

Supplemental Methods

Network Analysis

To allow cross comparability between depths, and in consideration that not all depths were collected on every sampling day, we kept only dates where samples had been collected at most depths. In cases where a few depths were missing, we selected sampling dates such that each depth was sampled a similar number of times. We arrived at following number of samples for each depth: 5m, 64; DCM 62; 150m, 62; 500m, 61; 890m, 62 samples.

The method interpolated missing values by filling in with zeroes after normalization (analogous to imputing with median values); this imputation is needed to calculate time lagged associations and causes the Spearman values to be more conservative than they would be if we imputed with other numbers. LSA output data were preliminarily filtered by identifying and retaining only those associations that had P values of less than 1%, (which corresponded to Q values of less than 1%, due to high proportion of statistically significant comparisons) and Spearman R values greater than 0.5 or less than -0.5. In order to focus our analysis on the most commonly occurring bacteria, we eliminated bacteria that could not be detected at least 25 times (corresponding to ~20% of the sampling dates) at a given depth. This step left 37% of the OTUs we had started with (averaged over all samples and depths), but maintained 91% of the total diversity, since the more common OTUs were also the most abundant ones. Note that this step means that different OTUs are removed at different depths depending on where those OTUs occur most commonly.

Organization of Networks

Our network from all of the data described above will be referred to as the “*all-vs-all*” network because it contains comparisons of all nodes at all depths versus each other depth. In order to focus on patterns in the bacterial community, we generated a restricted network containing only bacterial nodes but not environmental parameters from the “*all-vs-all*” network; this network will be referred to henceforth as the “*bacterial all-vs-all*” network. From the “*bacterial all-vs-all* network” we created a subnetwork in which we kept only edges representing positive correlations and nodes connected by those edges, which we call the “*positive bacterial all-vs-all*” network. These three networks, “*all-vs-all*”, “*bacterial all-vs-all*” and “*positive bacterial all-vs-all*” will henceforth be referred to as “parent” networks because other “child” networks were derived from them.

Five “*within depth child networks*” were generated from each of the three parent networks (“*all-vs-all*”, “*bacterial all-vs-all*” and “*positive bacterial all-vs-all*”), each of which only contained nodes from one

depth. Note that “child networks” are derived from the “parent networks” but only some nodes and edges are kept. From both the “bacterial all-vs-all” network and from the “positive bacterial all-vs-all” network, we generated 10 “*between depth child networks*”, containing bacterial nodes from all possible pairs of depths.

Density

Network density is defined in this manuscript as

$$E/E_p$$

Where E is the number of edges in a network and E_p is the number of edges that would be present if every node were connected to every other node. Thus, for one network

$$E_p = N * N - 1 / 2$$

where N is the number of nodes in the network. If we are evaluating the density of edges between two networks

$$E_p = N_1 * N_2$$

where N_1 is the number of edges in the first network and N_2 is the number of edges in the second network. High network density may indicate either that many groups of bacteria respond in common ways to a few environmental variables or else that many OTUs are somehow associated with other OTUs such as by symbiosis.

Application of Density Measurement

For the *bacterial all-vs-all* network and *positive bacterial all-vs-all network*, we determined three edge density attributes (1) the total density, (2) the density of edges that connected pairs of nodes from the same depth, and (3) the density of edges that connected pairs of nodes from different depths. For each of the *within depth bacterial networks* and *within depth positive bacterial networks* we determined (1) the density of edges connecting nodes within that depth and (2) the density of edges connecting that depth to nodes at other depths. Then for each *between depth network* we determined (1) the density of edges connecting pairs of nodes of at the same depth and (2) the density of edges connecting pairs of nodes from each of the two depths.

Calculation of Clustering and Path Length

Networks were transferred from Cytoscape to R using the “RCytoscape” package (Shannon et al. 2013). The “transitivity” and “average.path.length” functions from the “igraph” package (Csardi & Nepusz 2006) were applied to each “within depth network”, and the “bacterial all vs all” networks to determine “Clustering” and “Path Length” coefficients. For each “observed network” we generated 1000 random networks (for statistical comparisons) of the same number of nodes and edges as each of the original networks using the “erdos.renyi.game” function from the “igraph” package. We determined P values for the clustering coefficient of each observed network by determining the fraction of the 1000 derived random networks that had greater clustering coefficients than the observed clustering coefficient. We determined P values of the path length by determining the fraction of the 1000 derived random networks that had greater average path length than the observed clustering coefficients, the fraction that had lower average path length, taking the smaller of these fractions and multiplying by two for a two tailed test.