

Figure S1: Life cycle of *Trichosphaerium* sp. observation through measurement of amoebae sizes (maximum-diamond, average-solid circle and minimum-square) recorded per field of vision (20x magnification) for each day during the approximately 25-day experiment.

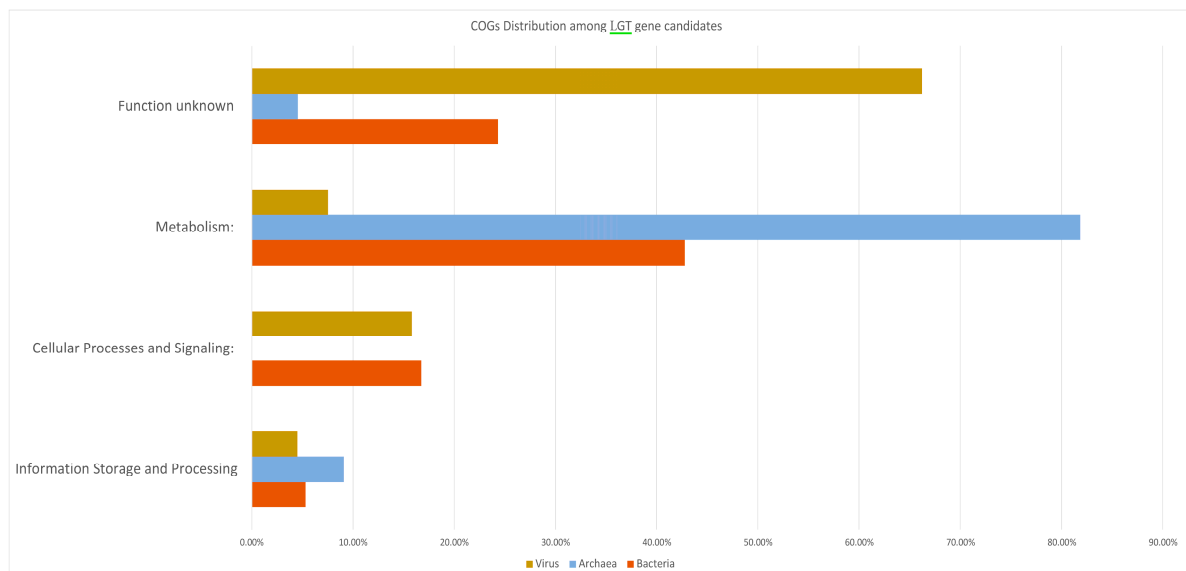


Figure S2: Functional categories based on Cluster Orthologous Groups (COGs) database of putative LGTs in *Trichosphaerium* sp. for bacteria, viruses and archaea.

