# RE: Concern about unequal sample sizes

Date: 7/12/2022

## Sample size tables

Samples per Treatment

|  |  |
| --- | --- |
| **Treatment** | **N** |
| Alone | 27 |
| Competition | 73 |

Samples per Block

|  |  |
| --- | --- |
| **Block** | **N** |
| 1 | 24 |
| 2 | 25 |
| 3 | 26 |
| 4 | 25 |

## Issues with unbalanced/unequal sample sizes

* Reduced statistical power
* IF variances are \*unequal\*

## Solutions:

* IF variances are similar, it is OK to Perform ANOVA anyways but keep in mind that power is reduced. 🡨 I argue that the distributions are fairly similar, see boxplots below.
* IF variances are not similar perform non-parametric test such as Kurk-Wallis
* Re-rerun analysis using a random and smaller subset of competition treatment, so they are comparable.

Violin plots

**Figure** showing distribution of various microbial community metrics according to treatment within *I. purpurea*.

Chart, radar chart

Description automatically generated

## Re-running analysis on smaller competition subset

Analysis was performed on data after randomly sub-setting for 30 samples from the competition treatment. Below I highlight and compare the main results from analysis on full data set and analysis on subset.

Linear Regression Results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent variable | Independent variable(s) | Full data set  B (p-value) | Subset data set  B (p-value) | Conserved? |
| Species richness | PC2 + Block + TRT | -5.734 (0.013) | -9.419 (0.0233) | Yes |
| Species evenness | PC2 + Block + TRT | 7.25 e-5 (0.027) | 1.382 e-04 (0.015) | Yes |
| Species Simpson diversity | PC4 + Block + TRT | 0.002 (p = 0.008) | 0.0027 (0.08) | No, but marginally significant |

ANCOVA results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent variable | Main independent variable(s) | Full data set  F-value (p-value) | Subset data set  F-value (p-value) | Conserved? |
| Relative fitness | Treatment × Species richness | 6.9 (0.01) | 0.032 (0.861) | No |
| Relative fitness | Treatment × Species evenness. | 7.34 (0.008) | 3.88 (0.064) | No, but marginally significant. |

Link to original analysis and results: <https://github.com/SaraMColom/Microbiome_2018/tree/modified/RCode#correlations-with-root-traits>

Link to analysis using subset of competition treatment and results: <https://github.com/SaraMColom/Microbiome_2018/tree/modified/AnalysisSubset>