



Figure S4 Intra-cluster link enrichment as a function of cluster number in M-Y and M-3D. We ran the hierarchical agglomerative clustering algorithm on the M-Y and M-3D datasets. In this algorithm, the number of clusters gradually decreases as clusters are merged together; to generate this data, we continued clustering all the way down to $N = 1$. Shown is the metric E , or intra-cluster link enrichment, at each value of N . Note that for both M-Y and M-3D the maximum value of $E(N)$ occurs when N is equal to the true number of distinct species present in the draft assembly.