2013

Bacterial Assemblages And Microbial Enzyme Activity In The Major Tributaries Of The Lower Mississippi River

Justin Millar
University of Mississippi

Follow this and additional works at: https://egrove.olemiss.edu/etd
Part of the Microbiology Commons

Recommended Citation
Millar, Justin, "Bacterial Assemblages And Microbial Enzyme Activity In The Major Tributaries Of The Lower Mississippi River" (2013). Electronic Theses and Dissertations. 906.
https://egrove.olemiss.edu/etd/906

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.
BACTERIAL ASSEMBLAGES AND MICROBIAL ENZYME ACTIVITY IN THE MAJOR TRIBUTARIES OF THE LOWER MISSISSIPPI RIVER

A Thesis
presented in partial fulfillment of required
for the degree of Master of Science
in the Department of Biology
The University of Mississippi

By

Justin J. Millar

December 2013
ABSTRACT

Spatial variation in the structure and function of bacterial communities is far more complex than previously believed. High throughput sequencing techniques have begun to elucidate the dynamic worldwide patterns of bacterial biogeography; however, there are still many ecosystems for which there is little to no representative data. Currently there are no high throughput-based analyses of bacterial assemblages in large rivers. Rivers form important connections between terrestrial land use and marine biogeochemistry and act as conduits for organisms and allochthonous nutrients. Bacteria dominate carbon cycling and nutrient processing in rivers, and therefore can influence regional and global elemental cycles. This study characterized the structure of bacterial assemblages of the major tributaries to the Lower Mississippi River using high throughput Ion Torrent PGM sequencing of the 16S rRNA gene, and the function of these communities using assays of extracellular enzyme activity. A total of 715,120 valid bacterial 16S rRNA sequences were classified into 30,211 operational taxonomic units (OTUs), although many of these OTUs were of very low proportional abundance. The most abundant individual OTU was identified as being related to the dominant marine Alphaproteobacteria *Pelagibacter*/SAR11 and was found in all rivers. The most abundant bacterial phylum was the *Cyanobacteria*, of which a number of genetically distinct OTUs were identified as being similar to *Prochlorococcus*. Community structure was related to river physiochemistry, although there was no consistent physicochemical factor responsible for these differences. While there were different bacterial communities in each river, the greatest
difference in assemblage structure was between free living cells and those associated with particles. Despite having different environmental characteristics and distinct bacterioplankton communities, there were no consistent differences in enzyme activity between rivers, although aspects of community structure could be linked to extracellular enzyme activity. Overall, this research shows that large rivers of the Mississippi River Basin harbor distinct bacterioplankton communities but also suggests that these structurally different bacterial assemblages may show similar ecological functions related to nutrient acquisition and organic carbon mineralization.
LIST OF ABBREVIATIONS OR SYMBOLS

ANOSIM – Analysis of similarity
ANOVA – Analysis of variance
Ark – Referring to sites or samples from the Arkansas River
CBH – Cellobiohydrolase
DNA – Deoxyribonucleic acid
DAPI – 4’,6-diamidino-2-phenylindole
DO – Dissolved oxygen
DOC – Dissolved organic carbon
Ill – Referring to sites or samples from the Illinois River
Miz – Referring to sites or samples from the Missouri River
NAGase – N-acetyl-β-D-glucosaminidase
NH₄-N – Ammonia
NMDS – Non-metric multidimensional scaling
NOₓ-N – Oxidized nitrogen
NO₂-N - Nitrite
NO₃-N - Nitrate
Ohi – Referring to sites of samples from the Ohio River
OTU – Operational taxonomic unit
PCA – Principle component analysis
PCR – Polymerase chain reaction
PO₄-P - Phosphate
RDA – Distance-based redundancy analysis

RNA – Ribonucleic acid

TDP – Total dissolved phosphorus

Temp - Temperature

Ten – Referring to sites or samples from the Tennessee River

TKN – Total Kjeldahl nitrogen

UpM – Referring to sites or samples from the Upper Mississippi River

YSI – Yellow Springs Instruments
ACKNOWLEDGEMENTS

I would like to acknowledge the people who through their effort and support made this project possible. This thesis was performed under the direction of Dr. Colin Jackson and Dr. Clifford Ochs, both of whom have been irreplaceable resources over the past two years. Beyond funding support, I leaned heavily on Dr. Jackson for his expertise in this field and particularly in the preparation and publication of these data. Dr. Ochs’ knowledge and field experience was invaluable throughout the rigorous sampling campaign. Without their guidance and insight, this project would not have come into fruition. My final committee member, Dr. Stephen Brewer, was an integral part of the prospectus and thesis preparation, particularly with regards to statistical analyses. I must also recognize Derek Bussan and Alexa Lampkin for aiding in all of the sampling collections. Beyond field work, support from the Jackson and Ochs labs was critically important, as was general input from across the Department of Biology at the University of Mississippi.

Special recognition must be paid to Jason Payne, who worked alongside me throughout every step of this entire thesis. Without Jason’s effort, support, and sacrifice, this study simply would not have become a reality. I must also acknowledge funding support from the National Science Foundation. Finally, I would like to thank my family, both back home and those I have made in Oxford, for all of their support.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS AND SYMBOLS</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>BACKGROUND</td>
<td>4</td>
</tr>
<tr>
<td>METHODOLOGY</td>
<td>11</td>
</tr>
<tr>
<td>RESULTS</td>
<td>19</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>46</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>56</td>
</tr>
<tr>
<td>VITA</td>
<td>65</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Water Chemical and Biological Variables ............................................................... 20
Tukey’s HSD Groups of Enzyme Activities ............................................................. 26
Classifications and Relative Abundances of Major OTUs ........................................ 35
LIST OF FIGURES

Map of Site Locations ............................................................................................................12
Prokaryotic Cell Density ........................................................................................................21
Extracellular Enzyme Activity ...............................................................................................24
Distance-based Redundancy Analysis of Enzyme Profiles .....................................................27
Tree Diagram of Bacterial Assemblages ..................................................................................29
NMDS Ordination of Particle Associated and Free Living Communities .........................30
NMDS Ordination for Free Living Communities by River ......................................................32
NMDS Ordination for Particle Associated Communities by River ......................................33
Community Structure by Relative Abundance of Major Groups ........................................37
S_{\text{Chao1}}, Inverse Simpson, and Shannon Indices of \alpha-Diversity ..................................39
Free Living Community NMDS with Environmental Variables .........................................41
Particle Associated Community NMDS with Environmental Variables ............................42
Free Living Community NMDS with Enzyme Activities .....................................................44
Particle Associated Community NMDS with Enzyme Activities ......................................45
INTRODUCTION

A major goal of ecological research is to develop a clear understanding of the factors that influence community structure in order to model species distributions across local, regional, and global scales. Biogeography describes the spatial distribution of biodiversity and the environmental and historical factors that shape this distribution (Ricklefs 1987). The extensive history of biogeography dates back to the 18th century, when Carl Linnaeus sought to explain the relationship between flora distributions and their geographic and environmental locations (O’Malley 2008). Predictive models of habitat and species distribution have expanded with the development of new research technology and statistical approaches (Guisan and Zimmerman 2000). Gathering representative biogeographic data for the application of these models is an important research objective, particularly for microbial communities which have less of a historical basis than plants or animals.

Molecular approaches, developed over the last few decades, have dramatically changed how microbial ecologists observe the spatial distributions of bacteria and other microorganisms (Pace 1997). Early culture-based approaches identified less than 1% of bacterial species (Stanley and Konopka 1985), which significantly underrepresented microbial diversity. Molecular sequence-based techniques developed by Carl Woese identified these unculturable bacteria and have revealed a far more diverse microbial world (Woese and Fox 1977). Norman Pace further advanced these molecular techniques by using a sequenced-based approach to analyze mixed communities from environmental samples (e.g. Stahl et al. 1985). These technological innovations have allowed microbial ecologists to more accurately describe bacterial community
and assess their spatial distributions. These findings have advanced earlier simplistic views on microbial communities, which suggested that the extreme dispersal potential of bacteria would lead to a ubiquitous condition where “everything is everywhere, the environment selects” (O’Malley 2008).

As biological communities vary across environmental gradients, the differences in the interactions of these communities with the environment can result in spatial patterns of critical ecosystem processes. These interactions are especially important in bacterial communities, which are often the primary regulators of nutrient cycling (Pace et al. 1997). As well as being used to determine community structure, molecular and biochemical techniques have also been developed to describe microbial interactions. For example, artificial substrates containing colorimetric or fluorometric tags can be used to determine the activity of extracellular enzymes in microbial communities (Sinsabaugh and Foreman 2001). These techniques have been used to describe biogeochemical processing and nutrient limitation in environmental samples (Sinsabaugh et al. 1997; Jackson et al. 2013). As researchers continue to apply new approaches to identify assemblage patterns in microbial communities, it is important to also address the functional characteristics of these communities. Assessing biodiversity and functionality simultaneously is a more direct means for describing their relationship, as opposed to drawing connections from separate studies.

Representative data across ecosystems is required in order to form a generalized understanding of microbial biogeography, and this baseline information is still lacking for many environments. In particular, there have been relatively few studies on microbial community structure in river networks, despite the importance of bacteria in riverine systems and the importance of rivers in global ecological processes. Rivers form important connections between
terrestrial and marine environments by acting as conduits for nutrients and organisms. A number of theoretical models describe changes in community structure along riverine systems (Vannote et al. 1980; Junk et al. 1989; Thorp and Delong 1994), but these only consider microbes in a cursory manner. Given that bacteria often dominate carbon cycling in large, well-mixed rivers (Cole et al. 2007), these conceptual models could not only benefit from incorporating a microbial viewpoint themselves, but they can also help describe observed patterns in microbial assemblages. The unidirectional flow of rivers also presents a physical barrier that may be analogous to those considered in the studies of hot springs (Papke et al. 2003; Witaker et al. 2003), and provides a direction of transmission to further describe patterns in microbial community structure. There is a pressing need to describe the community and functional characteristics of riverine bacterial assemblages in order to develop our understanding of microbial biogeography and its importance in vital environmental processes.
BACKGROUND

The brief history of modern microbial biogeography (roughly 25 years compared to over 200 years of research on plant and animal distributions) has yielded a complex and sometimes contradicting view of the spatial distribution of free-living bacteria. As conjectured by early microbiologists, environmental factors are often a primary influence on spatial patterns. In aquatic environments, changes in bacterial communities have been observed along depth (e.g. Ovreas et al. 1997) and salinity gradients (e.g. Crump et al. 2004). Similar patterns have been described in sediments for changes in salinity (e.g. Jackson and Vallaire 2009), and along temperature gradients in archaeal hot spring communities (e.g. Miller et al. 2009). Changes in carbon (Kuffner et al. 2012; Deng et al. 2013) and nitrogen inputs (Jackson and Vallaire 2009) can also influence the structure of microbial communities. Soil pH, in conjunction with the C-to-N ratio, has also been shown to influence microbial community composition (Hogberg et al. 2007). Meta-analyses of bacterial diversity across multiple spatial scales have suggested that a single environmental variable such as pH (Lozupone and Knight 2007) or temperature (Fuhrman et al. 2008) may be the primary driver of bacterial biogeography.

