPP gradually increased over time and was closely related to changes in GPP (Fig. 2). CR rates were high within 4 days indicating that an active heterotrophic microbial community capable of using stream water-derived dissolved organic carbon established quickly. Biofilms exhibited no relative change in DO consumption vs. production until late succession (64 days), which is likely attributable to high

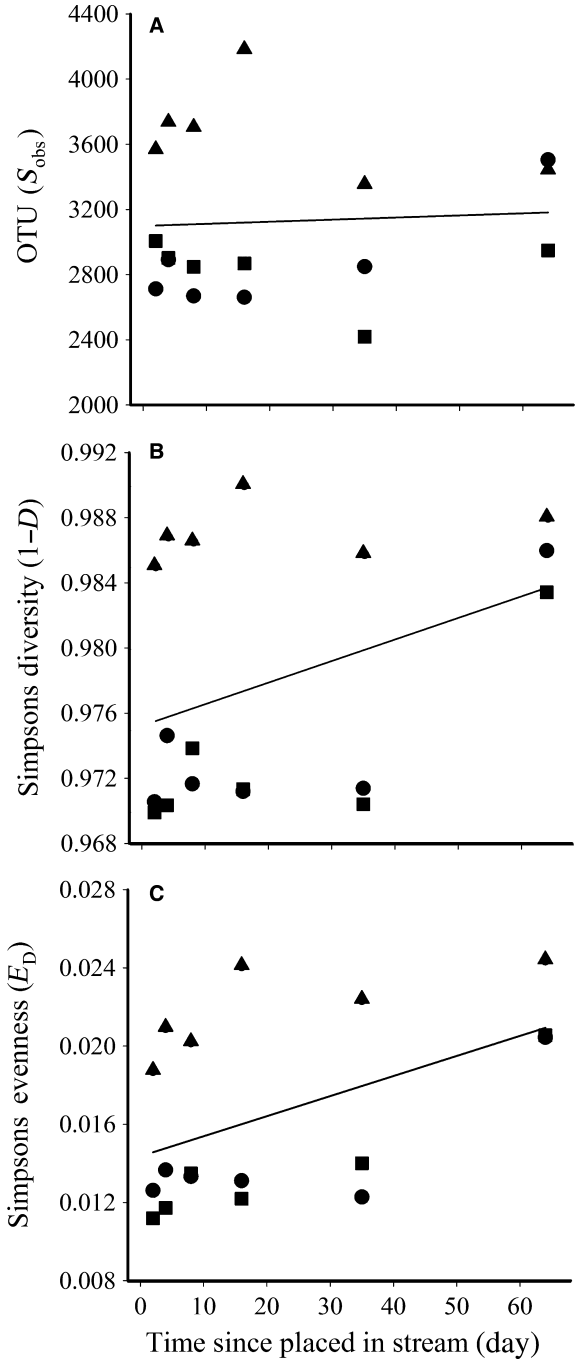

Fig. 3 Observed bacterial OTU richness (S_{obs}; panel A), the complement of Simpson's diversity (1 − D; panel B) and Simpson's evenness (ED; panel C) over time and across sites. All differed over time and across sites ($P < 0.01$, Adj. $R^2 \geq 0.71$). OTUs with a sequence count ≤10 were removed prior to these calculations. Different shapes represent the three site samples (circle = site 1, triangle = site 2, square = site 3). Regression statistics are listed in Table 2.

algal biomass during this period. Generalizations about the importance of allochthonous or autochthonous carbon sources during early biofilm succession may not be discernible. Pohlon *et al.* (2009) found that pioneer biofilm communities (7 days of growth) had low ratios of β-xylosidase: β-glucosidase enzymes suggesting that biofilms relied upon carbon sources derived from algae (as would be expected for net autotrophic communities). They found that late stages of succession (5 months) resulted in biofilms likely using compounds from allochthonous carbon sources. Contrary to this, biofilms became net autotrophic after 2 months of growth and exhibited a trajectory towards net autotrophy instead of one directed to net heterotrophy. Allowing longer time frames of primary succession to occur in this study system (e.g. several months) may eventually lead to communities dependent upon autochthonous sources, as seen in other studies examining ecosystem metabolism in prairie streams (Riley & Dodds 2012).