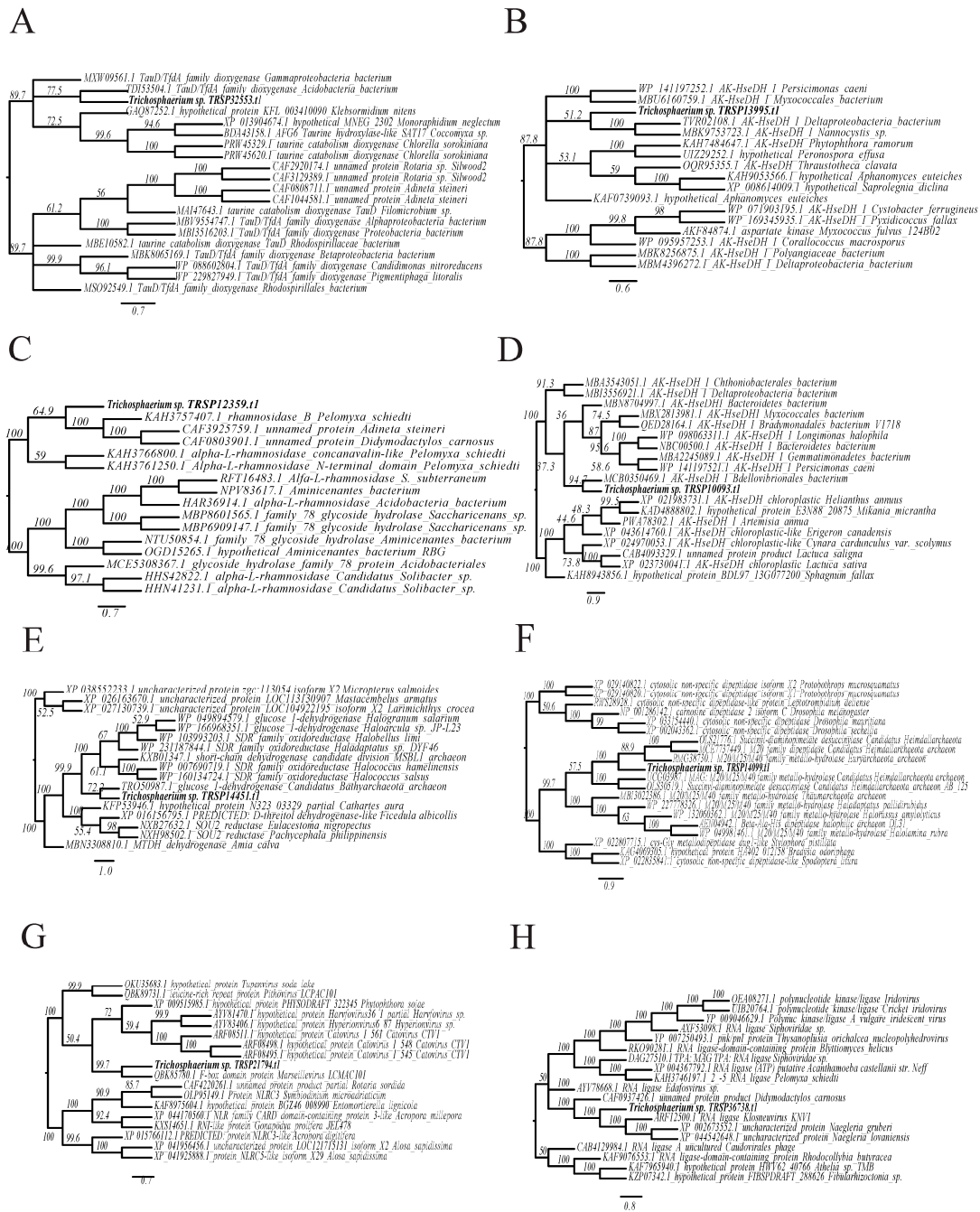


Figure S3. Phylogenetic reconstructions demonstrating putative lateral gene transfers (LGTs) in *Trichosphaerium* sp. genome among bacteria (a-d), archaea (e,f), and giant viruses (g,h). Clade supports at nodes are ML IQ-TREE 1000 ultrafast bootstrap values. All branches are drawn to scale.

