Ecological and Ontogenetic Components of Larval Lake Sturgeon Gut Microbiota Assembly, Successional Dynamics, and Ecological Evaluation of Neutral Community Processes

Shairah Abdul Razak,¹,² Kim T. Scribner¹,²,³

¹Department of Fisheries & Wildlife, Michigan State University, East Lansing, Michigan, USA
²Center for Frontier Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia
³Department of Integrative Biology, Michigan State University, East Lansing, Michigan, USA

Shairah Abdul Razak and Kim T. Scribner contributed equally to this article. Author order was determined based on the workload associated with data analysis and writing of the paper.

ABSTRACT  Gastrointestinal (GI) or gut microbiotas play essential roles in host development and physiology. These roles are influenced partly by the microbial community composition. During early developmental stages, the ecological processes underlying the assembly and successional changes in host GI community composition are influenced by numerous factors, including dispersal from the surrounding environment, age-dependent changes in the gut environment, and changes in dietary regimes. However, the relative importance of these factors to the gut microbiota is not well understood. We examined the effects of environmental (diet and water sources) and host early ontogenetic development on the diversity of and compositional changes in the gut microbiota of a primitive teleost fish, the lake sturgeon (Acipenser fulvescens), based on massively parallel sequencing of the 16S rRNA gene. Fish larvae were raised in environments that differed in water source (stream versus filtered groundwater) and diet (supplemented versus nonsupplemented Artemia fish). We quantified the gut microbial community structure at three stages (prefeeding and 1 and 2 weeks after exogenous feeding began). The diversity declined and the community composition differed significantly among stages; however, only modest differences associated with dietary or water source treatments were documented. Many taxa present in the gut were over- or underrepresented relative to neutral expectations in each sampling period. The findings indicate dynamic relationships between the gut microbiota composition and host gastrointestinal physiology, with comparatively smaller influences being associated with the rearing environments. Neutral models of community assembly could not be rejected, but selectivity associated with microbe-host GI tract interactions through early ontogenetic stages was evident. The results have implications for sturgeon conservation and aquaculture production specifically and applications of microbe-based management in teleost fish generally.

IMPORTANCE  We quantified the effects of environment (diet and water sources) and host early ontogenetic development on the diversity of and compositional changes in gut microbial communities based on massively parallel sequencing of the 16S rRNA genes from the GI tracts of larval lake sturgeon (Acipenser fulvescens). The gut microbial community diversity declined and the community composition differed significantly among ontogenetic stages; however, only modest differences associated with dietary or water source treatments were documented. Selectivity associated with microbe-host GI tract interactions through early ontogenetic stages was evident. The results have implications for lake sturgeon and early larval ecology and survival in their natural habitat and for conservation and aquaculture produc-
tion specifically, as well as applications of microbe-based management in teleost fish generally.

**KEYWORDS** community ecology, diet, freshwater fish, gut microbiota, ontogeny

One of the primary goals in community ecology is to understand community composition and diversity and the relative influences of the ecological forces underlying the patterns of distribution and abundance across space and time (1). In parallel with community ecology, the field of microbial ecology focuses on studying free-living and host-associated microbiota (2, 3). The microbial communities in host-microbial ecosystems have been extensively studied, and findings reveal that the microbiotas from different habitats are taxonomically diverse and that most have ecological importance. Yet, there are common ecological principles that govern microbial community assembly and composition (2–5).

The importance of the gut microbiota to fish has long been recognized, as microbial taxa perform important roles associated with nutritional provisioning, metabolic homeostasis, and immune defense (6–8). Compositional information pertaining to the fish gut microbiota (reviewed in references 9 and 10) has characterized microbial communities as being highly variable in composition across ontogenetic stages (11–14). Comparatively, the ecological processes governing microbial community formation, diversity, and dynamic compositional changes are typically less understood due to the challenges of detangling the complexities of many underlying factors (9, 15–19). Gut microbiota assembly usually occurs early in fish host development, involving factors such as fish host physiology, surrounding environmental conditions, and diet, providing opportunities to study microbial community formation and succession, including exchanges between the aquatic environment and the gut (11, 12, 20).

Examinations of the community composition and diversity of microbiota associated with fish gastrointestinal (GI) tracts would benefit from a statistical (5) exploration of the relative importance of four processes described by Vellend (1), specifically, selection, drift, dispersal, and speciation. These processes underlie the patterns in many other ecological communities and are also applicable to studies of microbial communities (5, 21). In our examination of the processes affecting the fish GI tract microbiota, we evaluated alternative predictions from neutrality theory, as introduced by Hubbell (22), with niche-based hypotheses (1). The ecological neutrality theory emphasizes the absence of differences in deterministic processes, such as ontogenetic shifts in the gut environment that place certain taxa at a selective advantage over others, including the per capita growth rate, death, and dispersal among species, and that assumes equal fitness across species (1, 22–24).

The neutral model in community ecology theory predicts that patterns in community composition and taxonomic diversity are likely the outcomes of stochastic processes of dispersal and drift (23, 25). By incorporating this conceptual framework together with advanced sequencing technologies, an improved understanding of the processes governing fish gut microbiota assembly during early life stages is possible. Understanding the principles associated with the bacterial colonization of GI tracts and with compositional changes across life stages will also help managers, particularly in aquaculture settings, to manipulate gut microbial communities to promote animal health, performance, and production (9, 26, 27).

Sturgeons belong to one of the oldest groups of the bony fish (Osteichthyes), and many are species of conservation concern. Of all 27 sturgeon species, lake sturgeon (Acipenser fulvescens) is the only sturgeon species endemic to the North American Great Lakes Basin and the only sturgeon in the genus Acipenser that spends its life solely in freshwater (28, 29). This species has experienced significant declines in abundance and distribution due to overfishing and the loss and degradation of its habitat (30). In recent years, sturgeon conservation aquaculture has increased greatly as part of restoration actions.

Successful lake sturgeon production can be limited due to high mortality, especially
during early ontogenetic stages. The low survival of sturgeon larvae is tied to the nutritional regimes associated with diet formulation, feeding schedule, food presentation, and preference (31, 32). After hatch, yolk-sac larvae gradually develop gastrointestinal (GI) tracts that eventually resemble the adult structures by 10 to 11 days posthatch (dph). Once the yolk-sac stage has completed (i.e., the mouth fully develops and opens), the fish start their transition to exogenous feeding, often on brine shrimp nauplii (*Artemia salina*) (33, 34). During this stage, early exposure to and the interaction of the fish gut with microbial colonists from surrounding water are possible, and shifts in the intestinal microbiota can be documented as the fish continue to develop. Early ontogenetic changes in host diet and physiology can shape gut community dynamics. Previous studies have demonstrated that the gut microbiota exhibits variations in community composition during early life stages (5, 9, 21) and that these can have important consequences during later stages (i.e., ontogenetic contingency [35–37], including disease susceptibility [38]).

Our objectives in this study were to characterize the gut microbiota of lake sturgeon during the early larval stages, at the time before the fish begin feeding, until 14 days following the onset of exogenous feeding, using Illumina MiSeq high-throughput sequencing of a portion of the 16S rRNA gene. Here, we define microbiotas/microbiomes as the assemblages of microorganisms existing in or associated with a defined habitat (in this case, the gastrointestinal tract), including active and interacting members as well transient or inactive members. The gut microbiotas colonizing lake sturgeon larvae were experimentally manipulated to quantify the associations of host factors, water supply, and diet on the lake sturgeon gut microbiota. An understanding of the ecological principles that govern the stability or transiency of fish microbial community composition and diversity is essential to successfully altering microbial communities for therapeutic and agricultural benefits. We also quantified whether the gut microbiota assembly was consistent with neutral expectations. Our findings have implications for the management of nutrition, disease, and potential probiotic use in lake sturgeon culture and characterize the dynamic relationships between the host ontogeny and the environmental epibiota associated with the temporal variability of the microbiota residing in the gut.