Succession of bacterial biofilm communities

Succession has been described as a random, stochastic arrival and assembly of taxa that, over time, converge into similar community types due to local, deterministic effects (Del Moral 2009). Studies evaluating succession in biofilms suggest that initial biofilm formation is largely stochastic, reliant upon species recruitment from the water column (Lyautey *et al.* 2005). After this initial colonization, biomass increases and microbial cells compete for resources, resulting in competitive exclusion and a decline in species diversity. Once the biofilm matures and multiple niches exist spatially, multiple species may become abundant due to a more complex, physical and chemical environment present (Lyautey *et al.* 2005; Besemer *et al.* 2007). In this study, we found that bacterial community composition changed substantially over both time and space (Fig. 4), and although diversity was high during early succession, it increased over time (Fig. 3). These patterns are somewhat counter to our hypotheses, but may indicate that we excluded the earliest biofilm successional stages due to sampling after 2 days instead of several hours (Pohlon *et al.* 2009). However, high alpha diversity initially may suggest that dispersal is unlimited and community assembly mechanisms, at least during early succession, are 'weak' (Weiher & Keddy 1999) and demographic stochasticity (e.g. random subsets of regional species pools), due to a plethora of open niche space, may control colonization and initial community establishment (Weiher *et al.* 2011; Cornell & Harrison 2014). Based on our analysis of phylogenetic turnover of community composition, bacterial communities are primarily influenced by stochastic processes during early succession (Fig. 5). Only after biofilms have gained significant biomass (day 35; Figs 1 and 5), do they display a greater phylogenetic turnover indicating a lack of niche-based

Table 3 Multiple linear regression model statistics for dominant bacterial genera relative abundance and successional time, site and their interaction. Estimate coefficients represent regression slope values unless found in the intercept component of each model

				Full model statistics			
Response variable	Estimate	*T*-value	*P*-value	*F*-statistic	d.f.	Adj. R^2	*P*-value
Luteolibacter				6.01	11	0.64	<0.01
Intercept	9.54	7.54	<0.01				
Time	0.33	1.01	0.33				
Time^2	−0.004	−3.43	<0.01				
Site 2	−4.28	−2.7	0.02				
Site 3	1.61	1.01	0.33				
Time × Site 2	0.03	0.62	0.55				
Time × Site 3	−0.02	−0.39	0.70				
Flavobacterium				35.53	12	0.91	<0.01
Intercept	7.88	12.55	<0.01				
Time	−0.07	−3.50	<0.01				
Site 2	2.29	2.58	0.02				
Site 3	6.92	7.79	<0.01				
Time × Site 2	−0.04	−1.46	0.17				
Time × Site 3	−0.11	−3.96	<0.01				
Gemmatimonas				15.34	11	0.84	<0.01
Intercept	2.35	7.65	<0.01				
Time	0.10	4.63	<0.01				
Time^2	−0.001	−4.15	<0.01				
Site 2	−0.77	−2.01	0.07				
Site 3	−1.60	−4.17	<0.01				
Time × Site 2	0.02	1.98	0.07				
Time × Site 3	0.04	2.86	0.02				
Hydrogenophaga				15.55	12	0.81	<0.01
Intercept	3.52	19.32	<0.01				
Time	−0.04	−6.27	<0.01				
Site 2	−1.99	−7.70	<0.01				
Site 3	−1.76	−6.81	<0.01				
Time × Site 2	0.04	5.17	<0.01				
Time × Site 3	0.05	5.63	<0.01				
Opitutus				5.6	12	0.57	<0.01
Intercept	1.25	6.08	<0.01				
Time	−0.004	−0.55	0.59				
Site 2	−0.43	−1.47	0.17				
Site 3	−0.67	−2.29	0.04				
Time × Site 2	0.03	2.73	0.02				
Time × Site 3	0.03	3.24	<0.01				

mechanisms driving community assembly during early succession. Dini-Andreote *et al.* (2015) found that stochastic processes were more influential during early succession in a salt marsh bacterial community likely due to random dispersal and immigration and a lack of nutrients that would have acted as an environmental filter. Similar to our study, they found that selection was the driving process later in time for bacterial community assembly due to sodium accumulation. In this study, we cannot discern specific environmental conditions that are linked to the switch to selection-based mechanisms. However, we do suggest that early succession is largely characterized by, but not necessarily exclusive to, passive dispersal and random colonization events. As algal, bacterial and other microbial biomass increases over time, a complex physical and chemical gradient establishes both vertically and laterally within biofilms (available nutrients, oxygen, redox gradients) and niche-based selection becomes the overriding process structuring bacterial communities. We also cannot exclude species interactions as a contributing selective force. Specifically, bacterial–bacterial and algal–bacterial interactions (among others) may be a biotic selective force particularly since algae are 'ecosystem engineers' in stream benthic biofilms (Besemer *et al.* 2007). Algae modify both physical structure and biofilm carbon