In addition to environmental variables, geographic influences can also play an important role in shaping bacterial assemblages. For example, Cho and Teidje (2000) identified endemic populations of fluorescent Pseudomonas at the genomic level, suggesting that these bacteria were not globally mixed. Similarly, an inverse relationship between population similarity and geographic distance has been observed at small (meters) scales in purple nonsulfur bacteria despite similar environmental conditions (Oda et al. 2003). Martiny et al. (2011) measured
community composition in ammonia oxidizing bacteria and determined that distance, from centimeters to continental divides, drives differences in community structure. These studies suggest that distance is important in separating microbial communities, and the compilation of local and regional studies has led to larger scale observations of microbial diversity (Lozupone and Knight 2007) that in some ways reflect similar patterns in plants and animals. For instance, the direct relationship between island/area size and diversity that has been described in plant (Kohn and Walsh 1994) and animal communities (Ricklefs and Lovette 1999) has also been observed in microbial assemblages (Horner-Devine et al. 2003; Bell et al. 2005).

The emerging view of microbial biogeography is complex, and the ecological significance of bacteria and other microbes makes understanding the dynamics of their distributions a critical area of research. Prokaryotic cells are the primary regulators of many global biogeochemical processes, store approximately 60-100% of the carbon contained in plants, and far more nitrogen, phosphorus, and other nutrients (Whitman et al. 1998). In addition to their role in various chemical cycles, bacteria can provide other ecological services, such as serving as bioindicators for nutrients and contaminants (Lemly and King 2000; Glassmeyer et al. 2005). Due to the many ecological roles that bacteria can fill it is important to not only describe patterns of bacterial biogeography, but also how this influences the functional characteristics of bacterial assemblages.

As the primary regulators of biogeochemical pathways, bacterial communities play a significant role in nutrient processing in many ecosystems. Microorganisms acquire specific nutrients from organic matter by producing extracellular enzymes that liberate target substrates from macromolecules (Sinsabaugh and Foreman 2001). Activity of extracellular enzymes has been used to measure biogeochemical processing (e.g. Sinsabaugh et al. 1997; Jackson et al.
2013), and as a proxy for detecting nutrient limitation (Hill et al. 2010). Extracellular enzyme activity has been found to covary with microbial community structure (Leff et al. 2012), and has been identified as a reasonable predictor of patterns in microbial succession (DeAngelis et al. 2013).

Understanding the underlying dynamics of microbial community structure and function is an emerging area for research. Data on these processes is particularly lacking for large riverine systems, which is surprising given the ecological importance of rivers. The transport of terrestrially-derived materials through lotic ecosystems forms an important connection between land use patterns and biogeochemical processes in freshwater and marine environments. This interaction is especially significant in the Mississippi River basin, which is the second and third highest global contributor of nitrogen and phosphorus, respectively, to the ocean (Howarth et al. 1996). The Mississippi River is the fourth longest and the tenth largest river in the world, and the most significant fluvial ecosystem in North America. The total drainage area of the Mississippi River and its tributaries covers approximately 40% of the continental United States, and discharges an annual average of 18,000 m$^3$ sec$^{-1}$ and 145 million metric tons of suspended sediment into the Gulf of Mexico (Brown et al. 2005; Meade and Moody 2010). A dramatic consequence of this interaction is the annual formation of the Gulf of Mexico “Dead Zone”, an area of hypoxia which develops as a result of nutrient flux from the Lower Mississippi River and has significant ecological and economic impacts (Rabalais et al. 2002; Turner et al. 2008). The Mississippi River watershed is one of the most agriculturally-influenced drainages in the world (Goolsby et al. 1999; David et al. 2010), and agriculture contributes to this nutrient flux as non-point sources of pollution. However, land-use patterns across the Mississippi River drainage are
heterogeneous, and therefore it is important to consider ecological processes based on its major
sub-basins.

Unlike most large rivers, which are fed by many small tributaries, the majority of the Mississippi River discharge is comprised of a few very large feeder rivers. The Lower Mississippi River watershed can be divided into five or six major sub-basins based on these major tributaries (Turner and Rabalais 2003, 2004). Each of these watersheds has its own distinct physiochemical profile based on their specific drainage areas. For instance, the Ohio River contributes twice as much discharge to the Lower Mississippi as the other tributaries, whereas land surface area and suspended sediment input are dominated by the Missouri River (Turner and Rabalais 2004). These watersheds also differ in land use patterns, for example the Upper Mississippi watershed has approximately twice the agricultural coverage as the Missouri and Ohio (Brown et al. 2005; Hill et al. 2011), and contributes more nutrients to the greater system relative to the other tributaries (Alexander et al. 2000, 2008; David et al. 2010). Together the properties of each of these rivers form the overall biogeochemical profile of the main system, establishing the basis of energy available for biological activity throughout the Lower Mississippi River and its influence in the Gulf of Mexico. Bacterial communities, as well as the rates at which nutrients are processed by these communities, may vary on a sub-basin basis, which may have important implications for the biogeochemical cycles throughout the entire Mississippi River ecosystem.

Currently there have been few studies on bacterial community structure in large rivers. Successional patterns of bacterial assemblages have been observed on the Changjiang River, which indicated that the community changed gradually and continuously longitudinally along the river, and that these changes were consistent over a two year period (Sekiguchi et al. 2002).
Molecular approaches were also used to identify changes in bacterial community composition along the Danube River and its major tributaries (Winter et al. 2007). The bacterial community in the Danube changed gradually, but there was also a distinct shift in community structure which was related to a phytoplankton bloom (Winter et al. 2007).

While these studies were the first to use molecular approaches to elucidate microbial community structure in large rivers, advances in research technology now allow a far more detailed description of natural bacterial assemblages. In particular, high-throughput DNA sequencing techniques can now be used to identify thousands of distinct bacterial species in a single sample of water. The high yield of high-throughput sequencing relative to earlier approaches is important in order to adequately sample total bacterial populations, which are often extremely dense, with freshwater communities averaging $10^6$ prokaryotic cells mL$^{-1}$ (Whittman et al. 1998). These high-throughput techniques, such as 454 pyrosequencing and Ion Torrent sequencing, can determine the nucleotide sequence of part of a specific gene, such as the 16S rRNA gene, in environmental DNA samples. These sequences can then be compared to databases of known sequences in order to determine the identity of bacterial populations (e.g. Fuhrman et al. 2008). High-throughput approaches allow researchers to identify many bacterial species within a community and describe patterns in the relative abundance of major groups of bacteria. In aquatic ecosystems, high-throughput sequencing approaches have described complex patterns in bacterial assemblages, particularly along environmental gradients including temperature and salinity (Lozupone and Knight 2007; Furhman et al. 2008).

There is a more substantial literature base for extracellular enzyme activity and its relation to nutrient processing in rivers. Enzyme activity in environmental samples can be estimated through the use of artificial substrates that are attached to colorometrically or
fluormetrically labeled tags (Sinsabaugh et al. 1997; Jackson et al. 2013). These approaches have been used to evaluate environmental enzyme activity in many aquatic ecosystems (e.g. Jackson et al. 1995; Martinez et al. 1996; Boschker and Cappenberg 1998), including rivers (Chappell and Goulder 1995; Sinsabaugh and Foreman 2001; Karrasch et al. 2003; Tiquia 2011). Physiochemical and nutrient profiles in rivers are directly connected to processes in their watershed. These properties are linked to microbial activity, and land-use patterns across the drainage basin have been shown to significantly influence enzyme profiles within their corresponding river or stream (Harbott et al. 2005, Williams et al. 2012). In regard to large rivers, Hill et al. (2010) determined that sediment enzyme activity could be used as an indicator of nutrient limitation in the rivers of the Upper Mississippi watershed (the Ohio, Missouri, and Upper Mississippi). While much of the enzyme activity in rivers occurs in sediment, planktonic bacterial communities (bacterioplankton) can also show appreciable enzyme activity. Planktonic microbial assemblages can exist as either free-living cells or as cells bound to suspended particles, and these different communities can have distinct enzyme profiles (Chappell and Goulder 1995; Simon 1995; Lehman and O’Connell 2002). It is therefore important to consider environmental heterogeneity within these microhabitats, as well as between watersheds, in order to develop a clear picture of the microbial diversity and enzymatic processes across a fluvial network.

The purpose of this project was to evaluate bacterial community structure and functional attributes in the major tributaries of the Lower Mississippi River. This network of large rivers provides an excellent natural system for studying microbial biogeography. The unidirectional flow of lotic systems provides a traceable path of transmission, and the distinct physiochemical profile of each tributary offer a range of environmental variation. These characteristics will allow
for a more descriptive analysis of the underlying dynamics that shape microbial community structure in large rivers, for which general patterns are still unclear (Dolan et al. 2005). Currently there have been no high-throughput sequence based analyses of bacterial communities in any of these rivers, and while sediment enzyme activity has been described in some of these rivers (Hill et al. 2010); planktonic enzyme activity has not been examined. Planktonic enzyme activity within the main channel may offer a more discrete view of microbial function relative to the immediate conditions of the river itself, as these microbial communities are likely to be continuously exposed to prevailing biochemical conditions. Furthermore, the connections between bacterial community structure and extracellular enzymatic activity have yet to be examined in large rivers, and certainly not with next-generation sequencing techniques. I used Ion Torrent sequencing of the 16S rRNA gene to describe bacterial assemblages in the Arkansas, Illinois, Ohio, Missouri, Tennessee, and Upper Mississippi rivers as they approach their confluences with the larger system. Functional characteristics of these communities were evaluated by determining the activity of a suite of enzymes related to the acquisition of carbon, nitrogen, phosphorus, and sulfur. Planktonic and particle-associated communities were separated by filtration, in order to identify the importance of microhabitat differences between watersheds. Community structure and enzyme activity was related to environmental variables, water chemistry, chlorophyll a, and prokaryotic cell density. Finally, the enzyme data was related to the bacterial communities to identify patterns in diversity, functionality, and environmental parameters. I hypothesized that the bacterial assemblages of each river would be distinct from one another, enzyme activity would differ between watersheds, and that these differences would primarily be driven by the specific chemical profiles of each river.
METHODOLOGY