(This research was conducted by Shairah Abdul Razak in partial fulfillment of the requirements for a doctoral degree from Michigan State University [39].)

**RESULTS**

**Sequencing summary and fish morphometric data.** The four treatment groups evaluated in this study are as follows: fish raised in natal stream water and fed live *Artemia* fish (denoted group S), fish raised in natal stream water and fed live *Artemia* fish mixed with retentate (denoted group Sp), fish raised in groundwater and fed live *Artemia* fish (denoted group GW), and fish raised in groundwater and fed live *Artemia* fish mixed with retentate (denoted group GWp). These group designations are used throughout the paper.

Rarefaction analyses showed that sequencing efforts were consistent across replicate samples and treatment groups at a depth of 5,775 sequences per sample, as denoted by a total percentage of coverage of higher than 98%. We were able to sample a large portion of the operational taxonomic units (OTUs) and diversity present while still retaining a large number of samples within fish of each age. After quality filtering, our 16S rRNA amplicon data set produced 6,034,269 high-quality reads. In total, we observed 4,137 OTUs (2,894 OTUs when omitting singleton OTUs), defined at 97% sequence identity. From 118 samples submitted for 16S rRNA sequencing, rarefaction at 5,775 sequences per sample eliminated five samples with coverage below this sequencing coverage.

A total of 56,989 sequence reads were obtained across four retentate samples after quality filtering. Eukaryotic profiling based on the 18S rRNA amplicon data set detected 1,131 taxa overall when rarefaction was achieved at 6,000 sequences per sample, yet only 201 taxa were retained after singletons were omitted. Many of these taxa detected
could not be classified beyond the class level, as 85% of the total sequences from nonsingleton taxa remained unclassified but are known dietary items described in other larval fish. Among the predominant classes detected were Podocopida (seed and muscle shrimps; 2.54%), Peridiniphycidae (dinoflagellates; 1.81%), and Bacillariophytina (diatoms), Spirotrichea (protozoan), Synurales (algae), and Peronosporomycetes (water molds) (each at 1.1%).

Morphometric data included mean ± standard deviation (SD) fish weight (in grams) and length (in millimeters) and revealed that fish growth was consistent across all four treatments throughout the duration of the study (Tables 1 and 2). By the end of the experiment, fish raised in stream water grew significantly larger than fish raised in UV-treated groundwater. No fish health issues were detected. The survival of fish was nearly 100% in all treatments and during all developmental stages.

Ontogenetic changes of dominant bacterial taxa in lake sturgeon larva rearing. We quantified the number of sequences that were represented by each phylum from the GI tracts of all fish sampled in each developmental period. We found that the gut bacterial communities comprised 26 microbial phyla; however, the most abundant phyla, covering more than 95% of all sequences, included, in order of abundance, Proteobacteria, Firmicutes, Bacteroidetes, Actinobacteria, Acidobacteria, Planctomycetes, Chloroflexi, and Fusobacteria. General patterns of bacterial phylum-level contributions to the gut microbiota were shown with regard to treatment group and the correspond-

<p>| TABLE 1 Length and weight of sampled larval lake sturgeon from the four treatment groups across developmental stages |</p>
<table>
<thead>
<tr>
<th>Stage</th>
<th>Treatment group</th>
<th>Mean ± SE length (mm)</th>
<th>Mean wt (g)</th>
<th>Temp (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefeeding</td>
<td>GW</td>
<td>18.97 ± 0.18A</td>
<td>0.034A</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td>GWp</td>
<td>18.95 ± 0.21A</td>
<td>0.034A</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>21.05 ± 0.19B</td>
<td>0.038B</td>
<td>15.7</td>
</tr>
<tr>
<td></td>
<td>Sp</td>
<td>21.18 ± 0.13B</td>
<td>0.038B</td>
<td>15.7</td>
</tr>
<tr>
<td>1 wk after active feeding</td>
<td>GW</td>
<td>28.63 ± 0.23C</td>
<td>0.087C</td>
<td>16.1</td>
</tr>
<tr>
<td></td>
<td>GWp</td>
<td>28.92 ± 0.26C,E</td>
<td>0.084C</td>
<td>16.1</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>30.03 ± 0.19D</td>
<td>0.096D</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td>Sp</td>
<td>29.81 ± 0.35D,E</td>
<td>0.095D</td>
<td>17.2</td>
</tr>
<tr>
<td>2 wk after active feeding</td>
<td>GW</td>
<td>36.48 ± 0.41F</td>
<td>0.178E</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td>GWp</td>
<td>35.71 ± 0.30F</td>
<td>0.166E</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>40.42 ± 0.22G</td>
<td>0.225F</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>Sp</td>
<td>40.40 ± 0.40G</td>
<td>0.217F</td>
<td>19.5</td>
</tr>
</tbody>
</table>

The mortality rate during the experimental duration was too small and is thus not presented here. The capital letters indicate significant differences in the means based on a post hoc Tukey HSD test (P < 0.05).

<p>| TABLE 2 Influence of temperature of rearing water on length |</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>P value (&gt; t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>7.141</td>
<td>3.587</td>
<td>1.991</td>
<td>0.05105</td>
</tr>
<tr>
<td>GWp</td>
<td>-0.022</td>
<td>0.381</td>
<td>-0.057</td>
<td>0.95469</td>
</tr>
<tr>
<td>S</td>
<td>0.451</td>
<td>0.738</td>
<td>0.611</td>
<td>0.54326</td>
</tr>
<tr>
<td>Sp</td>
<td>0.575</td>
<td>0.826</td>
<td>0.696</td>
<td>0.48883</td>
</tr>
<tr>
<td>1 wk after active feeding</td>
<td>7.677</td>
<td>0.833</td>
<td>9.212</td>
<td>4.35E-13*</td>
</tr>
<tr>
<td>2 wk after active feeding</td>
<td>15.965</td>
<td>0.800</td>
<td>19.952</td>
<td>&lt;2.00E-16*</td>
</tr>
<tr>
<td>Temp</td>
<td>0.857</td>
<td>0.269</td>
<td>3.186</td>
<td>0.00229*</td>
</tr>
<tr>
<td>GWp × 1 wk after active feeding</td>
<td>0.315</td>
<td>0.538</td>
<td>0.585</td>
<td>0.56097</td>
</tr>
<tr>
<td>S × 1 wk after active feeding</td>
<td>0.008</td>
<td>0.628</td>
<td>0.012</td>
<td>0.99014</td>
</tr>
<tr>
<td>Sp × 1 wk after active feeding</td>
<td>-0.336</td>
<td>0.672</td>
<td>-0.5</td>
<td>0.61903</td>
</tr>
<tr>
<td>S × 2 wk after active feeding</td>
<td>-0.752</td>
<td>0.538</td>
<td>-1.396</td>
<td>0.16777</td>
</tr>
<tr>
<td>Sp × 2 wk after active feeding</td>
<td>0.146</td>
<td>0.538</td>
<td>0.271</td>
<td>0.78756</td>
</tr>
</tbody>
</table>

*Analysis of covariance indicated that the temperature of the rearing water significantly influenced the length. The residual standard error was 0.659 on 60 degrees of freedom, the multiple R² value was 0.994; the adjusted R² value was 0.993, the F statistic was 864.3 on 11 and 60 degrees of freedom; and the P value for temperature was 0.00229. NA, not applicable. *, parameter with a significant P value.
ing developmental stages (Fig. 1). Across all treatment groups, the microbiota composition shifted from communities dominated by Proteobacteria and several other phyla prior to the initiation of exogenous feeding to communities dominated by Firmicutes after feeding began. The relative abundance of all other phyla was reduced after the fish began actively feeding.