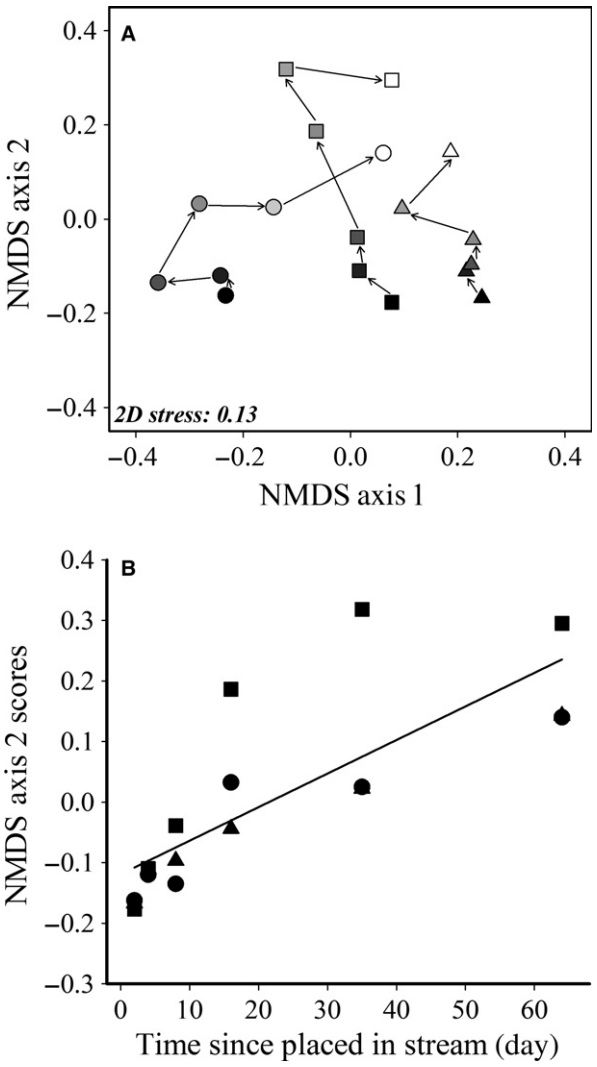

Fig. 4 Nonmetric dimensional scaling (NMDS) ordination of bacterial communities sampled across time (2D Stress = 0.13, Panel A). Communities sampled at each time point are denoted by a gradient of black/grey colour scheme (2-day communities are black; 64-day communities are white). Different shapes represent the three site samples (circles = site 1, triangle = site 2, square = site 3). Bacterial community composition significantly changed over time and site (Adj. $R^2 = 0.72$, $P < 0.01$) based on NMDS axis scores (Panel B).

cycling (Battin *et al.* 2003, 2016), thereby potentially influencing bacterial community structure via these changes. Further, algal communities are important components of stream biofilms, which can develop predictably after disturbances (Fisher *et al.* 1982; Pohlon *et al.* 2009). Although we did not measure algal community succession in this study, future studies should include both bacterial and algal community measurements to discern the strength of their interactions over time.

