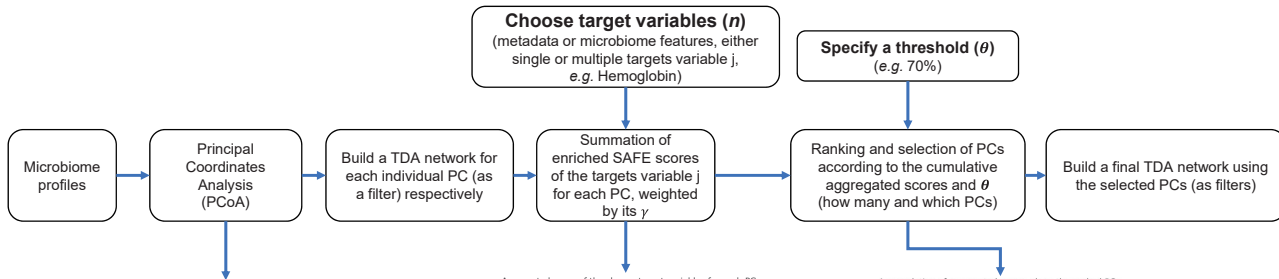
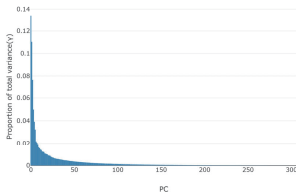


Supervised selection of PCs based on target variables of metadata or microbiome features



Proportion of total variance explained by each PC



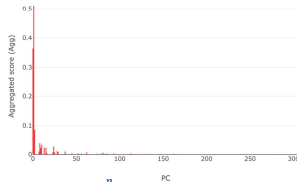
$$\gamma_i = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i}$$

λ_i : eigenvalue of PC_i

m : number of PCs

γ_i : proportion of total variance explained by PC_i

Aggregated score of the chosen target variables for each PC



$$Agg_i = \gamma_i \sum_{j=1}^n S_j$$

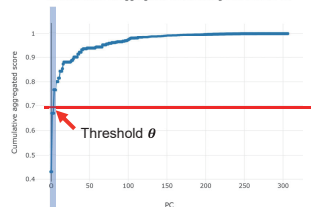
γ_i : proportion of total variance explained by PC_i

S_j : enriched SAFE score for target variable j

n : number of target variables

Agg_i : aggregated score of the target variables for PC_i

Accumulation of aggregated scores along the ranked PCs



Selected PCs: PC1 and PC2