

Supplementary Material

1 PSEUDO CODE

Algorithm 1 Association maximization with greedy forward stepwise selection	on
procedure SELECTIONENERGY($\mathbf{Z}', \mathbf{y}, \alpha$, patience) if PERMDISP2($\mathbf{Z}'\mathbf{y}$) < α then testStatisticFunction = $cF($)	> Determine test statistic
Metric = combinedF'	
else	
testStatisticFunction = $F_{n,\alpha}()$ Metric = 'discoF'	
end if	
$\mathbf{X} = \mathbf{Z}' \in \mathbb{R}^{n imes 3}$	Select first 3 columns
maxF = testStatisticFunction(baseSet, y)	
improvementTime = 0	
for $i \in [4, \dots, \mathbf{Z}']$ do	
$\mathbf{X}_{new} = \mathbf{X} \cup z'_{*,i}$	\triangleright Append <i>i</i> th column
newF = testStatisticFunction(newSet, y)	
diff = newF - maxF	
if diff $\geq eps$ then	
$\mathbf{X} = \mathbf{X}_{new}$	
maxF = newF	
improvementTime = 0	
else	
improvementTime = improvementTime +1	
end if	
if improvementTime > patience then	
Break	
end if	
end for	
return(X, testStat)	
end procedure	

2 SUPPLEMENTARY FIGURES



Figure S1. Feature selection computational time comparisons for balanced and unbalanced sampling designs between SelEnergyPerm, LASSO, RFE, RF, Information Gain, and Boruta across each scenario and dimension. Points are the mean for each experimental condition.



Figure S2. Comparison of SelEnergyPerm-selected log-ratio subset characteristics with Boruta, Information Gain Filtering, LASSO, and RFE across five simulation scenarios for the unbalanced sampling design. Using 200 simulations for each scenario-dimension by method we assessed: (Top Row) the clustering coefficient of log ratio networks formed by selected subsets returned from each method, (Middle Row) the magnitude of the association as measured by the cF-statistic on selected subsets returned from each method, and (Bottom Row) the number of log ratios returned by each method. Points are the mean for each experimental condition and error bars indicate 95% confidence interval.



Figure S3. Overall mean performance comparison for data generated from synthetic distributions aggregated across all scenarios and dimensions using MCC, Sensitivity, Specify, Positive predictive value (PPV), Negative predictive value (NPV), Youden Index, and False Positive Rate (FPR) metric. Error bars indicate standard error.



Figure S4. Overall mean performance comparison for data generated from 16S and WGS synthetic data aggregated across all scenarios and effect levels using MCC, Sensitivity, Specify, Positive predictive value (PPV), Negative predictive value (NPV), Youden Index, and False Positive Rate (FPR) metric. Error bars indicate standard error.