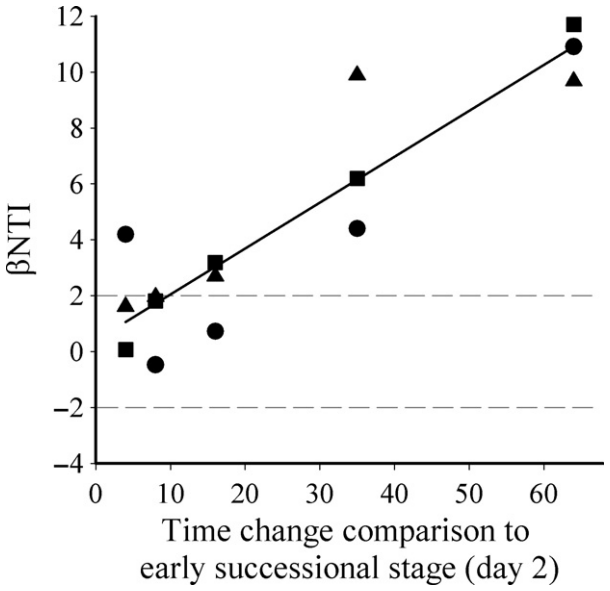

Fig. 5 Phylogenetic turnover in community composition (βNTI) compared to the earliest successional stage within each site. βNTI differed over time (Adj. $R^2 = 0.25$, $P = 0.02$), but not site ($P > 0.10$). Different shapes represent the three site samples (circles = site 1, triangle = site 2, square = site 3). Dashed lines represent 2 standard deviation units.

Spatiotemporal variation in community composition

Previous work on temporal or spatial dynamics of bacterial communities associated with apple flowers (Shade *et al.* 2013), leaf surfaces (Redford & Fierer 2009), deglaciated soils (Nemergut *et al.* 2007; Brown & Jumpponen 2014) and salt marshes (Dini-Andreote *et al.* 2015) has observed both temporal and spatial effects on community structure and assembly. The commonality across these ecosystems is that temporal patterns tend to be stronger than spatial patterns, suggesting that bacterial communities develop predictably over time, but with some degree of variability across locations (Shade *et al.* 2013) or spatial scales (Dini-Andreote *et al.* 2015). Our data indicate that location accounted for a large portion of variation in bacterial community composition (PERMANOVA results, Fig. 4), but communities displayed similar successional trajectories among sites, both evidenced by convergence of community composition later in time (NMDS, Fig. 4) and similar dynamics in phylogenetic turnover across sites (i.e. βNTI > 2 at approximately same successional stage, Fig. 5). Further, relative abundance of many dominant bacterial phyla and genera differed across sites and many taxa were correlated with environmental parameters that we measured implying that some community members are as influenced by space (or environmental differences across space) relative to time.

The majority of OTUs and sequences represented Bacteroidetes, Betaproteobacteria and other classes of Proteobacteria, and Verrucomicrobia. Other studies (Boucher *et al.* 2006; Besemer *et al.* 2007, 2012; Pohlon *et al.* 2009; Battin *et al.* 2016) have found that Proteobacteria (especially Betaproteobacteria and Alphaproteobacteria), Bacteroidetes and Cyanobacteria dominate in freshwater bacterial biofilms, but relatively few have documented dominance of Verrucomicrobia members in these biofilms (but see Boucher *et al.* 2006). Fierer *et al.* (2013) found that Verrucomicrobia are very abundant in tallgrass prairie soils but have remained underrepresented in previous studies because of sequencing or primer biases (Bergmann *et al.* 2011). Suspended freshwater bacterial communities can be compositionally very similar to soil inoculum (Crump *et al.* 2012). If true, it is unsurprising that the regional species pool in tallgrass prairie streams (i.e. bacteria in the water column) is dominated by species of Verrucomicrobia as well as Proteobacteria and Bacteroidetes that form, and continue to dominate, freshwater biofilms during succession. However, we did not target active communities in this study (i.e. 16S rRNA instead of rRNA genes), suggesting that some trends in our data may be due to deposition of inactive or dead cells. Similar to Besemer *et al.* (2012), future studies addressing these questions in aquatic environments should include sampling of both the local (e.g. biofilm, sediments, detritus) and regional (water column) species pools. In addition, collecting data on important environmental conditions characterizing local habitat as well as using both DNA- and RNA-based approaches will help gain a more complete understanding of large-scale processes governing microbial community assembly spatiotemporally.

In summary, our study suggests that microbial communities establish stochastically but develop predictably over time, regaining ecosystem productivity rates and with biofilms becoming increasingly reliant upon autochthonous carbon sources. Further, biofilm-associated bacteria converge to similar communities that are influenced by stochastic processes during early succession, but driven by selection after biofilms start to mature. Although we examined several microbial components over time (algal and bacterial abundance, ecosystem process rates, bacteria assemblages), additional research is needed to mechanistically link bacterial function and structure throughout ecological succession.

