

Simulation Studies

1. Simulation for comparing the performance of different methods on main simulated settings:

(1) Run the script `~/manuscript/simulation/1-simulation.R` to obtain the results of power and FDR for the proposed method (PoDA) and the benchmark methods under different simulation settings. Note that because the codes across different simulation settings differ only slightly, we have consolidated all settings into a single R script with corresponding annotations. To replicate a specific setting, simply uncomment the corresponding code block. All results can be found at `~/manuscript/simulation/Results`

(2) Run the scripts `~/manuscript/simulation/2-plot_A7.R` and `~/manuscript/simulation/2-plot_exceptforA7.R` to obtain the figures showing the results. All figures can be found at `~/manuscript/simulation/Plots`.

2. Simulation for evaluating the impact of different zero-handling strategies on PoDA, LinDA, and ANCOMBC:

(1) Run script `3-zero-simulation.R` to obtain the results of power and FDR for the proposed method (PoDA) and the benchmark methods under different simulation settings. Note that because the codes across different simulation settings differ only slightly, we have consolidated all settings into a single R script with corresponding annotations. To replicate a specific setting, simply uncomment the corresponding code block. All results can be found at `~/manuscript/simulation/Results`

(2) Run scripts `~/manuscript/simulation/4-barplot-A2&A10-LinDA&ANCOMBC.R`, `~/manuscript/simulation/4-barplot-A2&A10-PoDA.R`, `~/manuscript/simulation/4-linegraph-A1-LinDA&ANCOMBC.R`, and `~/manuscript/simulation/4-linegraph-A1-PoDA.R` to obtain the figures showing the results. All figures can be found at `~/manuscript/simulation/Plots`.

3. Simulation for evaluating the impact of different bias term update choices on PoDA:

Run script `~/manuscript/simulation/s-three_version_of_poda/examination_A1.R` to obtain the results and obtain the figures showing the results. The results can be found at `~/manuscript/simulation/s-three_version_of_poda`.

4. Simulation for evaluating the impact of different information-splitting repetition numbers on PoDA:

Run script `~/manuscript/simulation/s-rep_time_examination/s-rep_time_examination`

`.R` to obtain the results and obtain the figures showing the results. The results can be found at `~/manuscript/simulation/s-rep_time_examination`.

Real data studies

1. Perform differential abundance analysis on two PD datasets using different methods:

- (1) Run the script `~/manuscript/realdata/PD/0-pd16sdata.R` to obtain and process the data.
- (2) Run the script `~/manuscript/realdata/PD/1-type1error.R` to obtain the results on Type I error control for each method. Run script `~/manuscript/realdata/PD/2-plot_type1error.R` to obtain the figure showing the results. The figure can be found at `~/manuscript/realdata/Plots`.
- (3) Run scripts `~/manuscript/realdata/PD/3-wallen1_detection.R` and `~/manuscript/realdata/PD/3-wallen2_detection.R` to obtain the results of differentially abundant genera identified by each method on Wallen1 and Wallen2.
- (4) Run script `~/manuscript/realdata/PD/4-main_plot.R` to obtain the figures showing the results of differential abundance analysis. The figure can be found at `~/manuscript/realdata/Plots`.
- (5) Run script `~/manuscript/realdata/PD/5-plot_lub.R` to obtain the boxplots of the relative abundances in the Lub6 and Lub12 datasets of the 14 common genera identified by PoDA in both Wallen1 and Wallen2 datasets, which appeared in more than 50% of the subjects in both Lub6 and Lub12 datasets.

2. Perform differential abundance analysis on the CDI dataset using different methods:

- (1) Run script `~/manuscript/realdata/CDI/1-CDI_detection.R` to obtain the results of differentially abundant genera identified by each method on CDI.
- (2) Run script `~/manuscript/realdata/CDI/2-zero-replacement_robust.R` to obtain the results on the impact of different zero-handling strategies on method performance for this dataset.
- (3) Run script `~/manuscript/realdata/CDI/3-main_plot.R` to obtain the figure showing the results. The figure can be found at `~/manuscript/realdata/Plots`.