List of Supplementary information

Meter-scale variation and taxon-accumulation curves in riverine microbiomes within a single transect. Bingdi Liu, Lin Zhang, Jason Knouft, Fangqiong Ling, Submitted for review at FESE.

Supplementary Figures

Figure S1. Rarefaction curves derived from water microbiome samples taken along a one-meter spaced transect in the Meramec river. (A) Site 1-5. (B) Site 6-10. (C) Site 11-15. (D) Site 16-20. All samples have reached plateaus in sequencing efforts. For clarity, panels are organized with five samples in each panel.

Figure S2. Near-shore and near-center groups emerged in principle coordinate analysis (PCoA) of Jaccard distances. The PCoA resulting from Jaccard distances presented a dramatic difference in composition between near-shore and near-center groups. Circles indicate samples. Numbers indicate the distance between the site and the river shore. The orange polygon indicates the near-shore group, and the blue polygon indicates the near-center group.

Supplementary Tables

Table S1. Pearson Correlation between technical replicates of the same riverine microbiome samples.
Table S2. Top twenty most abundant bacterial families in the Meramec transect.
Table S3. ASVs identified as informative taxa by LefseR and associated log LDA scores.
Table S4. Model fits of statistical and ecological null models.

Supplementary Text: Applying ecological and statistical null models to riverine microbial communities .