Linear discriminant analyses (LDA) effect size (LEfSe) scores were computed for taxa differentially abundant among the three ontogenetic stages (prefeeding and 1 and 2 weeks after active feeding began) (40). Among all 45 detected OTUs that were statistically and biologically different between lake sturgeon gut microbial communities compared across stages, 5 OTUs explained the greatest differences (Fig. 2). All other OTUs with significant LEfSe scores are listed in Table S1 in the supplemental material.
The abundance of two OTUs (OTU82, *Alphaproteobacteria* genus *Devosia*; OTU29, *Gammaproteobacteria* genus *Vibrio*) at 1 week after feeding was different from that at the other two stages. Three OTUs (OTU9, unclassified *Clostridiaceae* from the phylum *Firmicutes*; OTU27, *Firmicutes* genus *Jeotgalibacillus*; OTU94, *Firmicutes* genus *Carnobacterium*) were present at a statistically significantly higher relative abundance in fish gut communities at 2 weeks after the initiation of feeding.

We tested whether changes in the larval gut microbiota occurred when fish were raised on the different experimental diets and in the different rearing water sources across life stages. Over the course of 3 weeks of development, we found that the microbiota diversity varied among age cohorts. Overall, the community diversity decreased as fish transitioned to active feeding (Fig. 3A and B). Multiple-factor analysis of variance (ANOVA) quantified the sources of the variability in the alpha diversity indices (inverse Simpson diversity index \[1/D\]) and indicated statistically significant differences in diversity as a function of sampling time (\[F_{inv}\] value for inverse Simpson diversity index \[F_{inv}\] = 11.31, number of degrees of freedom \[df\] = 2, \(P < 0.001\)), whereas OTU richness did not differ across sampling times (\[F_{rich}\] value for OTU richness \[F_{rich}\] = 0.91, \(df = 2, P = 0.407\)) or treatment (\[F_{rich}\] value = 1.92, \(df = 3, P = 0.137\)). No significant interaction was found between water and food treatments for the inverse Simpson diversity index or for OTU richness (for the inverse Simpson diversity index, \[F_{inv}\] value = 1.13, \(df = 6, P = 0.358\); for OTU richness, \[F_{rich}\] value = 0.38, \(df = 6, P = 0.892\)).

**Variation in gut bacterial community profiles in association with water, diet, and time.** To visualize the relationships between the gut microbial community composition and the composition of the environmental sources (water and diet), a principal-component analysis (PCoA) was performed to analyze the samples in a reduced-dimensional space using ordination plots. The variation in community membership among the gut microbiota from all treatment groups across time periods indicated age-dependent changes in the prevalent microbial taxa (Fig. 4A). The variation in community membership was less evident when comparisons were made across environmental (water and food) microbial communities (Fig. 4B) sampled concurrently during the same time period.

Statistical analyses of the beta diversity among sampling periods (fish developmental stages) revealed significant microbial community taxonomic compositional divergence (Table 3; permutational multivariate analysis of variance [PERMANOVA] test pseudo-\(F\) value = 2.077, \(R^2 = 0.059, P < 0.001\)). We rejected the null hypothesis of no temporal (developmental) differences in the multivariate centroid location characterizing the microbial community composition. In addition, the significant results of a test for the homogeneity of multivariate dispersion (PERMDISP) \((P < 0.001)\) indicated that the compositional dispersion within each group was heterogeneous (Tables 4 and 5).

Community composition data were further analyzed to decouple the treatment effects of water source and diet from the pervasive effects of sampling time. Under the
null hypotheses, the water type and the food treatments administered were not expected to significantly affect the gut community taxonomic composition, and there would be no interaction between water and diet treatments within each time point. During the prefeeding period (sampling period 1), we found significant water treatment effects on PCoA axis 6 since the fish had not begun feeding (Table 6). As the fish transitioned to active feeding, both water and food treatments interacted significantly to affect the gut community composition after 1 week of feeding (sampling period 2; PCoA axis 1; Fig. 5A). The effect of water treatment on PCoA axis 5 was significant (Table 6). During the third sampling period, after the second week of active feeding, the gut community composition was statistically significantly different between water treatments. There was no effect of diet (PCoA axis 1; Fig. 5B). Least-squares (LS) mean values for all important axes across all times are shown in Table 6.

Focusing on the first coordinate axis during sampling periods 2 and 3, following the initiation of feeding, we investigated whether there was evidence for the co-occurrence
of taxa in the gut communities. The genus *Clostridium* (OTU001), associated with the family *Clostridiaceae* (phylum *Firmicutes*), was found to have a strong, positive correlation with the first axis during the first week of feeding (Pearson correlation $r = 0.884$).

During the second week of feeding, two other bacterial taxa (*Sarcina* [OTU002] and an

**FIG 4** Principal-coordinate analysis among microbial communities originating from the lake sturgeon larval gut across all four different treatments at three developmental stages (prefeeding and 1 week and 2 weeks after active feeding initiation) (A) and from environmental samples, including water, detritus, and *Artemia* fish (B). These environmental samples were collected at times corresponding to the times of fish sampling. Axes represent the first two principal coordinates maximizing the variance in the data (PCoA1 and PCoA2). Dissimilarity was calculated based on Bray-Curtis distances. The fish gut microbial communities from the four treatment groups are denoted S, Sp, GW, and GWp, as defined in the text. Water-GW and Water-S are water samples from groundwater and river water, respectively. *Artemia*-GW and *Artemia*-S are brine shrimp prepared using the respective water sources, and detritus was collected from a sock filter. Var, variation.

**TABLE 3** Variability of fish gut microbial communities across treatment and ontogenetic stages$^a$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No. of df</th>
<th>Sum-of-square value</th>
<th>Mean square value</th>
<th>$F$ model</th>
<th>$R^2$</th>
<th>$P$ value ($&gt;F$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
<td>2</td>
<td>1.907</td>
<td>0.954</td>
<td>2.077</td>
<td>0.059</td>
<td>&lt;0.001$^*$</td>
</tr>
<tr>
<td>Treatment</td>
<td>3</td>
<td>1.471</td>
<td>0.490</td>
<td>1.068</td>
<td>0.045</td>
<td>0.120</td>
</tr>
<tr>
<td>Stage x treatment</td>
<td>6</td>
<td>2.863</td>
<td>0.477</td>
<td>1.039</td>
<td>0.088</td>
<td>0.170</td>
</tr>
<tr>
<td>Residuals</td>
<td>57</td>
<td>26.173</td>
<td>0.459</td>
<td>1.039</td>
<td>0.088</td>
<td>0.170</td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>32.414</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$The variability of the fish gut microbial communities across treatment and ontogenetic stages was analyzed by PERMANOVA. The results revealed that the ontogenetic stages significantly influenced the gut microbial community composition for at least one sample (PERMANOVA pseudo-$F$ value $= 2.077$, $R^2 = 0.059$, $P < 0.001$, number of permutations = 1,000). $^*$, parameter with a significant $P$ value.
unclassified genus [OTU004], both of which are members of the Clostridiaceae family), showed strong correlations (for Sarcina, \( r = -0.791 \); for the unclassified Clostridiales genus, \( r = 0.743 \)) with the first coordinate axis. Another taxon in the genus Deefgea, a member of the family Neisseriaceae (phylum Proteobacteria, order Gammaproteobacteria), was also present at a high abundance and was positively correlated on the same axis (\( r = 0.779 \)). The strong correlations of these genera are reflected by the presence of taxa at high relative abundances, quantified in Tables S2 and S3.