Acknowledgements

This research was supported by the National Science Foundation's Konza Prairie Long Term Ecological Research programme (DEB – 0823341) awarded to W.K.D. We thank John Brant and Matt Troia for assistance in the laboratory and field; Alina Akhunova and Hanquan Liang with MiSeq sequencing assistance; and Matt Troia and Lydia Zeglin for helpful comments which improved this manuscript. This is contribution no. 17-062-J from the Kansas Agricultural Experiment Station. J.C.S. was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of the Subsurface Biogeochemical Research Program's Scientific Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). PNNL is operated for DOE by Battelle under contract DE-AC06-76RLO 1830. A portion of the research was performed using Institutional Computing at PNNL. The authors declare no conflict of interest related to this work.

References

American Public Health Association (APHA) (1995) *Standard Methods for the Examination of Water and Wastewater*, 19th edn. American Public Health Association, American Water- works Association, and Water Environment Federation, Washington, District of Columbia.

Anderson MJ (2001) Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 626–639.

Araya R, Tani K, Takagi T, Yamaguchi N, Nasu M (2003) Bacterial activity and community composition in stream water and biofilm from an urban river determined by fluorescent in situ hybridization and DGGE analysis. *FEMS Microbiology Ecology*, **43**, 111–119.

Battin T, Kaplan LA, Newbold JD, Hansen CME (2003) Contributions of microbial biofilms to ecosystem processes in stream mesocosms. *Nature*, **426**, 439–442.

Battin TJ, Besemer K, Bengtsson MM, Romani AM, Packman AI (2016) The ecology and biogeochemistry of stream biofilms. *Nature Reviews*, **14**, 251–263.

Beaulieu JJ, Tank JL, Hamilton SK *et al.* (2011) Nitrous oxide emission from denitrification in stream and river networks. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 214–219.

Bergmann GT, Bates ST, Eilers KG *et al.* (2011) The underrecognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biology and Biochemistry*, **43**, 1450–1455.

Berry D, Mahfoudh KB, Wagner M, Loy A (2011) Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Applied and Environmental Microbiology*, **77**, 7846–7849.

Besemer K, Singer G, Limberger R *et al.* (2007) Biophysical controls on community succession in stream biofilms. *Applied and Environmental Microbiology*, **73**, 4966–4974.

Besemer K, Peter H, Logue JB *et al.* (2012) Unraveling assembly of stream biofilm communities. *The ISME Journal*, **6**, 1459–1468.

Bott TL (1996) Primary productivity and community respiration. In: *Methods in Stream Ecology* (eds Hauer FR, Lamberti GA), pp. 663–690. Academic Press, San Diego, California.

Boucher D, Jardillier L, Debroas D (2006) Succession of bacterial community composition over two consecutive years in two aquatic systems: a natural lake and a lake-reservoir. *FEMS Microbiology Ecology*, **55**, 79–97.

Brown SP, Jumpponen A (2014) Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Molecular Ecology*, **23**, 481–497.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.

Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–1624.

Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **366**, 2351–2363.

Clements FE (1916) *Plant Succesion: An Analysis of the Development of Vegetation*. Carnegie Institution of Washington, Washington, District of Columbia.

Connell JH, Slatyer RO (1977) Mechanisms of succession in natural communities and their role in community stability and organization. *The American Naturalist*, **111**, 1119–1144.

Cook RD (1979) Influential observations in linear regression. *Journal of the American Statistical Association*, **74**, 169–177.

Cornell HV, Harrison SP (2014) What are species pools and when are they important? *Annual Review of Ecology, Evolution, and Systematics*, **45**, 45–67.

Costigan KH, Daniels MD, Dodds WK (2015) Fundamental spatial and temporal disconnections in the hydrology of an intermittent prairie headwater network. *Journal of Hydrology*, **522**, 305–316.

Cowles HC (1899) *The Ecological Relations of the Vegetation on the Sand Dunes of Lake Michigan*. University of Chicago Press, Chicago, Illinois.

Crump BC, Amaral-Zettler LA, Kling GW (2012) Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME Journal*, **6**, 1629–1639.

Dang H, Lovell CR (2000) Bacterial primary colonization and early succession on surfaces in marine waters as determined by amplified rRNA gene restriction analysis and sequence analysis of 16S rRNA genes. *Applied and Environmental Microbiology*, **66**, 467–475.

