Ecology Appendix S1

Metabolic insight into bacterial community assembly across ecosystem boundaries

Nathan I. Wisnoski[1], Mario E. Muscarella[1,2], Megan L. Larsen[1,3], Ariane L. Peralta[1,4], Jay T. Lennon[1]

SUPPLEMENTAL METHODS

Surface water samples were obtained approximately every 25 m along a longitudinal transect from the lacustrine zone near the dam to the two major streams feeding University Lake (Fig. 1). We used a Quanta Hydrolab (OTT, Kempton, Germany) water sonde to measure temperature, dissolved oxygen, pH, and conductivity of the epilimnion at each site. We collected water samples at each site for biological and chemical analyses. We measured total phosphorus (TP) concentrations using the ammonium molybdate method (Wetzel and Likens 2000).

Sample preparation — For aquatic samples, we extracted total nucleic acids (RNA and DNA) from the filters using the MoBio PowerWater RNA extraction kit and the DNA elution accessory kit (Carlsbad, CA) and cleaned the extracts via ethanol precipitation. We treated RNA extracts with DNase (Invitrogen) to degrade DNA prior to cDNA synthesis via the SuperScript III First Strand Synthesis Kit and random hexamer primers (Invitrogen). For soil samples, we extracted DNA with the PowerSoil DNA isolation kit (MoBio, Carlsbad, CA). Once DNA and cDNA samples were cleaned and quantified, we amplified the 16S rRNA gene (DNA) and transcript (cDNA) using barcoded primers (515F and 806R) targeting the V4 hypervariable region (Caporaso et al. 2012). We purified sequence libraries using the AMPure XP purification kit (Bechman), quantified using the Quant-it PicoGreen dsDNA kit (Invitrogen), and pooled at equal molar ratios (final concentration:10 ng per library). After pooling, we sequenced the libraries on the Illumina MiSeq platform using 250 × 250 bp paired end reads (Reagent Kit v2) at

the Indiana University Center for Genomics and Bioinformatics Sequencing Facility. Paired-end raw 16S rRNA sequences reads were assembled into contigs, quality-trimmed, and aligned to the Silva Database (version 132) (Quast et al. 2013). Chimeric sequences were detected and removed using the VSEARCH algorithm (Rognes et al. 2016). We created OTUs by first splitting the sequences based on the RDP taxonomy (Cole et al. 2009), and then binning sequences in to operational taxonomic units (OTUs) based on 97% sequence similarity. All initial sequence processing was completed using the software package *mothur* (version 1.41.1, Schloss et al. 2009).

Nestedness of aquatic OTUs — We partitioned β-diversity into turnover and nestedness components following Baselga (2010) using the 'betapart' R package (Baselga and Orme 2012). We used the Sørensen family of β-diversity metrics, and isolated the turnover and nestedness components of each aquatic sample (separated into DNA and RNA pools) relative to the three soil samples. We then averaged the turnover and nestedness fractions to produce a final mean nestedness and mean turnover value for each aquatic sample. We also addressed nestedness more simply by creating a Venn diagram of OTUs from the soil community, the total aquatic community, and the active aquatic community to visualize the taxa shared between these portions of the meta-ecosystem.