Figure S1

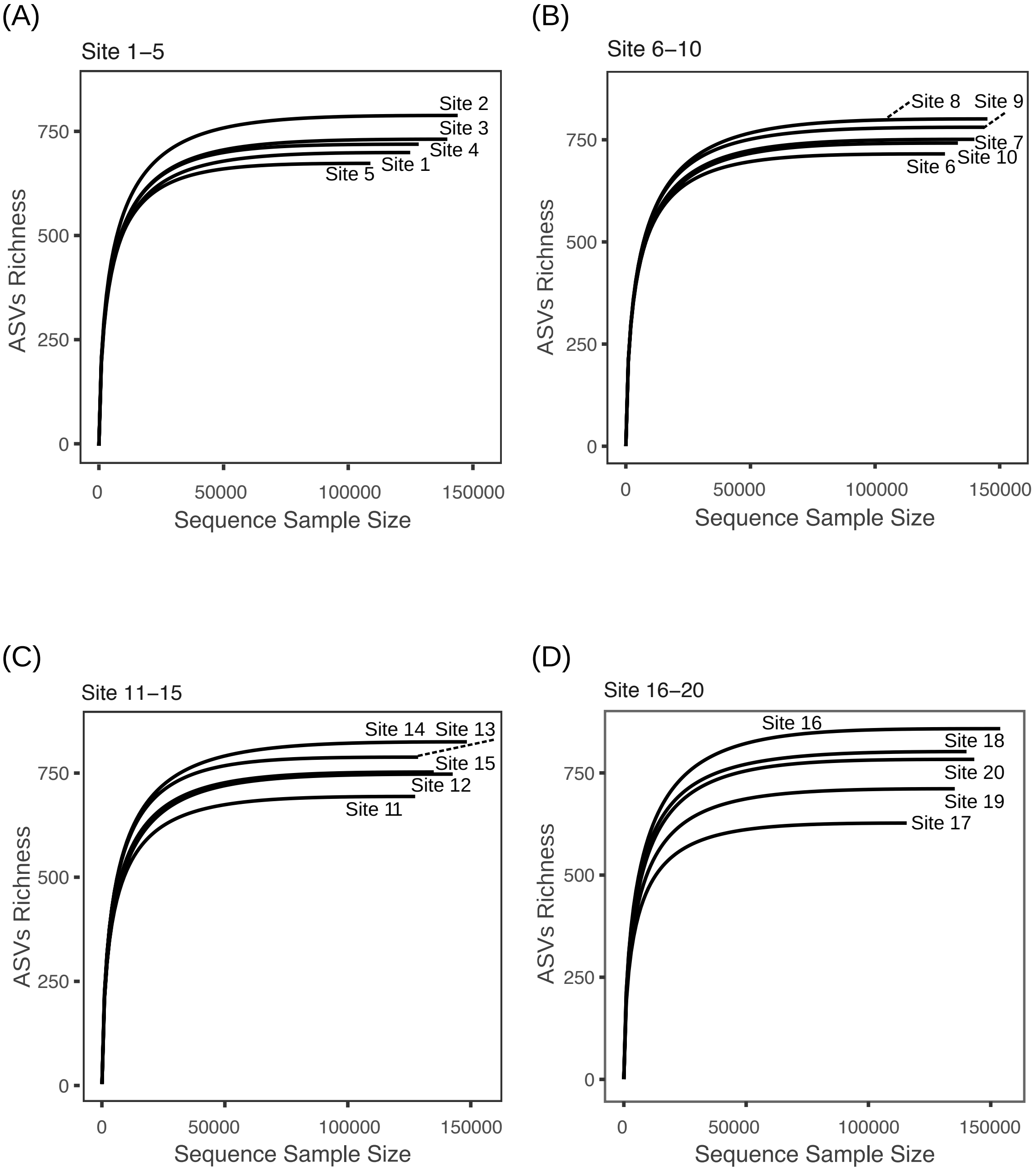
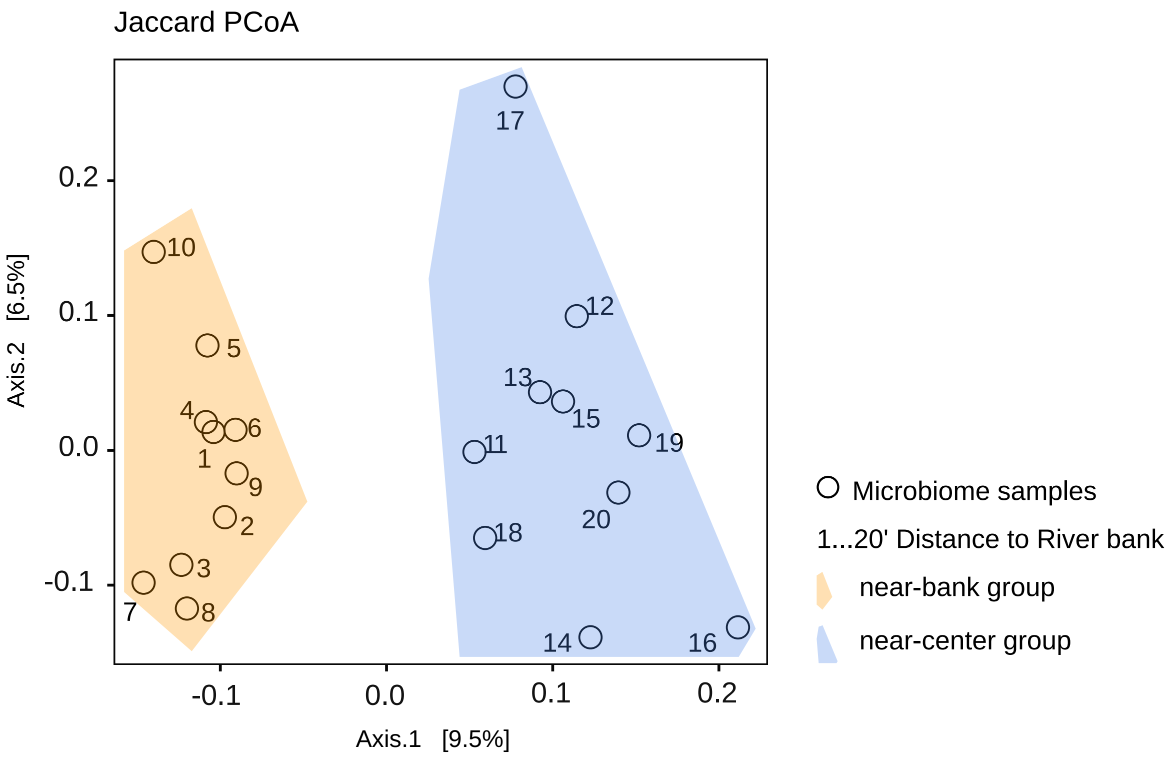

(A)
Site 1–5
Site 2
Site 3
Site 4
Site 1
Site 5
ASVs Richness
0
250
500
750
0
50000
100000
150000
Sequence Sample Size
(B)
Site 6–10
Site 8
Site 9
Site 7
Site 10
Site 6
ASVs Richness
0
250
500
750
0
50000
100000
150000
Sequence Sample Size
(C)
Site 11–15
Site 14
Site 13
Site 15
Site 12
Site 11
ASVs Richness
0
250
500
750
0
50000
100000
150000
Sequence Sample Size
(D)
Site 16–20
Site 16
Site 18
Site 20
Site 19
Site 17
ASVs Richness
0
250
500
750
0
50000
100000
150000
Sequence Sample Size

Figure S2



Jaccard PCoA
Axis.2 [6.5%]
0.2
0.1
0.0
-0.1
Axis.1 [9.5%]
-0.1
0.0
0.1
0.2
10
5
4
6
1
9
2
3
7
8
17
12
13
15
11
19
20
18
14
16
Microbiome samples
1...20' Distance to River bank
near-bank group
near-center group

Table S1. Pearson Correlation between technical replicates of the same riverine microbiome samples.