Neutral processes are not the dominant mechanisms generating and maintaining microbial community composition during early gut microbiome assembly. LS mean analyses of PCoA axis 1 at sampling point 3 (2 weeks after the initiation of feeding, 22 dph) indicated that the dispersal of bacteria from environmental (water) sources into the gut was important in shaping the community composition in the gut (Fig. 5B; Table 6). However, the analyses did not address the following question: is the fish gut community at each developmental stage a neutral subset of the source community? We wished to distinguish between species that were detected in the gut in association with neutral processes (e.g., dispersal associated with water source) and species that were detected in the gut in association with deterministic processes that may be associated with age-dependent gut conditions. To answer this question, we applied the neutral model theory, based on the neutral model for prokaryotes of Sloan et al. (24), to investigate the developmental processes underlying gut microbiome compositional and successional changes over time. The neutral model assumes that the community composition can be explained by the dispersal of microbial taxa from the environment (water) and ecological drift (stochastic change) within the source community (23). Based on LS mean analyses, we found that water was a more important source of gut community members than diet was.

Neutral processes were detected as important contributors to microbial community formation during each sampling period (developmental stage), as indicated by significant \( P \) values (Fig. 6a to c). These processes were significant, yet they were not strongly predictive. Using the \( R^2 \) value as a measure of the goodness of fit, we determined that the gut community composition from across all treatments poorly fit a neutral model, based on the low \( R^2 \) value of the model fit (\( R^2 \) value range, 0.215 to 0.292).

Across sequential sampling periods, as the fish aged, the number of shared microbial OTUs between the gut and the water decreased (Tables 7 and 8). At the prefeeding stage, larval GI tracts contained the most neutrally dispersed taxa (number of taxa, 114), but this number declined dramatically as the fish began active feeding (number of taxa, 45). At 2 weeks after active feeding, only 38 taxa were shared between the gut and water communities, and about half of these taxa were underrepresented or overrepresented in the gut (Table 7). We conclude from these analyses that the OTUs of the gut microbiomes are likely under selection.

### Table 4: Permutation test for homogeneity of multivariate dispersions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No. of df</th>
<th>Sum-of-square value</th>
<th>Mean square value</th>
<th>( F ) model</th>
<th>Permutation</th>
<th>( P ) value (&gt;( F ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage group</td>
<td>2</td>
<td>0.025</td>
<td>0.0124</td>
<td>12.518</td>
<td>999</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Residuals</td>
<td>66</td>
<td>0.065</td>
<td>0.001</td>
<td></td>
<td></td>
<td>0.807</td>
</tr>
</tbody>
</table>

*The multivariate dispersions (number of permutations = 999) indicate that the dispersion of the distances in microbial communities among three ontogenetic stages is significantly heterogeneous. *, parameter with a significant \( P \) value.

### Table 5: Pairwise comparisons

<table>
<thead>
<tr>
<th>Stage</th>
<th>Observed ( P ) value</th>
<th>Permuted ( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefeeding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 wk after active feeding</td>
<td>0.843</td>
<td>0.002*</td>
</tr>
<tr>
<td>2 wk after active feeding</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
</tr>
</tbody>
</table>

*The observed \( P \) values are given below the diagonal, and the permuted \( P \) values are given above the diagonal. *, pairwise comparisons with a significant \( P \) value.
The taxa present in the gut are not a completely neutral subset of the taxa present in the source water communities (see Tables S4 to S6, which detail the taxa that were overrepresented, underrepresented, and neutrally represented during all three stages).

Over the time course covering all ontogenetic stages, the abundance of several taxa consistently conformed to neutral expectations (OTU062, genus *Finegoldia*), whereas other taxa were overrepresented (OTU016, genus *Pseudomonas*; OTU041, genus *Jani-bacter*) or underrepresented (OTU015, genus *Turicella*) in the fish gut in relation to their frequencies in rearing water (Table 8).

**DISCUSSION**

Many basic microbial community ecology questions concerning the dynamics of a community composition at the onset of colonization and early ontogenetic changes are of importance to understand host-microbial relationships in the wild and under domestic (i.e., fish culture) conditions. Information pertaining to the source(s) of microbial communities that establish in the fish gut, the relative importance of neutral versus nonneutral factors during community assembly, and animal host age/growth effects on gut community establishment is incompletely known.

In this study, we were interested in characterizing the development of the fish gut microbiota during important early developmental stages. Our experimental system enhanced our understanding of the factors affecting the initial colonization and development of the lake sturgeon larval gut microbiota prior to and during the critical transition from endogenous to exogenous feeding. Our research contrasts with previous studies on other sturgeon species associated with diet and gut microbiota, which were performed on either larger or older fish. For example, two studies involved Siberian sturgeon (*Acipenser baeri*) (41) and white sterlet sturgeon (*Acipenser ruthenus*) (42) used fish at the juvenile stage (average weight, 15 g to 30 g; age, more than 3 months posthatch). Other studies used fish that had been exposed to cultivation for extended periods of time (53, 79).

This research setting provided an experimental system amenable to the evaluation of environmental factors (focusing both on the rearing water type and on the diet administered) and host-associated factors that affect the compositional dynamics of the intestinal microbiota in lake sturgeon, an important aquaculture teleost fish species and a species of conservation concern.

Lake sturgeon larvae were raised in constant, controlled environments through manipulation of the water and diet and evaluated over three sampling periods from 11 to 29 dph. Our results suggest that major compositional shifts in the gut community composition corresponded to the different developmental stages. The taxonomic profiles changed across sequential larval stages, extending from before the onset of exogenous feeding (prefeeding) through 29 dph, when the fish were actively feeding and when the GI tract structures resembled the adult anatomical structures (33, 43). The findings indicate that during these early ontogenetic periods, host physiological development likely serves as a strong deterministic force directing the formation of gut community composition.

**TABLE 6** Results of analyses of significant PCoA axes for each sampling period

<table>
<thead>
<tr>
<th>Linear regression model</th>
<th>Significant treatment</th>
<th>Stage, important axis</th>
<th>Food with supplement</th>
<th>Food without supplement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Prefeeding, PCoA axis 6</td>
<td>Water type</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 wk after active feeding, PCoA axis 1</td>
<td>Interaction (water type and food type)</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 wk after active feeding, PCoA axis 5</td>
<td>Water type</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 wk after active feeding, PCoA axis 1</td>
<td>Water type</td>
<td>0.020</td>
</tr>
</tbody>
</table>

*The four axes across all three stages that showed a significant effect of either water, food, or an interaction of both water and food treatments on the microbial community composition were axis 6 for prefeeding, axis 1 and axis 5 for 1 week after active feeding, and axis 1 for 2 weeks after active feeding. Interaction plots are shown in Fig. 5.*
communities, regardless of the food type or surrounding (water) environmental communities. The temporal shift in bacterial community composition has also been documented in other studies involving zebrafish (12, 20) and rainbow trout (Oncorhynchus mykiss) larvae (44) during periods of constant diet and environmental conditions. However, while the rainbow trout studies reported a strong influence of diet type and environmental factors on the gut microbiota composition, the findings from the zebrafish studies differed from those of the rainbow trout studies (12, 20, 44, 45).
Pronounced temporal changes in microbial community composition occurred from the prefeeding period to the exogenous feeding periods. Proteobacteria dominated the gut communities during the prefeeding stage, whereas Firmicutes dominated the communities after 2 weeks of feeding. Large decreases in community taxonomic diversity were also documented between the prefeeding and postfeeding periods. These results are concordant with the findings for other fish species, which documented Proteobacteria and/or Firmicutes to be among the most abundant phyla in fish gut communities. Studies on zebrafish indicated that Firmicutes and Proteobacteria are most common in larvae (12) but that the microbiota in adults was dominated by Fusobacteria (46). Studies on microbiota ontogenetic shifts in rainbow trout fry also indicated the presence of Proteobacteria or Firmicutes in association with either a marine-based or a freshwater plant-based diet offered during the first feeding (18). Meanwhile, the administration of prebiotics to Siberian sturgeon (Acipenser barii) shifted the composition of the gut microbiota primarily in the phylum Firmicutes (41).