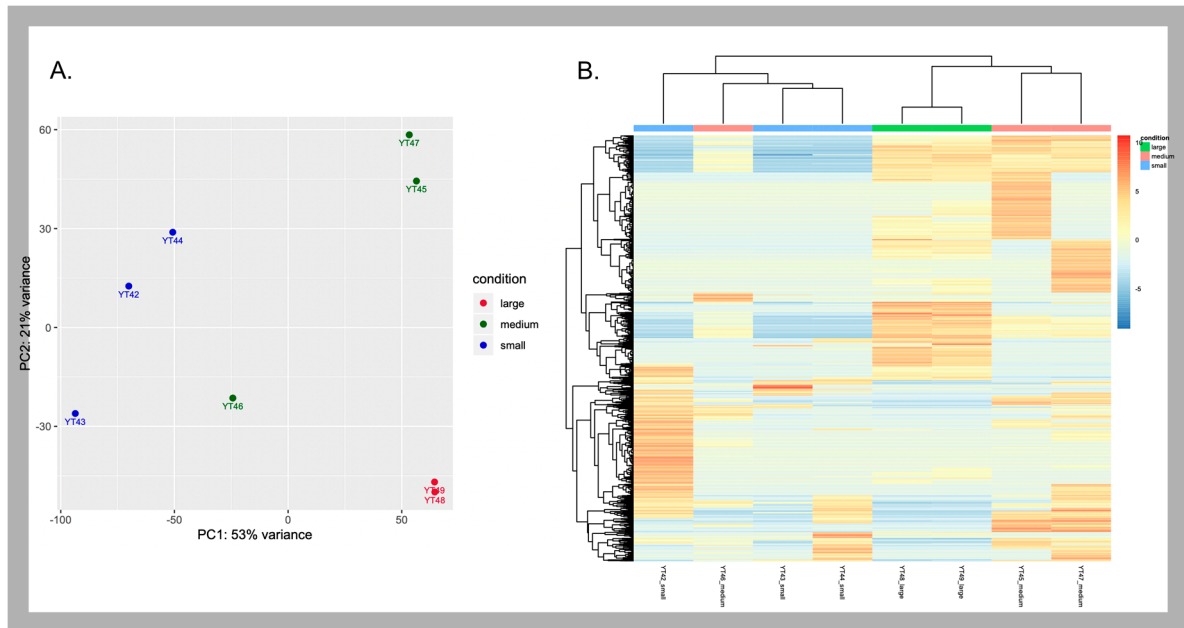


Figure S4. PCA plot (a) and Clustered heatmap (b) of differentially expressed genes (DEGs) for 8 samples including small (3 replicates), medium (3 replicates) and large (2 replicates) cells. PCA data for the plot were the transformed normalized counts of each sample generated from DESeq2. The color scale in (b) from red (highly expressed) to blue (low expression) represents the transformed, normalized counts from a variance stabilizing transformation.

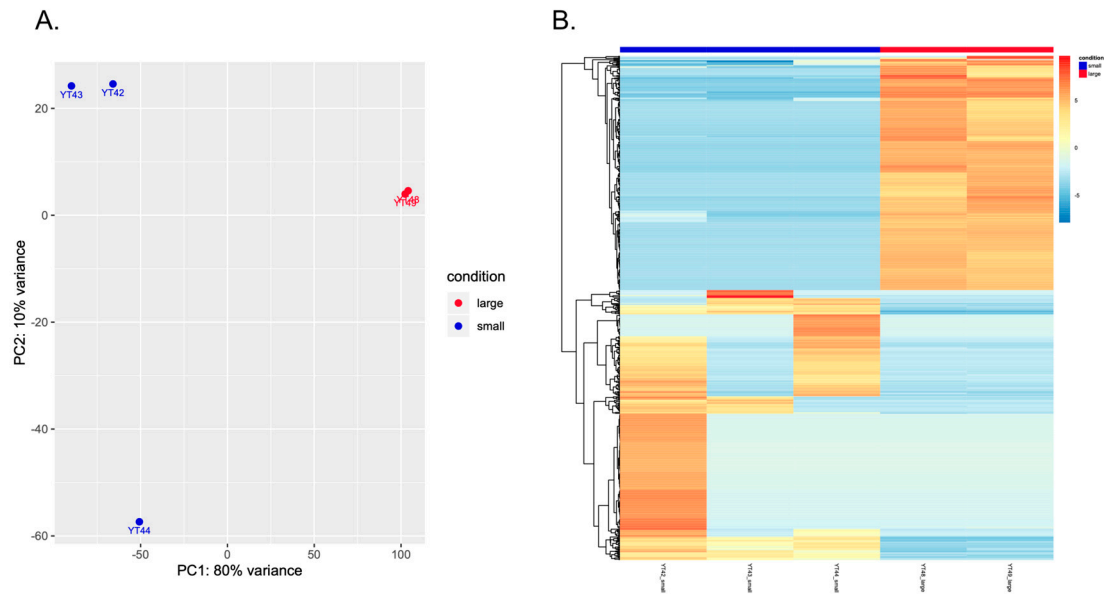


Figure S5. PCA plot (a) and Clustered heatmap (b) of differentially expressed genes (DEGs) for 5 samples including small (3 replicates) and large (2 replicates). PCA data for the plot were the transformed normalized counts of each sample generated from DESeq2. The color scale in (B) from red (highly expressed) to blue (low expression) represents the transformed, normalized counts from a variance stabilizing transformation.