Sample ID	Pearson correlation
Site_1	0.997
Site_2	0.999
Site_3	0.998
Site_4	0.998
Site_5	0.998
Site_6	0.998
Site_7	0.996
Site_8	0.999
Site_9	0.998
Site_10	0.999
Site_11	0.999
Site_12	0.999
Site_13	0.998
Site_14	0.998
Site_15	0.998
Site_16	0.999
Site_17	0.998
Site_18	0.998
Site_19	0.999
Site_20	0.998

Table S2. Top twenty most abundant bacterial families in the Meramec transect.

Phylum	Class	Order	Family	Relative abundance
Bacteroidota	*Bacteroidia*	*Flavobacteriales*	*Flavobacteriaceae*	29.06%
Proteobacteria	*Gammaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	21.91%
Actinobacteriota	*Actinobacteria*	*Frankiales*	*Sporichthyaceae*	7.42%
Bacteroidota	*Bacteroidia*	*Cytophagales*	*Spirosomaceae*	6.56%
Proteobacteria	*Gammaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	5.53%
Bacteroidota	*Bacteroidia*	*Flavobacteriales*	*Crocinitomicaceae*	3.19%
Cyanobacteria	*Cyanobacteriia*	*Chloroplast*	*Chloroplast*	3.03%
Actinobacteriota	*Actinobacteria*	*Micrococcales*	*Microbacteriaceae*	2.81%
Proteobacteria	*Gammaproteobacteria*	*Burkholderiales*	*Methylophilaceae*	2.36%
Proteobacteria	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	1.90%
Bacteroidota	*Bacteroidia*	*Chitinophagales*	*Chitinophagaceae*	1.83%
Proteobacteria	*Gammaproteobacteria*	*Burkholderiales*	*Rhodocyclaceae*	1.02%
Proteobacteria	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	0.96%
Bacteroidota	*Bacteroidia*	*Sphingobacteriales*	*Sphingobacteriaceae*	0.87%
Proteobacteria	*Gammaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	0.77%
Proteobacteria	*Gammaproteobacteria*	*Burkholderiales*	*Alcaligenaceae*	0.69%
Campilobacterota	*Campylobacteria*	*Campylobacterales*	*Arcobacteraceae*	0.44%
Actinobacteriota	*Acidimicrobiia*	*Microtrichales*	*Ilumatobacteraceae*	0.42%
Bacteroidota	*Bacteroidia*	*Sphingobacteriales*	NS11-12_marine_group	0.41%
Actinobacteriota	*Actinobacteria*	*Corynebacteriales*	*Mycobacteriaceae*	0.38%

Table S3. ASVs identified as informative taxa by LefseR and their associated log LDA scores

ASV	Family	Genus	LDA score (log 10)	Location group
97a35de693545ff794c1085bfa61a7a6	*Flavobacteriaceae*	*Flavobacterium*	-3.03	Near-shore
e53656fb00f7a9dcdb154b610ae779a4	*Spirosomaceae*	*Pseudarcicella*	-2.91	Near-shore
0f05f60bd4a4852ad33a80a3ab58f2e7	*Flavobacteriaceae*	*Flavobacterium*	-2.48	Near-shore
c09dc1ca17575194c0c4dea52cc7d8d5	*Crocinitomicaceae*	*Fluviicola*	-2.30	Near-shore
6c98232456f30a254b5c0d2a9d2b70e4	*Flavobacteriaceae*	*Flavobacterium*	-2.19	Near-shore
f2a9e81130dda9802ab8e7760a77f296	*Methylophilaceae*	*Methylotenera*	-2.17	Near-shore
cdc283691ec0ebd20791bdd254f6b959	*Flavobacteriaceae*	*Flavobacterium*	-2.17	Near-shore
005dc613581d3ccc2c2654bb39df847f	*Sporichthyaceae*	*Candidatus Planktophila*	-2.09	Near-shore
ace463b19f889926269b2722780bf1f5	*Crocinitomicaceae*	*Fluviicola*	-2.07	Near-shore
de3d1eb30333d43a7b69ef4fec973951	*Flavobacteriaceae*	*Flavobacterium*	-2.06	Near-shore
ee8cc4df843289f855eb1ddc51347d68	*Flavobacteriaceae*	*Flavobacterium*	-2.05	Near-shore
b3e4fe29c660e2525d1af58ccda40a88	*Flavobacteriaceae*	*Flavobacterium*	-2.05	Near-shore
db46080bf97fc5030b245298efe8bee7	*Microbacteriaceae*	*Rhodoluna*	2.01	Near-center
3bebd715e82d11b99d6a925fb7fcebd5	*Sphingomonadaceae*	*Sphingorhabdus*	2.03	Near-center
bac95eeb2a8263d52d81533d23518c14	*Rhodobacteraceae*	*not classified*	2.07	Near-center
f6f09be65bb70a3eb12c546cff667f7d	*Microbacteriaceae*	*Aurantimicrobium*	2.14	Near-center
216fefb4b0cbed1c5463db2ec24c5356	*Comamonadaceae*	*not classified*	2.34	Near-center
d0c1d5091d1495638bc415196e56ca6e	*Comamonadaceae*	*Limnohabitans*	2.51	Near-center
e545faf63c407fc1e58c9509bc95066f	*Comamonadaceae*	*not classified*	3.07	Near-center
65fdf0c0aa400eac203709fe638a5da2	*Burkholderiaceae*	*Polynucleobacter*	3.14	Near-center