Comparative analyses of the gut microbiota from 8 freshwater fish species encompassed...
ing fish with different feeding habits revealed that Proteobacteria and Firmicutes were the dominant phyla in all fish species (47). Another cross-sectional gut microbiota study performed on channel catfish also indicated the prevalence of Proteobacteria in fish at the earlier stage and the appearance of Firmicutes along with Proteobacteria when the fish reached 65 dph (48) or when the fish were stocked into a nursery pond (11).

We also found that the predominant taxonomic assemblages tended to consist of closely related taxa from the family Clostridiaceae. These included the genera Carnobacterium, Jeotgalibacillus, Sarcina, and Clostridium and an unclassified genus from the family Clostridiaceae. This might indicate that traits underlying assemblage membership were often shared among related host organisms. Taxa of a type of lactic acid bacterium, Carnobacterium, are common commensals of fish, in which they have been detected in the intestinal content and in the mucosal layer (49). One of these taxa, the genus Clostridium, was enriched following the transition to active feeding. In humans, this genus is part of the important commensal microbiota that begins to colonize the intestines of breast-fed infants as early as the first month of life and has been shown to play roles in modulating gut homeostasis over the entire life span (for a review, see reference 50). Studies in mice indicated that commensal clostridia populate specific regions in the intestinal mucosa, thus establishing a close relationship with gut cells that perform critical physiological functions (50).

Another taxon, Sarcina, was also abundant by the time that the fish reached 22 dph and were actively feeding. Members of the genus Sarcina are fermenting bacteria that are frequently found in the gastric contents and feces of human patients with gastrointestinal disorders. All the strains were obligate anaerobes, fermented cellulose, and required a carbohydrate for growth. One species, Sarcina maxima, produces ethanol as a major product but not butyric acid from glucose and can also produce acid and gas from sugar, like glucose, fructose, sucrose, maltose, lactose, galactose, and raffinose (51). S. ventriculi is widespread in the soil and may be considered part of the intestinal flora of human, although its significance remains unknown (51, 52). Given the biology of both microbial taxa and the benthic feeding habits of sturgeon, it is possible that Sarcina and Clostridium are predominant taxa in the distal gut of sturgeon in natural habitats. Sturgeon possess a valvular hindgut (spiral valve) that serves as the primary region of digestion and nutrient absorption and thus might provide an abundance of nutrients for bacteria like Sarcina and Clostridium to flourish (33, 53) to maintain gut physiological functions. Previous studies (53) showed that anaerobic bacterial fermentation takes place in the spiral valve, producing volatile fatty acids (VFAs) and hydrogen gas as by-products, supporting the idea that these bacteria may enhance digestive efficiency.

Two important Proteobacteria taxa present during all stages were Pseudomonas and Deelfgea. The presence of Pseudomonas at a relatively high abundance has also been reported as part of the lake sturgeon egg-associated community (54). Another study found Pseudomonas spp. to be present in abundance within the gastrointestinal tracts of juvenile coho salmon (Oncorhynchus kisutch) (55) as well as on eggs, but not in the water source or in food. This was likely attributed to the vertical transmission of a pioneering strain from the eggs to the fish GI tracts. Pseudomonas is also commonly

### TABLE 8 Taxa overrepresented, underrepresented, and neutrally represented in all three ontogenetic stages

<table>
<thead>
<tr>
<th>Representation</th>
<th>OTU (taxonomy)</th>
<th>Developmental stage in which OTU was detected</th>
<th>OTU detection frequencies in rearing water/fish gut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrally represented</td>
<td>OTU062 (genus Finegoldia)</td>
<td>All stages</td>
<td>0.091/0.087, 0.500/0.045, 0.333/0.250</td>
</tr>
<tr>
<td>Overrepresented</td>
<td>OTU016 (genus Pseudomonas)</td>
<td>All stages</td>
<td>0.455/0.957, 0.500/0.864, 0.667/0.375</td>
</tr>
<tr>
<td></td>
<td>OTU041 (genus Janibacter)</td>
<td>All stages</td>
<td>0.455/0.609, 0.500/0.318, 0.333/0.458</td>
</tr>
<tr>
<td>Underrepresented</td>
<td>OTU015 (genus Turicella)</td>
<td>All stages</td>
<td>0.818/0.609, 1.000/0.409, 0.6667/0.412</td>
</tr>
</tbody>
</table>
observed in the gut microbiota of mature fish (9, 56, 57). A number of diet-related studies have reported variability in the relative abundance of *Pseudomonas* spp. that were affected by differential food treatment (for a review, see reference 6).

Little information pertaining to *Deefgea* spp. is available. Studies have so far found only two species in the genus *Deefgea* (family *Neisseriaceae*, order *Betaproteobacteria*) (58). Those species were described as *Deefgea rivuli* and *Deefgea chitinilytica* (58, 59). However, these two taxa originated from hard water and wetland samples, respectively. The association of *Deefgea* isolates with fish was first documented in another study (60). That study reported six isolates of *D. chitinilytica* that were cultured from swabs of skin and internal organs of two freshwater ornamental fish species raised on farms, gold tench (*Tinca tinca*) and goldfish (*Carassius auratus auratus*). Several other bacterial taxa from the same family were described to be chitin-hydrolyzing species, and *D. chitinilytica* was suspected to have a similar function, too. Due to this, *Deefgea* may have significant importance and may be listed among opportunistic taxa that may play roles in infections of aquatic organisms (60).

Throughout the assembly of early gut bacterial communities, stochastic and deterministic factors associated with water and food epibiota play roles in shaping these ecological communities (9, 55, 61). As we have shown by ANOVA and multivariate analyses of the gut microbiota at each time point, the water rearing environment initially had a strong influence on the community composition during the prefeeding stage, as the gut communities reflected the aquatic communities of both water treatments. Water appears to serve as a primary inoculant before and during the transition from endogenous to exogenous feeding. The communities in the early stages of gut microbiota colonization were temporarily unstable in our study. Stochastic processes, such as the random recruitment of water epibiota into the gut, often occur at the beginning of exposure of the early upper GI tract and gill surface to ambient water (56). As the fish in our study developed further, the gut communities changed coincident with the initiation of active feeding and diverged in composition from the surrounding aquatic communities. Previous studies of the egg surface communities of lake sturgeon performed by our group also documented directional changes in bacterial community composition and diversity across sequential egg developmental stages (54). Another study on tilapia larvae (62) also documented changes in the gut community structure over time, with significant contributions of water bacterial communities.

As feeding continued, significant interactions were observed. We observed no significant differences in the gut communities between the supplemented and the nonsupplemented food treatment groups within the stream water environment. In contrast, significant differences in the gut microbiota of fish raised in groundwater were observed on the basis of the food treatment administered (a supplemented versus a nonsupplemented diet). The gut communities of fish raised in groundwater differed according to their diet, indicating that diet influences gut community membership. However, further analyses showed that the effect of diet on the gut community composition after the prefeeding stage was not evident. Only water treatment was significantly associated with the gut community composition at 2 weeks after active feeding. At this point, the gut communities differed between the stream water and the groundwater treatments.

We failed to detect significant differences in the gut microbial communities of fish raised using different feeding treatments. We were surprised to see that there was not much variation in the microbial taxonomic composition associated with water and food (*Artemia*), even though the water sources used for rearing the fish and preparing the live *Artemia* were different. A lack of community compositional heterogeneity may have been because the *Artemia* shrimp prepared in stream and groundwater had comparable nutrient properties. However, the ordination of detritus was segregated at the bottom right side of the PCoA plot, likely attributed to natural food items in the samples influencing their microbial community structure. We demonstrated the presence of natural food items in detritus/retentate based on eukaryotic profiling using 18S
rRNA sequence data. Several groups of lower eukaryotes were detected in the retentate collected from sock filters, including protozoans, diatoms, and dinoflagellates.