Site Selection and Sampling

Water samples were collected from six major tributaries of the Lower Mississippi River. Three sites on each of the Arkansas (Ark), Ohio (Ohi), Missouri (Miz), Tennessee (Ten), Upper Mississippi River (UpM) were sampled from July 6 to 17, 2012 (Fig. 1). An additional site on the Tennessee River was added to have a second representative site in Kentucky Lake, a large reservoir formed just before the Tennessee merges with the Ohio River. A surplus in supplies at the end of sampling allowed for two sites on the Illinois River (Ill) to be surveyed for community samples, but not enzyme activity. Rivers will subsequently be identified by the letter codes seen above, and sites within a river will be labeled alphabetically beginning with the most upstream stream site (i.e. sites furthest from the confluence will be labeled A). The Kentucky Lake samples are referred to as Ten C and Ten D. Sites within each river were separated by 50 – 100 river km, and were selected to be just above the confluence in order to describe the cumulative profile of each major tributary as it contributes to the greater Mississippi River system. All samples were collected between 8:00 am and 2:00 pm, and all sites on an individual river were surveyed on the same day beginning with the most downstream site and working upriver, expect for two distinct regions of the Tennessee River which were treated as individual systems. Environmental parameters (water temperature, dissolved oxygen content, conductivity, total dissolved solids, oxidation-reduction potential, pH, depth) were recorded at each site using a YSI multiprobe instrument. For sampling, the main channel of the river at each site was accessed by kayak, and water samples were collected at approximately 0.5 m depth. Each collection consisted of three
sterilized 1 L bottles for water physiochemistry and chlorophyll a concentration, and three sterilized 500 mL bottles for prokaryotic cell density, microbial extracellular enzyme activity, and bacterial community structure. Filled sample bottles were stored in river water during transportation to the laboratory (1-2 h).

**Water Physiochemistry**

Upon return to the laboratory, 200 mL subsamples from each of the three physiochemistry water samples were filtered through Whatman GF/F filters, and filters and filtrates frozen (-20 °C) until subsequent determination of chlorophyll a concentration and dissolved chemistry. At that time, filtrates were thawed and soluble reactive phosphorus (PO$_4$-P) and total dissolved phosphorus (TDP) were measured following the spectrophotometric methods described by Wetzel and Likens (2000). Dissolved organic carbon (DOC) was measured by nondispersive infrared absorption using a Teledyne-Tekmar Apollo 9000 Total Organic Carbon Analyzer, and nitrate (NO$_3$-N) was analyzed using a Dionex ion chromatograph. Ammonium (NH$_4$-N), nitrite (NO$_2$-N), and total Kjeldahl nitrogen (TKN) were analyzed using a Lachat autoanalyzer at the USDA Agricultural Research Service laboratory in Oxford, MS. In addition to water chemistry, turbidity of unfiltered water was measured using a Hach model 2100A turbidimeter. Chlorophyll a concentration was extracted from frozen filters by shaking in 90% ammonium hydroxide-buffered acetone for 24 h at 5 °C. Absorbance of chlorophyll a was read at 655 and 750 nm following the procedure of Wetzel and Likens (2000).
Fig. 1 Map of the fifteen site locations (shaded circles) in the Arkansas, Missouri, Ohio, Tennessee, and Upper Mississippi rivers that were sampled in this study. Water physiochemistry, prokaryotic cell density, extracellular enzyme activity, and bacterial community samples were collected at each site. Physiochemistry and community samples were also collected at an additional upstream site in the Upper Mississippi River, and two sites on the Illinois River near it’s confluence with the Mississippi River. Locations of larger U.S. cities are shown for reference.
**Prokaryotic Cell Density**

Two 8 mL subsamples from each of the three microbial water samples collected at each site were used to determine prokaryotic cell density microscopically using 4’,6-diamidino-2-phenylindole (DAPI) fluorescent staining. One 8 mL subsample was filtered through a 3 micron pore-size polycarbonate filter to remove larger suspended particles; the other subsample was not filtered. Unfiltered and filtered samples were preserved in a 2% (final concentration) buffered formaldehyde solution within 2 h of collection and stored at 4 °C (Crump et al. 1998). Bacteria were then processed for enumeration following the procedures of Velji and Albright (1986). Tetrasodium pyrophosphate detergent (2 mL) was added to each sample, which was then sonicated (20 W, 5 s) six times to break up particles. DAPI stain (20 μL of 100μg/mL) was then added to each sample, which was stained at 4 °C for 10 minutes. After staining, 0.4 μL of the final solution was filtered through a black 0.22 micron polycarbonate filter, which was used for enumeration. Twelve fields of view were averaged to estimate prokaryotic cell density (cells mL⁻¹) using an Olympus AX70 microscope under epifluorescence.

**Microbial Enzyme Activity**

The potential activity of enzymes related to the acquisition of carbon (β-glucosidase, cellobiohydrolase (CBH), N-acetyl-β-D-glucosaminidase (NAGase), and leucine aminopeptidase), nitrogen (leucine aminopeptidase and NAGase), phosphorus (phosphatase), and sulfur (sulfatase) was measured fluorometrically using substrates linked to methlyumbelliferyl or coumarin (Sigma-Aldrich Corp., St. Louis, MO, USA). Subsamples (100 mL) of each microbial water sample were filtered through 3-micron pore-size polycarbonate filters and 0.22-micron pore-size polyethersulfone filters to remove larger suspended particles.
and microbial cells, respectively. Filters were frozen for subsequent DNA extraction. Extracellular enzyme activity was determined in each of these filtrates and in the unfiltered samples. Enzyme assays were generally conducted within 4 - 5 h of sample collection, with samples maintained in river water at ambient temperature until processing. Assays followed the procedure demonstrated by Jackson et al. (2013). All samples, quenching, and substrate controls were replicated in triplicate on the same 96-well microplate and fluorescence measured at 12, 24, 36, and 60 minute intervals with an excitation wavelength of 350 nm and an emission wavelength of 450 nm using a BioTek FLx800 fluorometer (BioTek Instruments, Winooski, VT, USA). The fluorescence of leucine aminopeptidase, phosphatase, and β-glucosidase reactions were also measured after a 6 minute incubation as the activity of these enzymes has been found to peak rapidly (Jackson et al. 2013). Microplates were covered during incubation, which was at room temperature (24 °C). Assays for all enzymes were run concurrently. We report extracellular enzyme activity as the greatest observed value, determined after accounting for substrate, standard, and quenching controls. Potential activity is expressed as the amount of substrate consumed per unit time and volume (nmol⁻¹h⁻¹mL⁻¹).

**Bacterial Diversity**

Samples were prepared for community analysis within two months following field sampling. PowerWater DNA Isolation Kits (MoBio, Carlsbad, CA) were used to extract DNA from the 3-micron and 0.22-micron filters, which were used to characterize particle associated and free-living communities, respectively. Extraction products were subsequently amplified using PCR, which was necessary because of low yields of environmental DNA from some samples. The primers used for the PCR (Bac8/28f and Univ1492r) select for nearly the full
length bacterial 16S rRNA gene. The conditions for this PCR are described in Jackson et al. (2001), and consist of a 95°C denaturation step (2 minutes), followed by 23 cycles of 95°C (1 minute), 45°C (1 minute) and 72°C (2 minutes), and a final elongation step at 72°C (7 minutes). Samples were stored at -20°C until preparation for sequencing. Amplicons were diluted to an equal concentration of DNA per sample, and tagged 16S rRNA gene V4 variable region primers 515f and 806r (Caporaso et al. 2011) used in a single step 30 cycle PCR using the HotStar Taq Plus Master Mix Kit (Qiagen, USA) under the following conditions: 94°C (3 minutes), followed by 28 cycles of 94°C (30 seconds), 53°C (40 seconds), and 72°C (1 minute), and a final elongation step at 72°C (5 minutes). Sequencing was performed on an Ion Torrent PGM at a commercial facility (MR DNA, Shallowater, TX, USA) following manufacturer’s guidelines.

The output data following sequencing was converted into FASTA files for each of the samples, and sequencing quality files were generated. These files were processed using the bioinformatics software MOTHUR (Schloss et al. 2009) following the general community analysis procedures of Schloss et al. (2011). A series of modifications to the sequences prior to statistical analyses were used to ensure that these approaches used the most appropriate data to describe these communities. Sequences were denoised, and the barcodes and primers were removed as these do not contribute to the genetic diversity of the samples. Sequences containing ambiguously called bases were removed, along with any sequences with repeats of > 6 base pairs (bp) and lengths < 150 bp long. Sequences were then trimmed to retain a base quality score average of at least 25, and aligned using the SILVA rRNA database (Pruesse et al. 2007). Artifacts from PCR, such as chimeras, can make up a significant portion of sequences in high-throughput approaches. Chimeric sequences were identified using UCHIME (Edgar et al. 2006) and removed. The remaining sequences were classified based on the Greengenes classification.
Non-bacterial sequences, such as chloroplasts, were then removed from the data set and the remaining sequences were organized into operational taxonomic units (OTUs) based on 3% threshold of sequence dissimilarity (i.e. sequences with <3% difference were clustered into the same OTUs).

**Statistical Analyses**

In order to maintain an even sampling across all samples, the statistical analysis of β-diversity (between samples) was based on the average of 1000 subsamples of the lowest number of sequences in one sample (4,239 sequences). A distance matrix was constructed using theta dissimilarity index, which considers proportional abundances of shared and non-shared OTUs (Yue and Clayton 2005). This distance matrix was used to construct a tree diagram, ordinated using non-metric multidimensional scaling (NMDS), and conduct analysis of similarity (ANOSIM) in order to identify differences between groups of samples, such as particle-associated vs. free-living between the watersheds. Variation within the NMDS plots was then correlated (Pearson) to specific OTUs, as well as environmental and enzymatic data. \( S_{\text{Chao1}} \), Shannon, and inverse Simpson indices were used to described α-diversity (within sample), which were differentiated using ANOVA and post-hoc Tukey’s HSD in R.