Table S4. Model fits of statistical and ecological null models.

Models	R-squared	Root Mean Squared Error
Binomial model	-0.48	0.38
Poisson model	-0.48	0.38
Sloan's UNTB model	0.89	0.1

Supplementary Text: Applying ecological and statistical null models to riverine microbial communities

Methods

To investigate the processes of dispersal and ecological drift, we employed the Sloan version of neutral model as a null model(Sloan et al., 2006; Hubbell, 2011), which had been applied in other environmental and host-associated microbiome (Venkataraman et al., 2015; Burns et al., 2016; Ling et al., 2018). Briefly, the frequency of occurrence of each ASV was modeled from the average relative abundance of that ASV across the transect p_i, the neutral model parameter m, migration rates, and the total number of reads per sample N with the detection limit d =1/N using the probability density function of the beta distribution ϕ.

$$\Pr(ASV_i \text{ is present with a relative abundance } > d) = \int_d^1 \emptyset_i(x; N * m, p_i)dx. \tag{1}$$

A non-linear least square method implemented in minpack.lm was used (Elzhoov et al., 2016). The predicted prevalence freq.pred and confidence intervals were obtained by the function pbeta() and binconf() with alpha=0.05, method="wilson" and return.df=TRUE using package Hmisc (Harrell and Duponot, 2021). Then, the plot of the fitted model with the actual data points was generated. In addition to the Sloan's version of the neutral model, we employed the binomial distribution and Poisson distribution as statistical null models, and compared the performance of the Sloan's neutral model against these two models.

Last, the performance of three models were evaluated by computing their R-squared and Root Mean Squared Error where freq is the average of actual prevalence across all taxa:

$$freq(ASV_i) = \frac{Actual\ occurrence\ of\ ASV_i}{Number\ of\ samples} \tag{2}$$

$$R\ squared = 1 - \frac{\sum(freq_i - freq_i.pred)^2}{\sum(freq_i - \overline{freq})^2} \tag{3}$$

$$Root\ Mean\ Squared\ Error = \sqrt{\frac{\sum(freq_i - freq_i.pred)^2}{Number\ of\ ASV - 1}} \tag{4}$$

Results

Neutral model and random sampling models

The R-squared of neutral model from beta distribution was 0.89 and its Root Mean Squared Error was 0.1. Besides the neutral model, binomial and Poisson models were fitted, yet both resulted in R-squared smaller than zero, indicating poor fits (Table S4). The larger R-squared and smaller Root Mean Squared Error of the neutral model indicated it performed better than the other two statistical null models. Thus, the effects of dispersal and ecological drift were likely to be beyond that of random sampling of the metacommunity.

Reference

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., et al. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J* 10, 655–664. doi:10.1038/ismej.2015.142.

Elzhoov, T. V., Mullen, K. M., and Boker, B. (2016). minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK. Available at: https://cran.r-project.org/web/packages/minpack.lm/index.htmlv.

Harrell, F. E. Jr., and Duponot, C. (2021). *Hmisc: Harrell Miscellaneous*. Available at: https://cran.r-project.org/web/packages/Hmisc/index.html.

Hubbell, S. P. (2011). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton University Press doi:10.1515/9781400837526.

Ling, F., Whitaker, R., LeChevallier, M. W., and Liu, W.-T. (2018). Drinking water microbiome assembly induced by water stagnation. *ISME J* 12, 1520–1531. doi:10.1038/s41396-018-0101-5.

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* 8, 732–740. doi:10.1111/j.1462-2920.2005.00956.x.

Venkataraman, A., Bassis, C. M., Beck, J. M., Young, V. B., Curtis, J. L., Huffnagle, G. B., et al. (2015). Application of a Neutral Community Model To Assess Structuring of the Human Lung Microbiome. *mBio* 6, e02284-14. doi:10.1128/mBio.02284-14.