The dynamics of the community compositional change during gut assembly processes can be attributed to many processes (2, 5). First, the composition of the gut community could be determined by environmental selection. In this study, the environments were associated with the rearing condition of the fish, as fish raised within similar water environments were exposed to similar pools of microbial taxa present in each water source. Thus, the local community may also be under the influence of neutral processes. The taxa present in the gut may be a random draw of the species present in water.

With respect to the rearing temperature, fish raised in stream water were exposed to ambient temperatures of 12 to 18°C, which are within the typical range of temperatures in the Upper Black River, Cheboygan, MI, during spawning season. During the experiment, the temperatures of both the surface water and the groundwater were standardized as much as possible to be different by about 2°C. Although stream temperature fluctuations are considered ecologically relevant, we are aware that we cannot completely overrule the effect that a higher temperature may have on the physiology and metabolism of larval sturgeon (34, 63). This was especially true at the final stage of 2 weeks after active feeding initiation, as the stream water had a warmer temperature than the groundwater (Tables 1 and 2). Changes in fish physiology may affect gut anatomical development and contribute to a gut selective environment (28). Subsequently, maturation of the fish gut may impose selective pressures that favor particular subsets of taxa or that inhibit the growth of certain subsets of taxa.

Evaluations of the relative importance of neutral versus deterministic processes were achieved through the implementation of a neutral model. Our findings indicated that neutral processes (e.g., a random dispersal of taxa from water) were ongoing but were not the pervasive ecological force shaping the gut community during all three developmental stages. The low model predictability (low $R^2$ values) indicated that neutral processes are not dominant. The results suggest that deterministic processes are also responsible for changes in the gut community composition associated with development of the fish GI tract, as evidenced by reductions in the number of taxa between the gut and water (Fig. 6a to c). Meanwhile, the presence of overrepresented taxa, such as \textit{Pseudomonas} and \textit{Janibacter}, may play important roles for fish hosts. For example, \textit{Pseudomonas} species are important decomposers of organic matter in soil, water, and food products, but several species are also known to be pathogens in plants, animals, and human (64). \textit{Pseudomonas} is also commonly found as part of the fish intestinal and fish egg microbial communities (55). \textit{Janibacter}, on the other hand, is less common, yet a study has reported the presence of this genus in the midgut of mosquito, indicating its possible role as part of the gut microbiota (65).

We recommend that future studies include experimental designs that sample over a longer duration, including through manipulations of the water community. Further studies focusing on taxa that increase in relative abundance over time would be useful to establish the ecological functions that may be used at a greater efficiency in applications of microbe-based fish management in aquaculture. Such data could also be used to identify suitable sites for the release of fish from conservation hatchery programs.

Overall, the significant changes in the diversity and the taxonomic composition of the lake sturgeon gut microbiota that occurred were principally associated with early developmental stages in connection with the initiation of the first feeding. Following the initiation of exogenous feeding, the microbial communities diverged from those in the surrounding water community and the epibiota of the food provided. Our understanding and ability to control (i.e., through the application of probiotics) underlying deterministic and stochastic factors associated with the source of microbial inocula appeared to be tied in part to the communities in the water when feeding began. Advancing the use of microbial manipulation (pre- and probiotics) is a goal for the aquaculture industry to promote fish growth and health. Accordingly, advancing


understandings of the compositional dynamics that naturally occur in the gut microbiota of cultured fish species, like sturgeon, has relevance to commercial and conservation aquaculture. Future studies may profitably explore the effects of manipulations of communities in rearing water and food samples to understand the dynamics of microbial community assembly associated with these factors.

MATERIALS AND METHODS

Experimental design and feeding regime. Lake sturgeon larvae were produced from a single mated pair collected during the sturgeon spawning season on the Upper Black River, Cheboygan, MI, in May 2013. Full-sib individuals were used to reduce the potential variability in the microbiota associated with host genetic background. All individuals were raised in a flowthrough water system under four different rearing conditions. We used a 2-by-2 treatment factor design associated with water type and feeding regime. The water types included river water from the natal stream (S) and UV-treated groundwater (GW), reflecting water sources commonly used in traditional hatchery operations. These water types were used throughout the culture process, including food (brine shrimp or Artemia fish) preparation. Fish were fed either live Artemia nauplii, which are commonly used in sturgeon hatcheries (66), or Artemia nauplii supplemented with organic retentate, including detritus and aquatic zooplankton obtained from serial filtration through 100-μm- and 50-μm-pore-size filters used to filter river water entering the hatchery (referred to here as retentate). The presence of digestible taxa in the retentate was confirmed using massively parallel sequencing of the V9 region of the 18S rRNA gene.

Each treatment included six 3.0-liter polycarbonate tanks (Aquatic Habitat, Inc.) that served as biological replicates. Each tank held 70 individuals to achieve an estimated statistical power of 0.8. The power analyses were performed based on our preliminary microbial studies (54). The fish were exposed to the same water type beginning at the time of egg fertilization and during the incubation stages. A newly hatched larval group of 70 fish was then distributed into each treatment replicate at 10 days posthatch (dph). The food was offered at 12 dph. However, only at 16 dph, we began to consistently offer food when at least half of the fish were feeding. To ensure that the fish received consistent amounts of food throughout the experiment, previously established dry weight feeding rates for sturgeon (67) were utilized, whereby the larvae in all tanks were fed at 26% of body weight daily (BWD). Prior to the first feeding each day, the retentate was added to the freshly prepared Artemia fish for treatment groups Sp and GWp. The amount of retentate (in grams) collected daily varied depending on each day of collection, yet the food amount was adjusted accordingly to ensure that the fish in each treatment group consumed the same amount of food. Fish were fed to satiation three times daily. Mortality was recorded daily, and the body weight (in grams) was recorded once every third day. All experiments were conducted at the Sturgeon Streamside Rearing Facility, managed by the Michigan Department of Natural Resources (MDNR) and Michigan State University (MSU) at Onaway, MI, using an approved MSU Institutional Animal Care and Use Committee (IACUC) protocol (animal use and care procedure numbers 03/04-058-00 and 05/07-086-00).

Sample collection. We sampled 10 sturgeon larvae from each replicate for each treatment at three ontogenetic developmental stages: time 1, which was before active feeding (11 dph); time 2, which was after 1 week of fish active feeding (22 dph); and time 3, which was after 2 weeks of active feeding (29 dph). These time points were selected to capture the critical phases of GI tract development after the fish completely absorbed their yolk sac and the gastrointestinal tract anatomy was completed (which occurs at approximately 10 dph). Lake sturgeon started actively feeding at between 13 and 16 dph when food was first offered (33, 43). The fish were sampled for microbial community interrogation at each time point and were euthanized with an overdose of ethyl 3-aminobenzoate methanesulfonate (MS-222; Sigma-Aldrich, St. Louis, MO, USA) at the time of sampling. Each individual larva was photographed at the time of euthanization, and images were analyzed, using ImageJ software, to determine the total body length (in millimeters). The fish were then transferred to and preserved in 50-ml Corning centrifuge tubes containing 80% filter-sterilized ethanol until GI tract dissection and bacterial DNA extraction were performed.