Analysis of variance (ANOVA) was used to determine site-based patterns in water chemistry, chlorophyll a concentration, prokaryotic cell density, and extracellular enzyme activity. Tukey’s HSD was performed on all significant relationships to determine site-to-site relationships. Log-transformed odds ratios were used to interpret filtration effects on cell density and enzyme activity using unfiltered water as the total density/activity, and evaluated using ANOVA and Tukey’s HSD. A principle component analysis (PCA) based on the mean unfiltered
activity of each enzyme was performed to evaluate site dissimilarity based on overall enzymatic profile. Distance-based redundancy analysis (RDA) was then used to elucidate the influence of mean water physiochemistry, chlorophyll α, prokaryotic cell density, and environmental parameters on site dissimilarity. All parameters were z-transformed, and variables that were very highly correlated (R < 0.9) with other variables were removed. A Mantel test on the unfiltered enzyme profile dissimilarity matrix and geodesic site distance was performed to determine geographic significance.
RESULTS

*Water Physicochemistry and Cell Density*

Water chemistry differed between rivers in terms of both organic material and inorganic nutrient concentrations (Table 1). Dissolved organic carbon (DOC) concentration was more than threefold greater in the Upper Mississippi River than in the Tennessee or the Ohio, and significantly greater than in the Missouri and Arkansas (Tukey’s HSD, $\alpha = 0.05$). Total Kjeldahl nitrogen (TKN) was also significantly greater throughout the Upper Mississippi sites than in any other river, and lowest in the Tennessee River samples taken in Kentucky Lake (Table 1). Different sites sampled within the same river tended to have similar concentrations of DOC and TKN. Ammonium ($\text{NH}_4$-N) concentration was significantly higher throughout the Ohio River and the riverine Tennessee sites, particularly at sites Ohi C and Ten B, which had $\text{NH}_4$-N concentrations an order of magnitude greater than any site in the Upper Mississippi, Missouri, Arkansas, or Kentucky Lake (Table 1). Oxidized nitrogen ($\text{NO}_x$-N) concentration was much higher in the Upper Mississippi River, then the Missouri and Ohio and lowest in the Arkansas River and Kentucky Lake (Table 1). Most of the oxidized nitrogen was in the form of nitrate ($\text{NO}_3$-N), though $\text{NO}_3$-N and nitrite ($\text{NO}_2$-N) concentrations had similar patterns throughout the rivers. While overall oxidized nitrogen was highest in the Upper Mississippi River, the Ohio River had comparable levels of $\text{NO}_2$-N, and the most downstream site on the Ohio River had considerably higher nitrite concentrations than all other sites (Table 1). Total phosphorus (TDP) concentration followed a similar pattern to carbon and nitrogen; highest in the Upper Mississippi, then the Missouri, then the remaining sites, which also paralleled inorganic phosphate ($\text{PO}_4$-P)
measurements (Table 1). Notably, the sites with elevated ammonium concentrations (Ohi C and Ten B) also had high levels of phosphate relative to the other sites in their systems (Table 1).
<table>
<thead>
<tr>
<th></th>
<th>1.7 ± 0.4 µg</th>
<th>1.5 ± 0.4 µg</th>
<th>1.3 ± 0.4 µg</th>
<th>1.2 ± 0.4 µg</th>
<th>1.1 ± 0.4 µg</th>
<th>1.0 ± 0.4 µg</th>
</tr>
</thead>
<tbody>
<tr>
<td>77 ± 4.8 mg</td>
<td>92 ± 4.8 mg</td>
<td>94 ± 4.8 mg</td>
<td>96 ± 4.8 mg</td>
<td>99 ± 4.8 mg</td>
<td>101 ± 4.8 mg</td>
<td>103 ± 4.8 mg</td>
</tr>
<tr>
<td>126 ± 4.8 µg</td>
<td>132 ± 4.8 µg</td>
<td>138 ± 4.8 µg</td>
<td>144 ± 4.8 µg</td>
<td>150 ± 4.8 µg</td>
<td>156 ± 4.8 µg</td>
<td>162 ± 4.8 µg</td>
</tr>
<tr>
<td>13.7 ± 0.5 g</td>
<td>14.3 ± 0.5 g</td>
<td>14.9 ± 0.5 g</td>
<td>15.5 ± 0.5 g</td>
<td>16.1 ± 0.5 g</td>
<td>16.7 ± 0.5 g</td>
<td>17.3 ± 0.5 g</td>
</tr>
<tr>
<td>9.4 ± 0.5 g</td>
<td>9.8 ± 0.5 g</td>
<td>10.2 ± 0.5 g</td>
<td>10.6 ± 0.5 g</td>
<td>11.0 ± 0.5 g</td>
<td>11.4 ± 0.5 g</td>
<td>11.8 ± 0.5 g</td>
</tr>
<tr>
<td>0.9 ± 0.5 g</td>
<td>0.9 ± 0.5 g</td>
<td>0.9 ± 0.5 g</td>
<td>0.9 ± 0.5 g</td>
<td>0.9 ± 0.5 g</td>
<td>0.9 ± 0.5 g</td>
<td>0.9 ± 0.5 g</td>
</tr>
</tbody>
</table>

**Table I** Mean (±SE) of chemical and biological variables throughout the major tributaries of the Lower Mississippi River with Tukey's HSD groupings.
Fig. 2 Prokaryotic cell density in water samples collected from fifteen sites in the Arkansas (Ark), Missouri (Miz), Ohio (Ohi), Tennessee (Ten), and Upper Mississippi (UpM) rivers. Sites on a given river are designated alphabetically beginning with the most upstream location. Counts are shown as mean (± standard error) cell density per mL for unfiltered water (total of gray and white bars) and 3-micron filtered water (white bars only), and grouped according to Tukey’s HSD.
When all sites were considered, DOC and TKN were positively correlated \((R = 0.86, n=45\) for all physiochemistry and biological correlations), but the relationship of DOC to nitrate was weaker \((R = 0.58)\), and non-existent to nitrite or ammonium. DOC was also positively correlated with total phosphorus \((R = 0.77)\) and \(\text{PO}_4\)-P \((R = 0.81)\), as was TKN \((R = 0.69\) and 0.77, for total phosphorus and \(\text{PO}_4\)-P, respectively). The Missouri river had significantly higher concentrations of chlorophyll a than all of the other rivers (Table 1). Interestingly, the two sites that had the highest \(\text{NH}_4\)-N and phosphate concentrations (Ohi C and Ten B, Fig. 1) had significantly lower levels of chlorophyll a than all other sites (Table 1).

Patterns in total prokaryotic cell density were primarily driven by the Missouri river, which had significantly higher cell counts than all other rivers (Tukey’s HSD, \(\alpha = 0.05\); Fig. 2). A similar, although not as distinct pattern was also evident in samples which had been filtered through a 3-micron filter to remove particles (Fig. 2). Overall cell density was correlated with both filtered cell density \((R = 0.80)\) and chlorophyll a \((R = 0.81)\), but not with any of the measured physiochemical variables \((R<0.4\) for all). The ratio of free-living cell density to total cell density ranged from 0.75 in the riverine Tennessee River samples to 0.53 in the Missouri River. However, analysis of log-transformed odds ratios indicated there was no significant difference in this ratio between rivers or sites \((\text{ANOVA, } p>0.1)\).

**Extracellular Enzyme Activity**

Bulk enzymatic activity varied between the different extracellular enzymes with leucine aminopeptidase showing the highest activity \((\text{mean unfiltered activity of } 0.301 \text{ nmol h}^{-1} \text{ mL}^{-1})\), followed by phosphatase \((0.059 \text{ nmol h}^{-1} \text{ mL}^{-1})\), and then \(\beta\)-glucosidase \((0.019 \text{ nmol h}^{-1} \text{ mL}^{-1})\), NAGase \((0.019 \text{ nmol h}^{-1} \text{ mL}^{-1})\), sulfatase \((0.013 \text{ nmol h}^{-1} \text{ mL}^{-1})\), and cellobiohydrolase \((0.010\)
nmol h$^{-1}$ mL$^{-1}$). Activity of each enzyme also varied between sites, and between unfiltered water, 3-micron filtrate, and 0.22-micron filtrate (Fig. 3). Unfiltered, particulate-free, and cell-free activity of all enzymes varied substantially from site to site (ANOVA, p<0.01; Fig. 3). However, much of this significance is driven by a large variation in activity between just a few sites, which in some cases were even on the same river (Table 2). Technical issues on the final day of sampling forced the removal of Upper Mississippi site A from the enzyme analysis.

The relative proportion of particle- and cell-free activity varied for each enzyme (Figure 3). Particle associated activity accounted for approximately half of the total activity of β-glucosidase (52.4%), NAGase (49.0%), and leucine aminopeptidase (46.3%), but less for phosphatase (34.6%) and cellobiohydrolase (23.7%), and much less for sulfatase (8.7%). The relative percentage of cell-related activity was lower than particle associated activity for β-glucosidase (32.1%), leucine aminopeptidase (29.4%), and NAGase (18.4%), but about the same as the particle associated activity for phosphatase (32.5%), cellobiohydrolase (28.3%), and sulfatase (8.9%). Environmental or “free” enzyme activity (after the removal of cells) was detected for all enzymes and at nearly every site. This varied from less than a quarter of β-glucosidase (15.5%) and leucine aminopeptidase (24.4%) activity, to greater proportions of activity for NAGase (32.6%), phosphatase (32.9%), cellobiohydrolase (48.0%), and sulfatase (82.5%).

The only consistent correlation between the activity of different enzymes, regardless of filtration was between cellobiohydrolase and NAGase, (R = 0.75 – 0.95, n = 45). Activity of phosphatase and β-glucosidase was positively correlated in unfiltered samples (R = 0.64), but not following filtration. Relating enzymatic activity to water chemistry showed a weak positive correlation between cellobiohydrolase activity and DOC (R = 0.61) and a negative correlation
Fig. 3 Extracellular enzyme activity in water collected from fifteen sites in the Arkansas (Ark), Missouri (Miz), Ohio (Ohì), Tennessee (Ten), and Upper Mississippi (UpM) rivers. Sites on a given river are designated alphabetically beginning with the most upstream location. Values are the mean (± standard error) activity of leucine aminopeptidase (a), phosphatase (b), β-glucosidase (c), N-acetyl-β-D-glucosaminidase (d), cellobiohydrolase (e), and sulfatase (f) determined in unfiltered water (dark gray bars), 3 micron filtered water (light gray bars), and 0.22 micron filtered water (white bars) at each site.
between β-glucosidase activity and ammonium (R = -0.65). All other correlations between enzymatic activity and chemical variables were minimal (R < 0.5). β-glucosidase activity was weakly correlated with chlorophyll a concentrations (R = 0.53), but enzyme activity was not otherwise related to either chlorophyll a concentration or cell density.

In addition to identifying spatial and chemical relations to specific enzymes, a multivariate approach, principal component analysis (PCA), was used to combine each activity into an overall enzymatic profile (Fig. 4). Two axis ordination of the PCA explained 66.1% of the total variation in overall enzyme profiles (PCA axis 1 = 40.1%, PCA axis 2 = 26.0%) (Fig. 4). Most of the difference between enzyme expression in the ordination was along PCA axis 1. Phosphatase and β-glucosidase activity had similar influences on the ordination, expressed along PCA axis 2. The other correlated activities, NAGase and cellobiohydrolase, also have similar vectors which went along both axes. Overall there was less variability within samples from a site than between sites. Samples for the Upper Mississippi River were diametrically opposed to the Missouri River along PCA axis 1 (Fig. 4). The Kentucky Lake samples (Ten C and D) grouped more similarly with the Missouri River than the riverine Tennessee samples. Marginal

Redundancy analysis (RDA) was then used to relate these patterns to the physiochemical data. Based on the RDA, highly significant (p<0.01) variables in the total model were chlorophyll a, dissolved oxygen, depth, TKN, NH$_4$-N, and NO$_3$-N; significant (p<0.05) variables were temperature and oxidation-reduction potential; marginally significant (p<0.1) variables were pH, total organic carbon, and total phosphorus (Fig. 4).
Table 2: Tukey's HSD groupings of leucine aminopeptidase (LEU), phosphatase (PHOS), β-glucosidase (β-GLU), N-acetyl-β-D-glucosaminidase (NAG), cellobiohydrolase (CBH), and sulfatase (SULF) activity throughout the major tributaries of the Lower Mississippi River in bulk (total), particle-free (<3 micron), and cell-free (<0.22 micron) water samples. Group rankings based on means and ordered alphabetically.