Davey ME, O'Toole GA (2000) Microbial biofilms: from ecology to molecular genetics. *Microbiology and Molecular Biology Reviews*, **64**, 847–867.

Del Moral R (2009) Plant succession on pumice at Mount St. Helens, Washington. *American Midland Naturalist*, **141**, 101–114.

Dini-Andreote F, Stegen JC, Dirk van Elsas J *et al.* (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, E1326–E1332. doi:10.1073/pnas.1414261112.

Dodds WK (1997) Distribution of runoff and rivers related to vegetative characteristics, latitude, and slope: a global perspective. *Journal of the North American Benthological Society*, **16**, 162–168.

Dodds WK, Hutson RE, Eichem AC *et al.* (1996) The relationship of floods, drying, flow, and light to primary production and producer biomass in a prairie stream. *Hydrobiologia*, **333**, 151–159.

Dodds WK, Gido K, Whiles MR, Fritz KM, Matthews WJ (2004) Life on the edge: the ecology of Great Plains prairie streams. *BioScience*, **54**, 205–216.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.

Fierer N, Nemergut D, Knight R, Craine JM (2010) Changes through time: Integrating microorganisms into the study of succession. *Research in Microbiology*, **161**, 635–642.

Fierer N, Ladau J, Clemente JC *et al.* (2013) Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science*, **342**, 621–624.

Fine PVA, Kembel SW (2011) Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography*, **34**, 552–565.

Fisher SG, Gray LJ, Grimm NB, Busch DE (1982) Temporal succession in a desert stream ecosystem following flash flooding. *Ecological Monographs*, **52**, 93–110.

Franssen NR, Gido KB, Guy CS *et al.* (2006) Effects of floods on fish assemblages in an intermittent prairie stream. *Freshwater Biology*, **51**, 2072–2086.

Gihring TM, Green SJ, Schadt CW (2012) Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environmental Microbiology*, **14**, 285–290.

Gleason HA (1926) The individualistic concept of the plant association. *Journal of the Torrey Botanical Society*, **53**, 7–26.

Hubbell SP (2001) *The Unified Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, New Jersey.

Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology*, **12**, 1889–1898.

Jackson CR, Churchill PF, Roden EE (2001) Successional changes in bacterial assemblage structure during epilithic biofilm development. *Ecology*, **82**, 555–566.

Kembel SW, Cown PD, Helmus MR *et al.* (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463–1464.

Kemp MJ, Dodds WK (2002) Comparisons of nitrification and denitrification in prairie and agriculturally influenced streams. *Ecological Applications*, **12**, 998–1009.

Li Y, Wang C, Zhang W *et al.* (2015) Modeling the effects of hydrodynamic regimes on microbial communities within fluvial biofilms: combining deterministic and stochastic processes. *Environmental Science & Technology*, **49**, 12869–12878.

Lyautey E, Jackson CR, Cayrou J *et al.* (2005) Bacterial community succession in natural river biofilm assemblages. *Microbial Ecology*, **50**, 589–601.

Melo LF, Bott TR (1997) Biofouling in water systems. *Experimental Thermal and Fluid Science*, **14**, 375–381.

Mulholland PJ, Helton AM, Poole GC *et al.* (2008) Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature*, **452**, 202–246.

Murdock JN, Dodds WK, Gido KB *et al.* (2011) Dynamic influences of nutrients and grazing fish on periphyton during recovery from flood. *Journal of the North American Benthological Society*, **30**, 331–345.

Nemergut DR, Anderson SP, Cleveland CC *et al.* (2007) Microbial community succession in an unvegetated, recently deglaciated soil. *Microbial Ecology*, **53**, 110–112.

Oksanen J, Blanchet FG, Kindt R *et al.* (2016) vegan: Community Ecology Package. R package version 2.3-3. Available from http://CRAN.R-project.org/package=vegan.

Pohlon E, Marxsen J, Küsel K (2009) Pioneering bacterial and algal communities and potential extracellular enzyme activities of stream biofilms. *FEMS Microbiology Ecology*, **71**, 364–373.

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, **26**, 1641–1650.

R Development Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Australia. Available from http://www.R-project.org.

Redford AJ, Fierer N (2009) Bacterial succession on the leaf surface: a novel system for studying successional dynamics. *Microbial Ecology*, **58**, 189–198.