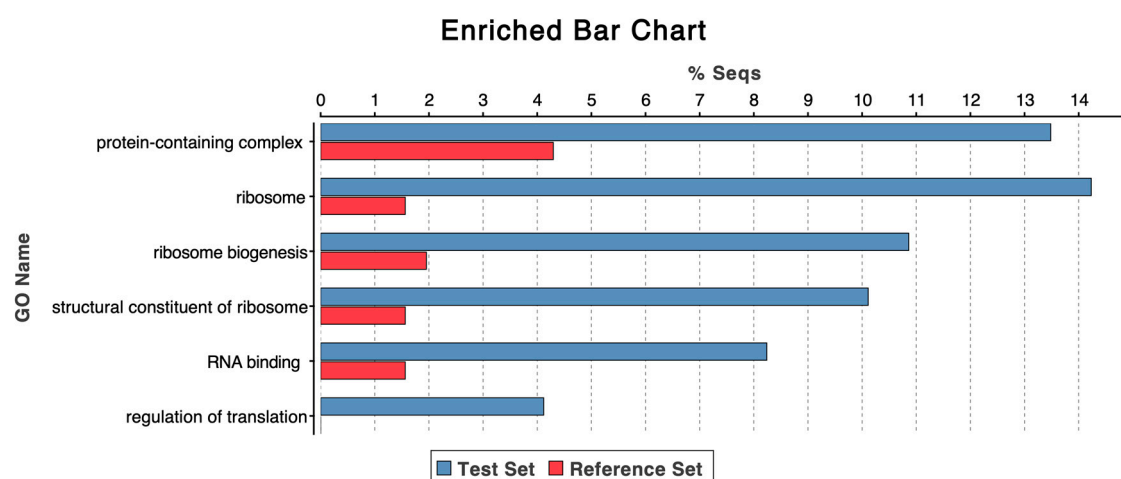


Figure S6. GO (gene ontology) enrichment of cellular processes in small (red bar) and large (blue bars) cells of *Trichosphaerium* sp.