<table>
<thead>
<tr>
<th></th>
<th>Arkansas</th>
<th>Missouri</th>
<th>Ohio</th>
<th>Tennessee</th>
<th>Upper Mississippi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A  B  C</td>
<td>A  B  C</td>
<td>A  B  C</td>
<td>A  B  C</td>
<td>A  B</td>
</tr>
<tr>
<td><strong>LEU</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22 µm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 µm total</td>
<td>bc abc abc</td>
<td>ab abc abc</td>
<td>c abc abc</td>
<td>a abc abc ab abc abc</td>
<td></td>
</tr>
<tr>
<td>3 µm total</td>
<td>a abcd bed</td>
<td>a cde abcd</td>
<td>de cde de</td>
<td>abc de ab cde cde abed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h bcd abc</td>
<td>a cde abcd</td>
<td>fgh abc gh</td>
<td>bcd gh ab abcd</td>
<td>efg def</td>
</tr>
<tr>
<td><strong>PHOS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22 µm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 µm total</td>
<td>a cd abcd</td>
<td>d ab cd</td>
<td>bcd d d</td>
<td>abc abcd cd abcd</td>
<td>abcd abcd</td>
</tr>
<tr>
<td>3 µm total</td>
<td>bc cde a</td>
<td>cde ab bcd</td>
<td>cde de e</td>
<td>bcd bc cd cde cd cd</td>
<td></td>
</tr>
<tr>
<td></td>
<td>def ef ab</td>
<td>cde bc cde</td>
<td>ef a f</td>
<td>bcd ef ef def</td>
<td>def def</td>
</tr>
<tr>
<td><strong>β-GLU</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22 µm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 µm total</td>
<td>a cd cd</td>
<td>abc bed</td>
<td>bcd ed</td>
<td>cd d ab cd</td>
<td>ab cd</td>
</tr>
<tr>
<td>3 µm total</td>
<td>ab bcd abc</td>
<td>abc cde bcd</td>
<td>def a ef</td>
<td>abcd f bcd bcd</td>
<td>cd cde</td>
</tr>
<tr>
<td><strong>NAG</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22 µm</td>
<td>c ef cde</td>
<td>cd ed ef</td>
<td>ef def cdef</td>
<td>b cdef ef ef</td>
<td>def a</td>
</tr>
<tr>
<td>3 µm total</td>
<td>ab b ab</td>
<td>abc ab b</td>
<td>de de de</td>
<td>b de de e</td>
<td>de a</td>
</tr>
<tr>
<td></td>
<td>c de c</td>
<td>de cd e</td>
<td>de de de</td>
<td>b de de e</td>
<td>de a</td>
</tr>
<tr>
<td><strong>CBH</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22 µm</td>
<td>bc cd bed</td>
<td>cd bcd d</td>
<td>cd cd cd</td>
<td>b bcd cd d</td>
<td>bcd a</td>
</tr>
<tr>
<td>3 µm total</td>
<td>c cde cd</td>
<td>de cde e</td>
<td>cde cde e</td>
<td>b de de de</td>
<td>cde a</td>
</tr>
<tr>
<td></td>
<td>b cde b</td>
<td>cdef cde def</td>
<td>ef boce f</td>
<td>bc ef ef ef</td>
<td>cde a</td>
</tr>
<tr>
<td><strong>SULF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22 µm</td>
<td>b ef a</td>
<td>ef def ef</td>
<td>c c c c</td>
<td>c cd ef cd</td>
<td>a cde</td>
</tr>
<tr>
<td>3 µm total</td>
<td>b g a</td>
<td>fg defg ef</td>
<td>be cd cde</td>
<td>cde cdef ef g</td>
<td>a cde</td>
</tr>
<tr>
<td></td>
<td>b e a</td>
<td>de cde cde</td>
<td>bcd b bc</td>
<td>b bcd cde cde</td>
<td>a bcd</td>
</tr>
</tbody>
</table>
**Fig. 4** Distance-based redundancy analysis of overall enzyme profile related to environmental and physiochemical variables. Ordination based on principle component analysis of mean activity of leucine aminopeptidase (LEU), phosphatase (PHOS), β-glucosidase (βGLU), cellobiohydrolase (CBH), N-acetyl-β-D-glucosaminidase (NAG), and sulfatase (SULF) in unfiltered water samples. Individual sample scores are connected to centroid labeled by river (Arkansas (A), Ohio (O), Missouri (M), Tennessee (T), and Upper Mississippi (U)) and site (A-D, starting with most upstream location). Enzyme scores are bolded and underlined. Vectors based on Pearson correlations show the environmental variables related to the ordination: cell density, chlorophyll, conductivity (cond), depth, dissolved oxygen (DO), pH, carbon (DOC, TKN), nitrogen (NH₄-N, NO₂-N, and NO₃-N) and phosphorus (TP).
Community Analysis

A total of 108 filters were collected (54 selecting for particle-associated communities and 54 selecting for planktonic communities) over the course of this study. Nine of these samples produced a poor yield following the DNA extraction and amplification steps and could not be used for the community analysis. The remaining 99 samples were successfully sequenced using Ion Torrent PMG sequencing. A total of 942,204 potential valid sequences were produced from the sequencing run with an average of 156 bp. Quality analyses in MOTHUR identified 50,689 potential chimeras and 176,395 chloroplast-derived sequences. After removing these sequences, a total of 715,120 sequences were classified into 30,211 OTUs. The majority of these OTUs (66.8%) were represented by a single sequence.

Spatial patterns between samples were evaluated based on theta-based dissimilarity matrix and visualized in a tree diagram (Figure 5). The most distinct pattern was between the particle-associated and planktonic samples across all rivers (ANOSIM, \( P<0.001 \)). NMDS ordination demonstrates a clear separation of these samples and the OTUs that significantly shape this distribution (Figure 6). The most influential group driving this division was OTU01, which classified as *Pelagibacter* in the SAR-11 group of the *Alphaproteobacteria*. This group accounted for over 16% of sequences in the free-living samples but only 0.5% of particle associated sequences (Table 3). Other taxa that were significantly more represented in planktonic communities than on particles were OTU15 (*Gammaproteobacteria* – Ellin339, 4.1% of planktonic sequences vs. 0.5% of particle-associated sequences), OTU51 (*Bacteriodetes* – *Algoriphagaceae*, 6.7% vs. 0.1%), OTU06 (*Actinobacteria* – ACK-M1 cluster, 8.4% vs. 0.6%), OTU08 (*Betaproteobacteria* – *Methylophilales*, 3.8% vs. 0.1%), OTU37 (*Betaproteobacteria* – unclassified, 2.0% vs. 0.8%). Particle associated samples also had a taxa which were more
Fig. 5 Tree diagram of bacterioplankton community β-diversity in the Arkansas (green), Illinois (purple), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers based on Ion Torrent sequencing of the 16S rRNA gene. Particle-associated samples are in the shaded portion of the tree, and free living communities are in the un-shaded portion. Mean diagram from of 1000 subsampling iterations based on lowest number of site sequences (n=4,239).
Fig. 6 NMDS ordination of bacterial communities from the majority tributaries to the Lower Mississippi River based on 16S rRNA sequences that were particle associated (filled circles) or free living (open circles). These groups were analyzed with all taxa (a, Stress=0.24, ANOSIM P<0.001) and excluding *Cyanobacteria* (b, Stress=0.26, ANOSIM P<0.001). Major OTUs are described in Table 3.
prevalent than in free living samples, the most important being multiple OTUs (02, 03, 04, 05, 13, 16, 20, 27, 29, 36) which classified as being related to the *Cyanobacteria* group *Prochlorococcus*. These combined OTUs accounted for nearly 30% of the particle associated sequences and 17% of the free living sequences. The high frequency of *Cyanobacteria* compared to other groups may have had a substantial role in separating planktonic and particle-associated communities; however, removal of this group from the dataset yielded similar patterns in community structure including the important OTUs driving those patterns (Figure 6b).

Particle associated and free living assemblages were extremely distinct from one another, so in order to describe differences between rivers these samples were separated for further community analysis (Figure 7). Bacterioplankton communities separated by individual tributaries (Figure 7a, ANOSIM all comparisons P<0.001). Pairwise comparisons between rivers indicated that the relationships which were not significantly different (P<0.001) were between the Arkansas and Illinois (P<0.136) and the Arkansas and the Ohio (P=0.003), likely due to the high site variability in the Arkansas (Figure 7a). The most evident separations were between the Upper Mississippi and the Missouri communities, and between the Ohio and the Tennessee communities. Both *Pelagibacter* and *Prochlorococcus* (OTUs 03, 04, and 16) significantly influenced these patterns. River-specific assemblages remained significant after removing *Cyanobacteria* sequences (Figure 7b, ANOSIM all comparisons P>0.001), with the same pairwise relationships (Arkansas communities did not separate from the Ohio or Illinois communities, P= 0.017 and 0.04, respectively).

Particle associated bacterioplankton assemblages also differentiated between rivers (Figure 7c, ANOSIM all comparisons P<0.001). Similar to patterns in the planktonic communities, the Upper Mississippi and Missouri samples separated primarily across NMDS
Fig. 7 NMDS ordinations of free-living bacterial communities from the Arkansas (green), Illinois (purple), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers based on 16S rRNA sequences. Communities were analyzed using all taxa (a, Stress=0.18) and removing Cyanobacteria (b, Stress=0.18). Major OTUs are described in Table 3.
Fig. 8 NMDS ordinations of particle associated bacterial communities from the Arkansas (green), Illinois (purple), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers based on 16S rRNA sequences Communities were analyzed using all taxa (a, Stress=0.24) and after removing Cyanobacteria (b, Stress=0.26). Major OTUs are described in Table 3.
dimension 1 while the Ohio and Tennessee samples separated mainly across NMDS dimension 2 (Figure 7c). The only non-significant (ANOSIM, P<0.001) relationship observed was between the Arkansas and Illinois communities (P=0.01). Again the Prochlorococcus-like OTUs were important in shaping this distribution. Specifically, one OTU (OTU04) was prevalent in the Upper Mississippi while a different OTU (OTU02) was more influential in the Illinois and Ohio samples. Removal of Cyanobacteria did not change the overall patterns in the particle associated communities, which were still distinct between rivers (Figure 7d, ANOSIM all comparisons P<0.001). Each pairwise comparison between river samples was significant (ANOSIM, P<0.001), with the exception of the Arkansas and the Ohio samples (P=0.013), where the most downstream site on the Arkansas grouped with the Ohio communities (Figure 7d). This grouping was largely driven by OTU32 (Fusobacteria - Fusobacterium). As with all of the other sampling groups, the most distant communities were between the Upper Mississippi and Missouri and the Tennessee and Ohio samples (Figure 7d).