Riley AJ, Dodds WK (2012) The expansion of woody riparian vegetation, and subsequent stream restoration, influences the metabolism of prairie streams. *Freshwater Biology*, **57**, 1138–1150.

Sartory DP, Grobbelaar JU (1984) Extraction of chlorophyll a from freshwater phytoplankton for spectrophotometer analysis. *Hydrobiologia*, **114**, 117–187.

Schloss PD, Westcott SL, Ryabin T *et al.* (2009) Introducing mothur: open-source, platform-independent, community- Supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**, 7537–7541.

Shade A, McManus PS, Handelsman J (2013) Unexpected diversity during community succession in the apple flower microbiome. *mBio*, **4**, e00602–e00612, DOI:10.1128/mBio.00602-12.

Siboni N, Lidor M, Kramarsky-Winter E, Kushmaro A (2007) Conditioning film and initial biofilm formation on ceramic tiles in the marine environment. *FEMS Microbiology Letters*, **274**, 24–29.

Sousa WP (1979) Experimental investigations of disturbance and ecological succession in a rocky interdial algal community. *Ecological Monographs*, **49**, 227–254.

Stegen JC, Lin X, Konopka AE *et al.* (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal*, **6**, 1653–1664.

Stegen JC, Lin X, Fredrickson JK *et al.* (2013) Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, **7**, 2069–2079.

Veach AM, Dodds WK, Jumpponen A (2015) Woody plant encroachment, and its removal, impact bacterial and fungal communities across stream and terrestrial habitats in a tallgrass prairie ecosystem. *FEMS Microbiology Ecology*, **91**, fiv109. doi:10.1093/femsec/fiv109.

Vellend M (2010) Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, **85**, 183–206.

Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*, 4th edn. Springer, New York, New York.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.

Wang J, Shen J, Wu Y *et al.* (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *The ISME Journal*, **7**, 1310–1321.

Webb CO, Ackerly DD, Kembel S (2011) *Phylocom: Software for the Analysis of Phylogenetic Community Structure and Character Evolution*. Users Manual Version 4.2. http://phylodiversity.net/phylocom/phylocom_manual.pdf.

Weiher E, Keddy P (1999) Assembly rules as general constraints on community composition. In: *Assembly Rules: Perspectives, Advances, Retreats* (eds Weiher E, Keddy P), pp. 251–271. Cambridge University Press, Cambridge, UK.

Weiher E, Freund D, Bunton T *et al.* (2011) Advances, challenges, and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **366**, 2403–2413.

Wrona FJ, Prowse TD, Reist JD *et al.* (2006) Climate change effects on aquatic biota, ecosystem structure and function. *Ambio*, **35**, 359–369.

Zhou J, Deng Y, Zhang P *et al.* (2014) Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E836–E845. doi:10.1073/pnas.1324044111.

This study was conceived and designed by A.M.V., W.K.D. and A.J. Sample collections were performed by A.M.V., microbial abundance and biofilm metabolism data collected by A.M.V. and W.K.D., and library preparation and bioinformatics analysis performed by A.M.V., S.P.B. and A.J. Statistical analyses were performed by A.M.V. and J.C.S., and the manuscript was written by all authors.

Data accessibility

16S rRNA gene sequence data are accessible at NCBI SRA: SRP077840. Microbial abundance, metabolism and bacterial OTU abundance and taxonomy data are located at Dryad doi: 10.5061/dryad.m2r17.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Map of the three locations where tiles were submerged in the main reach of Kings Creek at Konza Prairie Biological Station.

Fig. S2 The amount of precipitation received in Manhattan, KS during the study period (location of gage: Manhattan Airport).

Fig. S3 Mothur-calculated rarefaction curve for each successional time point.

Table S1 Abiotic site characteristics (temperature, dissolved inorganic nitrogen, and turbidity) collected during sampling times.

Table S2 The 12 bp Molecular Identifier Tags for each sample across time (2, 4, 8, 16, 35, 64 days), site (1, 2, 3), and replicate within a site (1, 2, 3) used in the secondary PCR.

Table S3 Multiple regression statistics for environmenta data collected and time, site, and time and site interactions.

Table S4 Multiple regression statistics for dominant bacterial phyla and time, site, and time and site interactions.