Expression of Sexual related Genes

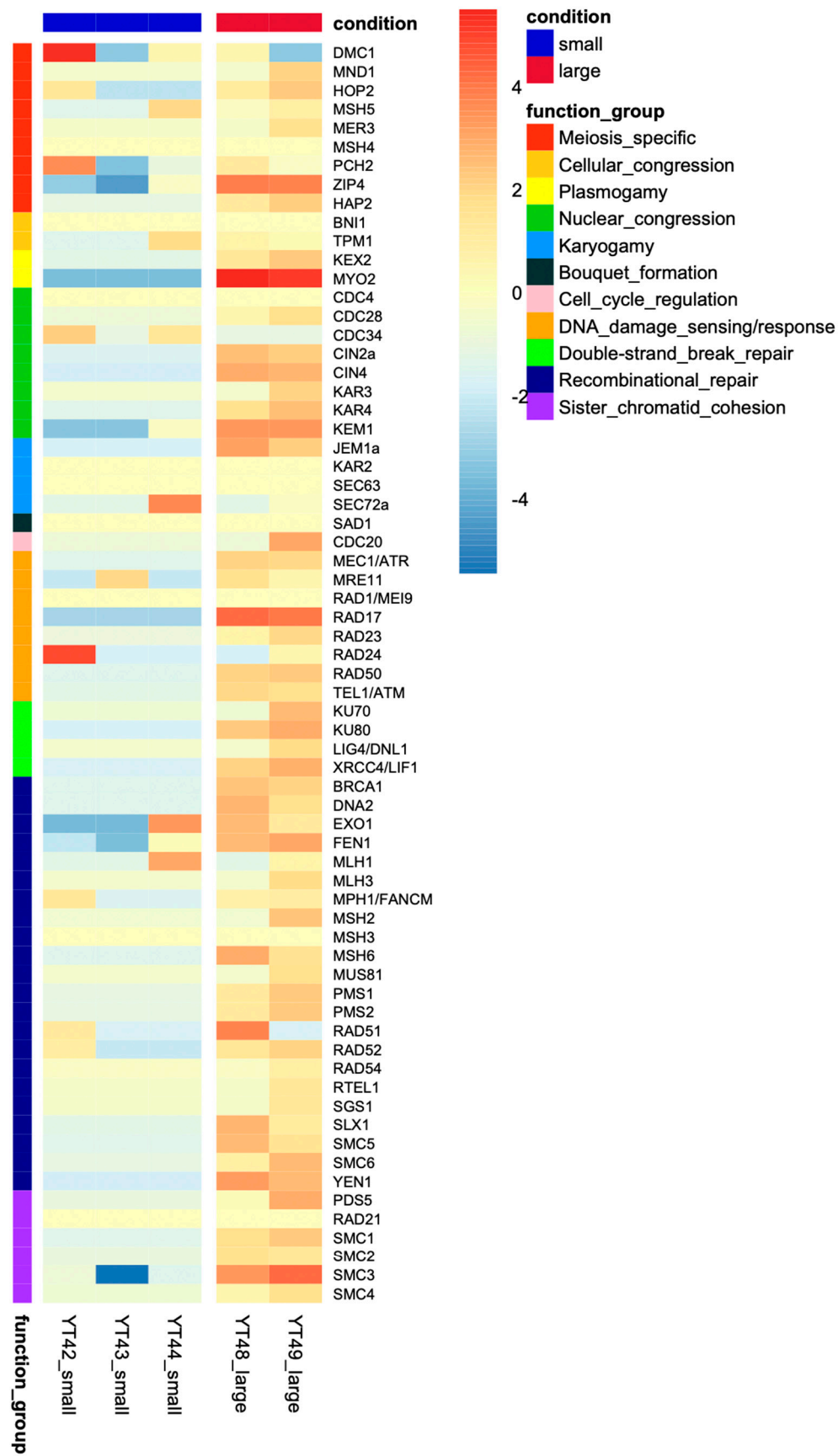


Figure S7. Heatmap of 67 genes (meiosis and sexual-related) grouped according to their functional categories in small (3) and large (2) cell samples. The color scale from orange (highly expressed) to blue (low expression) represents the transformed, normalized counts from a variance stabilizing transformation.

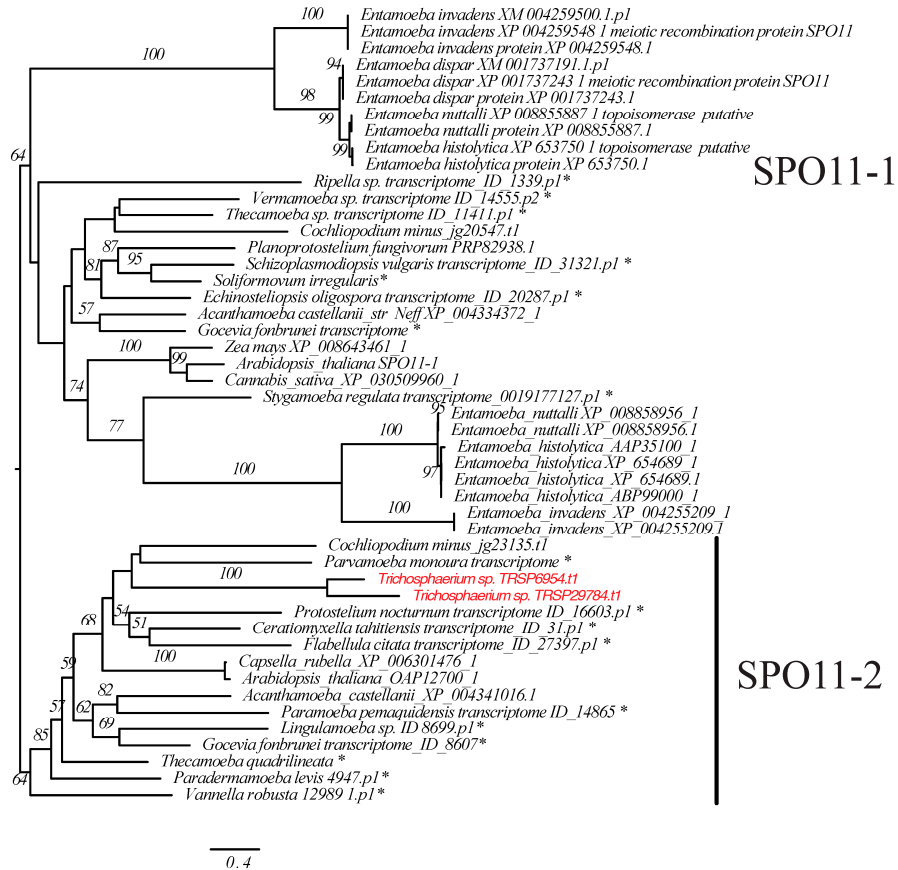


Figure S8. Phylogenetic reconstructions of SPO11 paralogs from amoebozoans and other eukaryotes. Amoebozoans without genome data are represented by an asterisk (*) to indicate the data come from RNA-seq data. Clade supports at nodes are ML IQ-TREE 1000 ultrafast bootstrap values. All branches are drawn to scale.