Throughout the entire dataset, Cyanobacteria were the most prevalent bacterial lineage (30.4% of all sequence, Figure 8) and the majority of cyanobacterial sequences were associated with Prochlorococcus-like populations (77% of cyanobacterial sequences). Cyanobacteria were proportionally more abundant in the particle-associated sequences (40.5%) than the planktonic samples (19.3%). The relatively importance of this lineage differed between rivers. For instance, the Ohio and Missouri had distinctly lower proportions of Cyanobacteria in particulate samples, and in planktonic assemblages which were dominated by Bacteroidetes and Alphaproteobacteria (Figure 8). Both Proteobacteria and Bacteroidetes were also prevalent phyla (22.0% and 14.6%, respectively). In free living communities, Alphaproteobacteria were the most abundant subphyla of the Proteobacteria (17.8% of all sequences), which almost completely (>93%)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pelagibacter</td>
<td>(Alphaproteobacteria)</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.7</td>
<td>1.0</td>
<td>1.3</td>
<td>0.1</td>
<td>0.4</td>
<td>0.7</td>
<td>1.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Prochlorococcus</td>
<td>(Cyanobacteria)</td>
<td>25.8</td>
<td>38.0</td>
<td>25.9</td>
<td>47.1</td>
<td>6.7</td>
<td>16.7</td>
<td>9.6</td>
<td>14.1</td>
<td>9.4</td>
<td>16.7</td>
<td>9.6</td>
</tr>
<tr>
<td>Uncl. ACK</td>
<td>M1 (Actinobacteria)</td>
<td>6.6</td>
<td>0.1</td>
<td>3.3</td>
<td>1.4</td>
<td>7.3</td>
<td>1.1</td>
<td>0.8</td>
<td>0.1</td>
<td>3.3</td>
<td>1.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Uncl. Verrucomicrobiae</td>
<td>(Verrucomicrobia)</td>
<td>0.4</td>
<td>1.0</td>
<td>0.1</td>
<td>0.4</td>
<td>5.8</td>
<td>12.5</td>
<td>3.2</td>
<td>12.5</td>
<td>3.2</td>
<td>12.5</td>
<td>3.2</td>
</tr>
<tr>
<td>Uncl. Methylophilales</td>
<td>(Betaproteobacteria)</td>
<td>4.5</td>
<td>0.1</td>
<td>3.1</td>
<td>0.1</td>
<td>3.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Uncl. Terrimonaceae</td>
<td>(Bacteroidetes)</td>
<td>1.3</td>
<td>0.1</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>0.6</td>
<td>4.6</td>
<td>0.7</td>
<td>4.6</td>
<td>0.7</td>
<td>4.6</td>
</tr>
<tr>
<td>Uncl. WPS-1 group</td>
<td>(Planctomycetes)</td>
<td>2.8</td>
<td>12.7</td>
<td>0.6</td>
<td>1.6</td>
<td>0.3</td>
<td>1.4</td>
<td>0.2</td>
<td>1.4</td>
<td>0.2</td>
<td>1.4</td>
<td>0.2</td>
</tr>
<tr>
<td>Uncl. Saprospirales</td>
<td>(Bacteroidetes)</td>
<td>3.9</td>
<td>4.0</td>
<td>4.1</td>
<td>4.9</td>
<td>4.8</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Ellin339</td>
<td>(Gammaproteobacteria)</td>
<td>2.7</td>
<td>0.2</td>
<td>7.0</td>
<td>0.5</td>
<td>5.4</td>
<td>0.9</td>
<td>2.9</td>
<td>0.9</td>
<td>2.9</td>
<td>0.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Uncl. Cyanobacteria</td>
<td>17, 30, 33, 47, 52, 62, 69, 74, 80</td>
<td>4.5</td>
<td>8.6</td>
<td>18.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.6</td>
<td>2.8</td>
<td>0.6</td>
<td>2.8</td>
<td>0.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Microcystis</td>
<td>(Cyanobacteria)</td>
<td>0.1</td>
<td>1.3</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Uncl. Flavobacteriales</td>
<td>(Bacteroidetes)</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Clostridium bifermentans</td>
<td>(Firmicutes)</td>
<td>0.3</td>
<td>1.7</td>
<td>0.2</td>
<td>1.2</td>
<td>0.1</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Uncl. Planctomycetacia</td>
<td>(Planctomycetes)</td>
<td>0.4</td>
<td>2.7</td>
<td>0.9</td>
<td>3.8</td>
<td>1.1</td>
<td>7.8</td>
<td>0.5</td>
<td>3.8</td>
<td>0.5</td>
<td>3.8</td>
<td>0.5</td>
</tr>
</tbody>
</table>

**Table 3** Major taxa in free-living (Free) and particle-associated (Part) bacterial communities of major tributaries to the Lower Mississippi River. Numbers represent the mean % of that taxonomic group in sequences (mean ± SE) obtained from each river (n=5 to 9). Taxa that were unclassified according to Greengenes are noted as "Uncl."
consisted of sequences related to the *Pelagibacter*/SAR11 group. *Deltaproteobacteria* were the only subphylum to be proportionally more abundant in particle associated assemblages than free living bacterioplankton.

Some of the rarer phyla also differed between river bacterial communities (Figure 8). *Verrucomicrobia* and *Planctomycetes* were proportionally more prevalent in the Missouri particulate and free living communities than in other rivers. Planktonic assemblages in the Ohio and the Tennessee rivers had higher proportions of *Actinobacteria* in the sequence dataset compared to the other samples. These communities in the Ohio also had a noticeably higher prevalence of sequences classified as *Acidobacteria*, which were an order of magnitude greater than in the other assemblages. In total, 50 distinct bacterial phyla were classified; approximately half being candidate phyla although many of the other lineages were represented by just a handful of sequences. Greengenes taxonomy schemes classified 7% of the total sequences as belonging to domain *Bacteria* but not to a specific phylum.

Within-sample (α) diversity had highly significant variation (ANOVA, p<0.01) between microhabitat (particle associated vs. planktonic), rivers, and microhabitat within rivers for all three diversity indices (Figure 9). Each method suggested that particle associated assemblages were more diverse than planktonic communities (mean values of $S_{\text{Chao1}} = 2,389$ vs. 1,406, inverse Simpson = 38.5 vs. 23.1, and Shannon = 4.95 vs. 4.34, respectively). $S_{\text{Chao1}}$ identified the greatest diversity in the particle associated assemblages of the Missouri (mean = 3,447) and Upper Mississippi rivers (mean = 2,897), and the least amount of diversity in the planktonic communities found in the Ohio River (mean = 990). These patterns were similar to the Shannon index (Figure 9) and the score of these two indices were correlated (R=0.79). However the inverse Simpson index yielded somewhat different patterns, rating particle associated diversity in
Fig. 9 Community structure of free-living and particle-associated bacterial assemblages in the major tributaries of the Lower Mississippi River based on 16S rRNA gene sequences. Major groups accounting for >1% in any sample type are shown. Bars are mean proportions from 5–9 replicate samples in a river with a mean of 7,223 sequence reads per sample.

% Bacterial sequences in community

 Particle-associated
Free-living

 Mississippi
Tennessee
Ohio
Illinois
Arkansas
Mississippi
Tennessee
Ohio
Illinois
Arkansas

Other

Predominant genera

Other
the Tennessee River as high as in the Missouri River (Figure 9b). Free living bacterioplankton communities in the Ohio River also ranked as the least diverse assemblages. Inverse Simpson index scores correlated well with those from the Shannon index (R=0.84), but not as strongly to the $S_{\text{Chao1}}$ results (R=0.47).

Patterns in bacterial community structure between rivers may be related to specific physiochemical profiles (Figure 10). Variables related to phosphorus content ($\text{PO}_4$-$\text{P}$ and TDP) and organic C/N content (DOC and TKN) were highly correlated with each other (R=0.99 and 0.91, respectively), so only one parameter was used. Both of these had significant influence on particle associated and free living communities with and without *Cyanobacteria*. The influence of phosphorus concentration, which was highest in the Illinois and Upper Mississippi rivers (Table 1), was mainly expressed across NMDS dimension 1 (Figure 10). DOC had similar patterns to phosphorus throughout the rivers and in the NMDS ordinations, though it tended to be strong when *Cyanobacteria* were included (Figure 10a, c). NO$_3$-$\text{N}$ also tended to have a similar influence on the community structure as phosphorus and DOC, whereas NH$_4$-$\text{N}$ typically opposed these three parameters along NMDS dimension 1(Figure 10).

Nutrient-related parameters tended to be expressed across the first NMDS axis, while generally environmental and chemical variables were more associated with the second NMDS axis (Figure 10). Temperature and pH could be related to river based difference, and were higher in the Illinois and Arkansas than in the Tennessee. Parameters related to suspended particles (turbidity and chlorophyll a) had a noticeable influence on particle associated communities, but not free living bacterioplankton (Figure 10c, d). These environmental variables can be used to summarize communities across rivers: higher TDP/NO$_3$-$\text{N}$/DOC (Upper Mississippi), higher TDP/NO$_3$-$\text{N}$/DOC plus turbidity/chlorophyll a (Illinois), higher turbidity/chlorophyll a
Fig. 10 Diversity index scores for free living and particle associated bacterial assemblages from the major tributaries of the Lower Mississippi River based on 16S rRNA gene sequences. $S_{\text{Chao1}}$ (a), inverse Simpson (b), and Shannon (c) indices were normalized to the sample with the lowest number of sequences ($n=4,239$ sequences). Values are the mean (±SE) values calculated from 5-9 replicate river samples. Groups based on Tukey’s HSD (p<0.05).
(Missouri), higher NH$_4$-N with lower TDP (Ohio), higher NH$_4$-N with lower TDP/NO$_3$-N/temperature/pH (Tennessee), and higher temperature/pH with lower TDP/NO$_3$-N (Arkansas).