Other environmental samples, including water and food, were also collected during each sampling period. Water samples (250 ml) from stream water and groundwater were collected from the hatchery reservoir tank and filtered through 0.45-μm-pore-size filters with a 47-mm-diameter filter membrane (Fisher Scientific), using a hand pump to obtain the aquatic microbial communities on the filter paper. These filters were then transferred to and preserved in 50-ml Corning centrifuge tubes containing 80% ethanol until bacterial DNA extraction was performed. For food samples, approximately 200 μl of food was pipetted into a 2-ml Eppendorf tube and was preserved in the same manner as the fish and water samples. All environmental samples from the different treatment groups are denoted by Artemia-GW (food treatment in which Artemia was prepared using groundwater), Artemia-S (food treatment in which live Artemia was prepared using stream water), detritus (food treatment in which retentates were collected from a sock filter prior to mixing with live Artemia), Water-GW (UV-treated groundwater), and Water-S (treatment with stream water).

DNA extraction and 16S rRNA amplicon sequencing. The gut microbiota from lake sturgeon larvae was surveyed using high-throughput sequencing of the V4 region of the 16S rRNA gene. The distal gut (spiral valve) from lake sturgeon larvae was sampled for each larva following aseptic techniques. The distal gut was defined as the section extending from the beginning of the intestine to the spiral valve. Exterior surfaces were swabbed with 100% ethanol before dissections of the whole digestive tract using sterile instruments. Dissections were performed as previously described (68) with slight modification. The intact GI
tracts were cut from the fish body cavities, and the excised GI tract was immediately transferred into filter-sterilized 80% ethanol for DNA isolation. Due to the small size of the gut, a composite of at least four GI tract samples from larvae were combined for each tank replicate within each treatment group at each time point. Each tube containing GI tract samples was first centrifuged for 15 min at 4°C to pellet the tissues and bacteria, before the DNA was extracted. A MoBio PowerSoil DNA isolation kit (Carlsbad, CA, USA), including a bead-beating step, was used for following the protocols for low-biomass samples, as suggested by the manufacturer, with slight modification. The integrity of each DNA sample was assessed based on the amplification of approximately 1.4 kbp of the 16S rRNA gene (with forward primer 27F and reverse primer 1389R), followed by gel agarose electrophoresis, and the DNA concentration was determined using a microplate spectrophotometer (BioTek, Winooski, VT, USA).

A total of 20 filter membranes (consisting of two biological replicates from each of the water treatments, stream and UV-filtered water) containing trapped aquatic microbial communities and 20 food samples (two biological replicates of fresh brine shrimp prepared using the respective water sources from both water treatments) were randomly collected five times throughout the period of fish rearing. The membranes and food samples were first vortexed vigorously for 5 min to wash the bacteria from the membrane or food. The homogenates were then centrifuged (10,000 rpm, 4°C, 15 min) to pellet the tissues and bacteria before the DNA was extracted. The pellets were suspended with 500 μL of buffer solution from the power bead tube of the MoBio PowerSoil DNA isolation kit (Carlsbad, CA, USA) and then transferred back into the power bead tube. The isolation of DNA from the water and food bacterial communities proceeded using protocols for low-biomass samples, as suggested by the manufacturer.

One hundred nineteen DNA samples, consisting of 79 samples from fish gut, 36 samples from water and food (here referred as environmental microbial communities), and 4 positive-control samples (microbial DNA present in samples from activated sludge and a pure culture of *Escherichia coli*), were then qualitatively sufficient bacterial DNA, as indicated by the presence of PCR amplicon bands following gel electrophoresis, and later were submitted for sequencing at the Michigan State University Research Technology Support Facility (RTSF; East Lansing, MI, USA). All of the sequencing procedures, including the construction of the Illumina sequencing library, emulsion PCR, and MiSeq paired-end sequencing (with the V2 platform) of the V4 region (~250 bp; with primers 515F [GTGCCCA GCMGCGCGGTATA] and 806R [TAATCTTGGGATCAGATGCAGG]) (69), followed standard Illumina (San Diego, CA, USA) protocols. For reference, this primer pair amplifies the region from positions 533 to 786 in the *E. coli* strain 83972 sequence (70). Michigan State’s Genomics RTSF (https://rtsf.natsci .msu.edu/genomics/) provided standard Illumina quality control, including base calling by Illumina real-time analysis (v1.18.61), demultiplexing, adaptor and barcode removal, and real-time analysis (RTA) conversion to the Illumina Bcl2Fastq (v1.8.4) program.

**Sequence processing.** Sequence data were processed using the default sequencing data analysis pipeline and computing workflow. Briefly, paired-end sequence merging, quality filtering, denoising, singleton sequence removal, chimera checking, and taxonomic assignment of the operational taxonomic unit (OTU) selection were conducted using a protocol within an open-source workflow implemented by the mothur (v1.36.1) program (69).

Alignments were performed using the SILVA-based bacterial reference database. *De novo* OTU clustering was performed using hierarchical algorithms as the default option offered in mothur to cluster sequences defined with 97% identity (71), whereas the taxonomic assignment was achieved using the training file set provided by mothur, which was derived from the Ribosomal Database Project (RDP).

To minimize the effects of undersampling while maintaining as large a data set as possible, the final OTU table was rarefied to a depth of 5,775 sequences per sample. Five DNA samples with low sequence depths were discarded in downstream analyses. Rarefaction analyses were performed to evaluate the sampling coverage for each sample based on the selected sequence depth.

**Retentate samples and eukaryotic 18S rRNA V9 metabarcoding for diet characterization.** To assess the biodiversity inventories and edible items in the detritus recovered from the sock filter, we performed massively parallel amplicon sequencing of the V9 region using a total of four retentate DNA samples (including both biological and technical replicates). Samples were submitted for sequencing to the Michigan State University Research Technology Support Facility (RTSF; East Lansing, MI, USA). All sequencing procedures, including the construction of the Illumina sequencing library, emulsion PCR, and MiSeq paired-end sequencing (with the V2 platform) of the V9 region (~175 bp; with primers Euk 1391F [GTACACACCGCCCGTC] and Euk Br [TGATCCTTCTGCAGGTTCACCTAC]) (72), followed standard Illumina (San Diego, CA, USA) protocols. Michigan State’s Genomics RTSF (https://rtsf.natsci .msu.edu/genomics/) provided standard Illumina quality control, as mentioned above.

Sequence reads were analyzed based on the protocol provided elsewhere (69), as mentioned above. Lower eukaryotic profiling was achieved with slight modification during sequence alignment and taxonomic identification. mothur program-provided SILVA reference files were customized based on *Saccharomyces cerevisiae* V9 region sequences and later properly formatted for taxonomic classification to the lowest level possible for eukaryotic analysis in mothur. Due to the short length of the amplified region, the sequence data allowed identification only to higher taxonomic levels, such as the genus, family, and order levels.

**Alpha diversity and temporal and differential abundance of OTUs.** All measures of community diversity and similarity, including the inverse Simpson diversity index (1/D) and the OTU richness of each sample, were calculated from the sequence data within mothur to quantify alpha (α) diversity. To test for significant differences in diversity indices among treatment groups (water and diet) and among time periods, a multiple-factor ANOVA was performed on the summary files provided by mothur using the programming and statistical software R (v3.0.2) base package. The test was followed by Tukey’s honest
significant difference (HSD) post hoc tests. P values below 0.05 indicated significant differences in pairwise mean comparisons.

A custom R code was used to calculate the relative abundance and identify the dominant phyla and taxa (OTUs) in all communities across sampling times (gut microbiota, water, and Artemia fish-associated epibiota). Codes were written and implemented using the packages dplyr and reshape2. The relative abundance of all taxa within community samples was calculated, and taxa with a normalized abundance that exceeded 0.1% were considered predominant taxa. The 20 most abundant taxa were a subset of the total number of OTUs based on the 0.1% cutoff, and relative abundances were prepared in a tabular format to show the temporal variability in fish gut community composition. The remaining taxa were grouped as “others.”