In addition to covariance between bacterial assemblages and environmental factors, it is also possible that distinct communities could have unique enzymatic profiles. These patterns between rivers were assessed for free living and particle-associated communities. Of the six enzymes assayed, four were found have significant covariance (p<0.05) with patterns in the free living communities (Fig. 13). Leucine aminopeptidase and phosphatase pulled in the same direction along both NMDS axes, although leucine aminopeptidase had a greater overall influence. Patterns in a sequence type classified as *Clostridium biermentans* (OTU21) were expressed similarly to these enzyme activities, which may indicate a relationship between abundance of these bacteria and leucine aminopeptidase and phosphatase activity. NAGase was expressed primarily along the first NMDS axis towards the Missouri River region of the ordination, with a similar draw as a specific *Prochlorococcus*-like OTU (OTU05). Conversely, sulfatase activity was almost exclusively expressed along the second NMDS, directed towards the Upper Mississippi River region of the ordination and with a different *Prochlorococcus*-like OTU (OTU02).

Only two enzymes were significantly related (p<0.05) to particle associated communities patterns (Fig. 14). Celllobiohydrolase was expressed across both of the NMDS axes and directed toward the Upper Mississippi River samples, where one site (UpMC) had significantly higher activity than other sample sites (Table 2). Sulfatase activity was expressed directly along the second NMDS axis. Similar to the ordination with free living communities, the sulfatase vector in the particle-associated community NMDS also followed an OTU identified as
Fig. 11 NMDS ordinations of free-living bacterial assemblages sampled from the Arkansas (green), Illinois (purple), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers, and the covarying physicochemical variables. Communities were analyzed using all taxa (a, Stress = 0.18) or excluding Cyanobacteria (b, Stress = 0.18). Vectors display environmental factors that were significantly (p<0.05) correlated with NMDS axes scores for each plot, and the direction and extent of influence.
Fig. 12 NMDS ordinations of particle associated bacterial assemblages sampled from the Arkansas (green), Illinois (purple), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers, and the covarying physicochemical variables. Communities were analyzed using all taxa (a, Stress = 0.24) or excluding Cyanobacteria (b, Stress = 0.26). Vectors display environmental factors that were significantly (p<0.05) correlated with NMDS axes scores for each plot, and the direction and extent of influence.
**Fig. 13** NMDS ordination of free-living bacterial assemblages from the Arkansas (green), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers based on 16S rRNA sequences and enzymatic activity (Stress = 0.18). OTUs classified in Table 3. Length of vectors is proportional to significance (p<0.05).
Fig. 14 NMDS ordination of particle associated bacterial assemblages from the Arkansas (green), Missouri (yellow), Ohio (red), Tennessee (orange), and Upper Mississippi (blue) rivers based on 16S rRNA sequences and enzymatic activity (Stress = 0.24). OTUs classified in Table 3, length of vectors displays significance (p<0.05).
Prochlorococcus, though this was yet another type (OTU04), and these vectors were oriented towards the Arkansas River communities.
DISCUSSION

High-throughput sequencing allows for a far more in-depth analysis of bacterial community structure than earlier molecular or culture-based approaches. The sequencing analysis in this study represents a much more detailed view of bacterial communities compared to other studies on bacterial biogeography in large rivers. For example, the DGGE and clone library approach used by Winter et al. (2007) on a study of the Danube River yielded 43 distinct OTUs over a similar number of samples (98) to this study. A study on the Changjiang River described only 14 OTUs in community analysis, although this was across only 15 samples (Sekiguchi et al. 2002). By comparison, the analyses of community structure in this study (which identified over 30,000 OTUs) benefit from having three orders of magnitude more taxonomic information compared to the previous descriptions of large river bacterial assemblages. While high throughput sequencing studies have been conducted in other lotic systems, they have been limited to lower order streams (Yergau et al. 2012; Besemer et al. 2013) or to specific regions of larger rivers (Fortunato et al. 2012). This study also represents the first comparative study of bacterial community structure across a network of large rivers, and the first to link community structure to functional aspects of riverine bacteria. The major tributaries of the Lower Mississippi River offer an excellent system for describing these patterns and relating community structure to environmental variable and functional attributes. Each of these rivers has distinct drainage basins which vary in climatic and land-use patterns (Goolsby 1999; Turner and Rabalais 2003, 2004), and the influence of these environmental differences impacts the structure of their bacterial communities.
One of the clearest distinctions in assemblage structure was between the Missouri and Upper Mississippi River. The Upper Mississippi River basin is among the most agriculturally influenced watersheds in the world (David et al. 2010). This anthropogenic influence greatly contributes to the nutrient profile of the Upper Mississippi, which had nitrogen and phosphorus concentrations approximately twice as high as the Missouri River. Water chemistry and nutrient availability can directly impact functional aspects of microbial communities, such as in the activity of extracellular enzymes (Hoppe et al. 1988; Romani and Sabater 2000; Shackle et al. 2000). Molar ratios of total nitrogen (sum of TKN and NO₃-N) and total phosphorus for the Upper Mississippi (23:1) and the Missouri (24:1) were similar, and suggested moderate P-limitation (relative to N) in each of these rivers when compared to Redfield’s suggested ratio (16:1 for N:P; Redfield 1958). The similar N:P ratios might also explain why the two rivers showed similar phosphatase activity. The ratio of total carbon to nitrogen was higher in the Upper Mississippi River (4:1) than in the Missouri River (3:1), as was its total C:P ratio (98:1 versus 67:1), which suggests that the Missouri is more C-limited. Interestingly both of these rivers had similar enzymatic profiles, which were dominated by peptidase and phosphatase, not the glucosidases. Hill et al. (2010) found greater levels of β-glucosidase than leucine aminopeptidase or phosphatase in the Missouri, however the relative differences were much greater in the upper reach than they were in the lower reach (which still included many sites that were upstream from this study). One of the major enzymatic differences between the two systems was that the Upper Mississippi, not the Missouri, had greater overall NAGase activity. Given that NAGase can be tied to nitrogen mineralization, this could indicate a greater degree of nitrogen limitation, although that would be surprising given the agricultural influence of the Upper Mississippi River drainage basin.
Patterns in enzyme activity could be related to the types of nutrient inputs from their watersheds. Carbon and nitrogen inputs can have significant influences on glucosidase and peptidase activity in aquatic environments (Foreman et al. 1998; Shackle et al. 2000; Harbott and Grace 2005). Foreman et al. (1998) conducted substrate addition experiments to observe the enzymatic response to varying nutrient input. As expected, the addition of protein induced leucine amino peptidase production (Foreman et al. 1998). Agricultural practices, particularly livestock land use, could increase protein concentrations in these rivers which may explain the higher levels of leucine aminopeptidase activity. Foreman et al. (1998) also found that glucosidase activity actually increased when glucose was added, which would usually be thought of reducing β-glucosidase activity. This is likely due to tradeoffs in C, N, and P acquisition. While bacterial productivity is generally carbon-limited, the acquisition of C is restricted by the need to produce enzymes that acquire N and P from organic sources. Therefore the activity of C-acquiring enzymes can be dependent on the activity of N-acquiring enzymes, which in turn can be dependent on the activity of P-acquiring enzymes (Sinsabaugh and Moorhead 1994).

However, extracellular enzyme activity in rivers, including those in this study, has also been found to be influenced by factors other than water chemistry (Hill et al. 2010), such as the structure of the bacterial communities that produce these enzymes. Variation of enzyme activity between lotic systems can be associated with differences in bacterial community structure (Findlay and Sinsabaugh 2006), so that relating enzyme activity profiles to bacterioplankton community structure may help explain patterns in microbial activity between these rivers.

The bacterial communities in the Upper Mississippi and Missouri rivers were distinctly different from one another in both free living and particle associated assemblages. Both microhabitats had a substantially higher proportion of *Prochlorococcus*-like OTUs in the Upper
Mississippi River, although removing *Cyanobacteria* from the analysis did not change the overall structural patterns. Planktonic communities in the Upper Mississippi samples also tended to have higher proportions of sequences similar to *Planctomycetacia* and *Gemmatimonadales* than those taken from the Missouri. Particle associated differences between these two rivers included an relatively higher abundance of *Methylocystis*-like sequences in the Missouri and more *Verricomicrobiae*-like sequences in the Upper Mississippi. Based on the NMDS ordinations, these differences in community structure were generally related to total carbon and ammonia (which was highly correlated with TKN), indicating that nutrient concentrations were influencing these assemblages. Among the free-living community, the Missouri samples appeared to be more associated with leucine aminopeptidase and phosphatase activity than the Upper Mississippi, despite not have a significantly higher overall activity. Additionally, sulfatase activity was important in distinguishing these communities. This may indicate that C-limitation has not produced an overall glucosidase-dominated enzymatic profile in the Missouri because of lower nitrogen and phosphorus inputs, whereas an excess of primary nutrients in the Upper Mississippi shifts metabolic constraints more towards secondary elements such as sulfur, and that these chemical regimes have shaped the bacterial assemblages. For the particle associated communities, there was a strong relationship between cellobiohydrolase activity and the Upper Mississippi River samples. Cellobiohydrolase activity was the highest in the Upper Mississippi River, although it is interesting that this covariance was tied to the particle associated communities. This may be the result of more diverse particulate C-sources in this river because of increased agricultural inputs, which led to greater metabolic investment into recalcitrant substrates such as cellulose. The remnants of corn and other agricultural crops could represent an important source of particulate cellulose.
A second distinct difference in community structure was between the bacterial assemblages in the Ohio and the Tennessee rivers. Both of these rivers were less nutrient-rich than the other major tributaries to the Lower Mississippi, and the Ohio River was more P-limited based on Redfield ratios. The surface water stoichiometry of the Tennessee River was comparable to Upper Mississippi and Missouri, with the lowest N to P ratio observed (21:1), whereas the N to P ratio was much greater in the Ohio River (46:1). The balance between P and N was expressed in structural differences between these communities. In the free living assemblages where *Cyanobacteria* were removed, the Ohio communities were expressed similarly to the ammonia vector, whereas the Tennessee communities were placed towards the phosphate vector. Interestingly this pattern is reversed in the particle associated communities, suggesting that these communities could be experiencing different nutrient regimes which may influence their metabolic activities. Particle associated bacteria may well acquire nutrients from the particles themselves, and not be as dependent on dissolved nutrients as free living cells, which might explain these different patterns. Further emphasizing differences between the two microhabitat types, there appeared to be a strong relationship between the free living communities in the Tennessee River and NAGase activity, however the particle associated communities in this system were not associated with any C-related enzyme activity. Particle associated communities are likely influenced by the chemical profiles of the particles themselves, as well as the surrounding water, which may drive functional and structural differences from free living assemblages. Indeed, differences in bacterial productivity between particle associated and free living communities have been observed in the Lower Mississippi River, and have been linked to the availability of particulate phosphorus (Ochs et al. 2010).
About one quarter of the free living sequences from the Ohio River samples were classified as a single OTU, which was identified as a member of the *Alphaproteobacteria* similar to a subgroup of the marine SAR11/*Pelagibacter* clade. The abundance of this single taxon was responsible for much of the structural differences between the Ohio and Tennessee assemblages. There is little known about the ecology of the freshwater SAR11 subgroup, referred to as LD12 (Bahr et al. 1996; Zwart et al. 1998), however its abundant worldwide distribution may mean that it has functional importance throughout freshwater systems (Logares et al. 2009; Newton et al. 2011). The difference in the proportional abundance of this OTU between the Ohio and Tennessee communities may be connected to the specific nutrient profiles of these rivers, and could have a significant influence on their biogeochemical properties. In marine environments, *Pelagibacter* is able to dominate bacterial populations because it is particularly well adapted to nutrient and energy limitation (Morris et al. 2002). In contrast, the LD12 freshwater cluster appears to be more successful in highly productive lakes (Salcher et al. 2011) and may show a high affinity for amino acid assimilation (Salcher et al. 2011), presumably through the production of peptidases. This OTU was detected in every river and was the most abundant OTU overall, which reinforces the idea of its widespread distribution. In contrast, the second most detected taxon, the aforementioned *Prochlorococcus*-like sequences, showed substantial phylogenetic variation between rivers, with different OTUs being more dominant in some river systems. Based on these data, it appears that nutrient regimes in the Ohio and Tennessee rivers have different impacts on the free living and particle associated assemblages, and that these influence are expressed in structurally and functionally distinct communities.

The contrast between particle associated and free living assemblages was evident in every river, and, overall, was the most striking pattern across the entire dataset. Similar patterns in
microhabitat community differentiation have been observed in many aquatic environments (DeLong et al. 1993; Acinas et al. 1999; Crump et al. 1999). Particle associated communities tended to be more diverse than free living assemblages, and in some cases had distinct relationships to environmental and functional parameters. Interactions with surfaces can induce physiological changes in bacterial populations and can have important consequences for their overall ecology (Van Loosdrecht et al. 1990). Increased habitat heterogeneity on suspended particles compared to the open water likely provides more functional niches for bacterial populations to occupy. Suspended particles may also offer increased habitat stability in large rivers, where high flow rates and turbulence have significant influence on nutrient regimes and productivity (Ochs et al. 2013). Adhesion to surfaces also allows for the formation of biofilms, which have distinct assemblage structure and ecological properties compared to planktonic cells. While bacterial biofilms could not be directly observed in this study, their community dynamics (Brümmer et al. 2000; Hullar et al. 2006) and functional importance (Battin et al. 2001) have been described in lotic systems. These factors make suspended sediments in rivers hotspots for microbial activity (Luef et al. 2007; Ochs et al. 2010), which greatly contribute to the overall microbial activity in the major tributaries of the Lower Mississippi River.

This project was designed using replication approaches that ensured that all observed patterns were an accurate representation of microbial processes. Replication has been identified as an important but often under-appreciated aspect of microbial research (Prosser 2010). Replication was prioritized in this study by collecting multiple samples at each site, and by surveying multiple sample sites on each river. Variation among sites is evident on the NMDS ordinations. Rather than averaging this variability across sites, it was included in each analysis as a means of conservatively describing community patterns. Even with the increased error of
treating all of the samples from a single river as replicates (despite being from different sites), there were still apparent differences in both free living and particle associated communities throughout each river. Incorporating site-variability strengthens these observations. Always sampling the most downstream site first and moving against the current ensured that the same “packet” of water was not re-surveyed at different sites. Sites were selected just above (20-150 km) each tributaries’ confluence allowing for a description of the terminal microbial community and environment as it enters the larger system. The one exception was the Tennessee River, which was sampled twice in the lentic-like Kentucky Lake region (near its confluence the Ohio River) and twice in a more river section upstream. Despite being environmentally distinct from each other, these two reaches had significantly similar bacterial assemblages, which suggests that characteristics across the drainage basin are more influential than local habitat differences.

The method used here for separating particles and cells has been used in many other studies (e.g. Crump et al. 1999; Besemer et al., 2005). While it is mostly effective, it is worth noting that filtration is not perfect. As particles collect on the filter, they begin to clog pores and restrict smaller objects from passing through. This would likely have resulted in planktonic cells being included in the particle associated assemblages. Additionally, it is possible that small tears in the 3 micron filter allowed particle associated cells to be incorporated into the free living communities or that vacuum filtration stripped some loosely attached cells from particles. These methodological issues would falsely increase the amount of similarity between free living and particle associated assemblages. Thus, the striking differences observed between these two types of communities in this study are actually conservative and may underestimate the differences between them.
The over 30,000 distinct OTUs that were classified in this study may be indicative of amplification and/or sequencing errors although the influence of these potential errors was minimized. Clustering OTUs using a 97% 16S rRNA sequence similarity (Stackbrandt and Goebel 1994) reduces amplification artifacts, and subsampling for community analyses decreases the influence of rare OTUs, as they are less likely to be selected than more abundant sequences (Gihring et al. 2012). While subsampling reduces the total number of potential observations, the number used (4,239) is an appropriate size for analyzing α- and β-diversity in microbial communities (Lundin et al. 2012). Furthermore, community analyses were based on theta index distances, which accounts for relative abundance and further reduces the impact of rare sequences. Based on the results, distinct communities were identified in six rivers and two microhabitats, and the average number of OTUs per sample was similar to other high-throughput sequencing studies (Yergau et al. 2012; Besemer et al. 2013). Ion Torrent was selected over 454 pyrosequencing because the lower cost allowed for all replicates to be sequenced and the relatively high yield of reads provides a broad view community structure. The tradeoff is short read lengths, although in this case they were still long enough (average of 156bp) for microbial community comparison (Liu et al. 2007). While this limits the depth to which sequences could be classified, it also represents another conservative methodological approach for describing the community differences, which again strengthens the observed patterns.

Across the main tributaries of the Lower Mississippi River, bacterioplankton associated with particles were clearly different from free living assemblages. Both of these community types showed distinct patterns based on river of origin, which were related to the physiochemical characteristics of their specific river. While a multivariate approach suggested that some rivers may differ in overall enzymatic profiles, it was difficult to identify specific patterns in enzyme
activity. This is despite the rivers not only being environmentally distinct, built also supporting different bacterioplankton communities. Thus, while microbial communities between these rivers differ structurally, their functional attributes, at least in terms of potential enzyme activity, appear to be much more similar. This could be the result of different bacteria occupying similar functional niches, or the result of the more widespread taxa having a greater influence in substrate acquisition. Identifying how these patterns change spatially and temporally and how they are expressed when these rivers merge is a potential area for future research.
LIST OF REFERENCES


Redfield AC (1958) The biological control of chemical factors in the environment. American Scientist 46(3): 230-221


VITA

Education

University of Mississippi
Master of Science (Biology, 4.0 GPA, ec. 2013)
2012-2013

Michigan State University
Bachelor of Science (Ecology, Evolution, and Organismal Biology, 3.4 GPA)
2007-2011

Honors and Awards

ASLO Early Career Travel Grant ($500)
2013

University of Mississippi Department of Biology Travel Grant ($200)
2013

Most Outstanding Teaching Assistant in Biology Award
2010

Michigan State University – Lyman Briggs College Dean’s List
2010-2011

Peer-Reviewed Publications

Millar JJ, Jackson CR, Payne JT, Ochs CA. 2013. Patterns in extracellular enzyme activity throughout the major tributaries to the Lower Mississippi River. *Biogeochemistry* (In preparation)


Professional Presentations

Millar JJ, Payne JT, Ochs CA, Jackson CR. Bacterial community structure in major tributaries of the Lower Mississippi River is driven by habitat differences at regional and micro-scales. American Society of Microbiology Meeting, New Orleans, LA, October 2013


Lampkin AL, **Millar JJ**, Payne JT, Ochs CA, Jackson CR. Particle-associated and soluble phosphatase activity in major rivers of the lower Mississippi Basin. American Society for Limnology and Oceanography Aquatic Sciences Meeting, New Orleans, LA, February 2013

**Millar JJ**, Payne JT, Ochs CA, Jackson CR. Microbial extracellular enzyme activity in large rivers of the Mississippi River Basin. American Society of Microbiology Meeting, Starkville, MS, October 2012

### Teaching Experience

<table>
<thead>
<tr>
<th>Institution</th>
<th>Year</th>
<th>Role Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Mississippi</td>
<td>2012</td>
<td>Microbiology Lab Teaching Assistantship - Prepared lab materials, taught sections, wrote quizzes, graded reports, quizzes, and exams</td>
</tr>
<tr>
<td>The Island School</td>
<td>2011</td>
<td>Statistics - Designed and taught school-wide course focused on research statistics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mangrove Ecology - Co-taught field-based research course culminating in poster symposium for representatives of the Bahamian Ministry of Health</td>
</tr>
<tr>
<td>Michigan State University</td>
<td>2010-2011</td>
<td>Applied Research Techniques Teaching Assistantship - Assisted in designing and teaching course on private sector research culminating in presentation for ConAgra scientists</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Molecular Biology Lab Teaching Assistantship - Taught two open and two closed lab sections, prepared lab materials, graded quizzes, reports, and exams</td>
</tr>
</tbody>
</table>

### Research Experience

<table>
<thead>
<tr>
<th>Institution</th>
<th>Year</th>
<th>Role Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Mississippi</td>
<td>2012-2013</td>
<td>Research Assistantship - NSF-funded project on microbial biogeography of the Mississippi River system: field planning and sampling, molecular research techniques (enzyme assays, DNA extraction, PCR, DGGE), bioinformatics analyses</td>
</tr>
<tr>
<td>Cape Eleuthera Institute, The Bahamas</td>
<td>2011</td>
<td>Research Fellowship - Influence of climate change on mangrove ecology in the Caribbean: field sampling, wet-lab manipulation experiments on fish respirometry, regional outreach</td>
</tr>
<tr>
<td>Mote Marine Laboratory</td>
<td>2010</td>
<td>REU Internship - Coastal Florida fisheries habitat ecology and monitoring: extensive field sampling, manipulation of large 10-year dataset, symposium presentation</td>
</tr>
<tr>
<td>Michigan State University</td>
<td>2008-2011</td>
<td>Research Associate - Early-life demographic patterns across animal phyla: collection of peer-reviewed literature, extrapolation and meta-analyses</td>
</tr>
<tr>
<td>Field/Lab Technician</td>
<td></td>
<td>Soil ecology in agricultural systems: field sampling, soil entomology and biogeochemistry research (microscopy, live culture, C:N pyrolysis)</td>
</tr>
</tbody>
</table>