Linear discriminant analysis (LDA) effect size (LEfSe) scores were determined using freely available online software (https://huttenhower.sph.harvard.edu/galaxy/) (40) to determine microbial communities that exceeded 0.1% were considered predominant taxa. The 20 most abundant taxa were a subset of the total number of OTUs based on the 0.1% cutoff, and relative abundances were prepared in a tabular format to show the temporal variability in fish gut community composition. The remaining taxa were grouped as “others.”

Linear discriminant analysis (LDA) effect size (LEfSe) scores were determined using freely available online software (https://huttenhower.sph.harvard.edu/galaxy/) (40) to determine microbial communities that exceeded 0.1% were considered predominant taxa. The 20 most abundant taxa were a subset of the total number of OTUs based on the 0.1% cutoff, and relative abundances were prepared in a tabular format to show the temporal variability in fish gut community composition. The remaining taxa were grouped as “others.”

Linear discriminant analysis (LDA) effect size (LEfSe) scores were determined using freely available online software (https://huttenhower.sph.harvard.edu/galaxy/) (40) to determine microbial communities that exceeded 0.1% were considered predominant taxa. The 20 most abundant taxa were a subset of the total number of OTUs based on the 0.1% cutoff, and relative abundances were prepared in a tabular format to show the temporal variability in fish gut community composition. The remaining taxa were grouped as “others.”

Linear discriminant analysis (LDA) effect size (LEfSe) scores were determined using freely available online software (https://huttenhower.sph.harvard.edu/galaxy/) (40) to determine microbial communities that exceeded 0.1% were considered predominant taxa. The 20 most abundant taxa were a subset of the total number of OTUs based on the 0.1% cutoff, and relative abundances were prepared in a tabular format to show the temporal variability in fish gut community composition. The remaining taxa were grouped as “others.”

Beta diversity. We used several packages implemented in the program R to perform comparative community compositional analysis of beta (β) diversity and other community ecological statistics using the tabulated OTU data set of the predominant taxa. Briefly, we used the vegan function to generate estimates of β-diversity (BC), distance among sample microbial communities (73). Subsequently, we used the cmdscale function to perform ordination (principal-coordinate analyses [PCoA]) based on BC distance (74). The ggplot2 and ggplot22 packages (75) were used to create ordination plots to visually compare the gut and environmental bacterial community compositions among samples collected from different treatments and among sampling periods based on the two largest eigenvalues. Two multivariate hypothesis tests were implemented using two functions. The adonis function was used to perform multivariate hypothesis testing on differences between the locations of the centroids of treatment group coordinate ordinations based on permutational multivariate analyses of variance (PERMANOVA). The betadisper function was used to perform a test for the homogeneity of multivariate dispersion (PERM-DISP) test on community BC matrices (76, 77). These tests were employed because of the nonparametric and skewed nature of the microbial community compositional data. The OTUs that had the highest correlations with the PCoA x and y component axes were identified based on Pearson correlation coefficients using the corr function.

Influence of water and diet treatments of GI tract microbiota. To analyze the treatment effects of water type and diets on the fish gut microbiota, PCoAs were performed separately on the fish gut communities for all four treatment groups within each developmental stage. Important PCoA axes denoted by eigenvalues comparatively larger than the average eigenvalues were selected. Linear regression models were fit, where each axis was a response variable given the predictor variables of water type and food type. Under the null hypothesis, we expected the gut community composition at each time point to be unaffected by treatment. The axes represent the linear correlations of the bacterial taxonomic compositions present in gut communities. The axes that showed significant effects of treatments or interactions between water and diet treatments were analyzed using the lsmeans function to determine the effect of each factor (or combinations of factors in the interaction) on the bacterial taxonomic composition and relative abundance in larval fish GI tracts.

Tests evaluating whether GI tract communities were a neutral subset of the environmental (source) communities. We used a neutral community ecological model adapted from previous work (24) to explore the relative importance of neutral processes (i.e., dispersal and ecological drift) and selection in the gut microbiota at a given time of sampling (prefeeding and 1 week and 2 weeks after active feeding began). This model also distinguished members of the gut microbiota whose presence was consistent with dispersal from surrounding environmental communities (water as a source) and those that deviated from the neutral model (i.e., those that were over- or underrepresented in the gut relative to the water source(s)). In general, the model predicts that taxa that are abundant in the source will be widespread taxa which occur in a set of target communities (larval fish gut), since these taxa are more likely to disperse by chance and be randomly sampled by an individual host, while rare taxa are more likely to be lost from fish hosts due to ecological drift (23, 24).

To quantify the relative importance of stochastic processes (neutral drift and gene flow) versus deterministic processes (selection) on the lake sturgeon gut community composition during the three ontogenetic stages, we adapted the approach of Venkataraman and colleagues (25) using their custom R scripts, with slight modification, referring to our data set. Water sources that were interrogated at the time that each group of fish was sampled were considered sources for the fish gut microbiome. The water source microbial communities were created by pooling surveys of both stream water and groundwater taxa at each respective sampling time, whereas the target communities were those microbial taxa that were collectively detected in the gut of fish larvae across all treatments.

Under the neutral model, the probability of observing a microbial OTU in the gut (target) was based on the relative frequency of occurrence in the source (water) at the time of sampling. The relative abundance of a given OTU in the source community was calculated as the number of sequences with the OTU in the source community/total number of sequences in the source community. Similarly, the
empirically observed frequency of detection for each OTU in the gut was calculated as the number of gut samples in which the OTU was detected/total number of gut samples across all treatments. Next, the expected frequency of detection was calculated based on a beta probability distribution for each OTU shared between the gut and the source community. Briefly, the lower limit of this probability density function was the relative abundance of the least-abundant OTU in the source community, while the upper limit was 1. The neutral model of Sloan et al. (24) was fit to an overall fitting parameter (Ntm) and the relative abundance of the OTU in the source community. Ntm is the minimized sum of squares of residuals between observed detection frequency and predicted detection frequency for total number of OTUs (25). The fitting parameter reflects the dispersal of microbes from the source community to the fish gut. The fitting of this parameter was performed using a least-squares approach (25). Finally, the variability around this expected detection frequency was calculated using 95% binomial proportion confidence intervals (the Wilson score interval method) with the Hmisc package in R (80). Calculating this for all OTUs yields the best-fit neutral model curve and the 95% confidence intervals shown in Fig. 6. The goodness of fit of this curve was assessed using the coefficient of determination \((R^2)\) (77, 78). The level of fit of the neutral model served as an adequate representation of the gut community composition if the \(R^2\) value was high, whereas the significance of neutral processes was indicated by the \(P\) value.

**SUPPLEMENTAL MATERIAL**

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 0.1 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.1 MB.

**SUPPLEMENTAL FILE 3**, XLSX file, 0.1 MB.

**SUPPLEMENTAL FILE 4**, XLSX file, 0.1 MB.

**ACKNOWLEDGMENTS**

Funding for this study was provided by the Michigan Department of Natural Resources, the Great Lakes Fishery Trust, and Michigan State University Ag-Bio Research. S.A.R. was supported by the Universiti Kebangsaan Malaysia (UKM) and the Ministry of Higher Education (MOHE), Government of Malaysia, and a dissertation completion fellowship from the College of Agriculture and Natural Resources while attending MSU.

We thank Brian Maurer, Andrew Denhaart, and Nick Sard for helping with data analyses and R. We also thank John Bauman, Kari Dammerman, Nathan Barton, Sarah Watson, Troy Smith, Adam Unstead, Lindsay Adams, and James Garavaglia for assistance in construction of the experiment and collection of the data. Mohamed Faisal and Terence Marsh provided valuable advice on the experimental design, analyses, and interpretation of the data.

**REFERENCES**


