mStat_generate_report_pair_example

Powered by MicrobiomeStat (Ver 1.1.1)

2023-10-11

Contents

1. Data overview and summary statistics

1.1 Parameter setting

Parameter	Value
data.obj	data.obj
feature.dat.type	count
group.var	group
test.adj.vars	NULL
vis.adj.vars	NULL
strata.var	sex
subject.var	subject
time.var	time
change.base	1
alpha.obj	alpha.obj
alpha.name	shannon, observed_species
alpha.change.func	log fold change
depth	NULL
dist.obj	dist.obj
dist.name	BC, Jaccard
pc.obj	NULL
prev.filter	0.1
abund.filter	1e-04
feature.analysis.rarafy	TRUE
feature.change.func	relative change
bar.area.feature.no	30
heatmap.feature.no	30
dotplot.feature.no	20
vis.feature.level	Genus
test.feature.level	Genus
feature.mt.method	none
feature.sig.level	0.1
feature.box.axis.transform	identity
base.size	18
theme.choice	bw
custom.theme	NULL
palette	NULL
pdf	TRUE
file.ann	NULL
pdf.wid	11
pdf.hei	8.5

1.2 Summary statistics

```
mStat_results <- mStat_summarize_data_obj(data.obj = data.obj,
                                          time.var = time.var,
                                          group.var = group.var,
                                          palette = palette)
```

Histogram of Sample Counts over Time

Boxplot of Sequencing Depth over Time
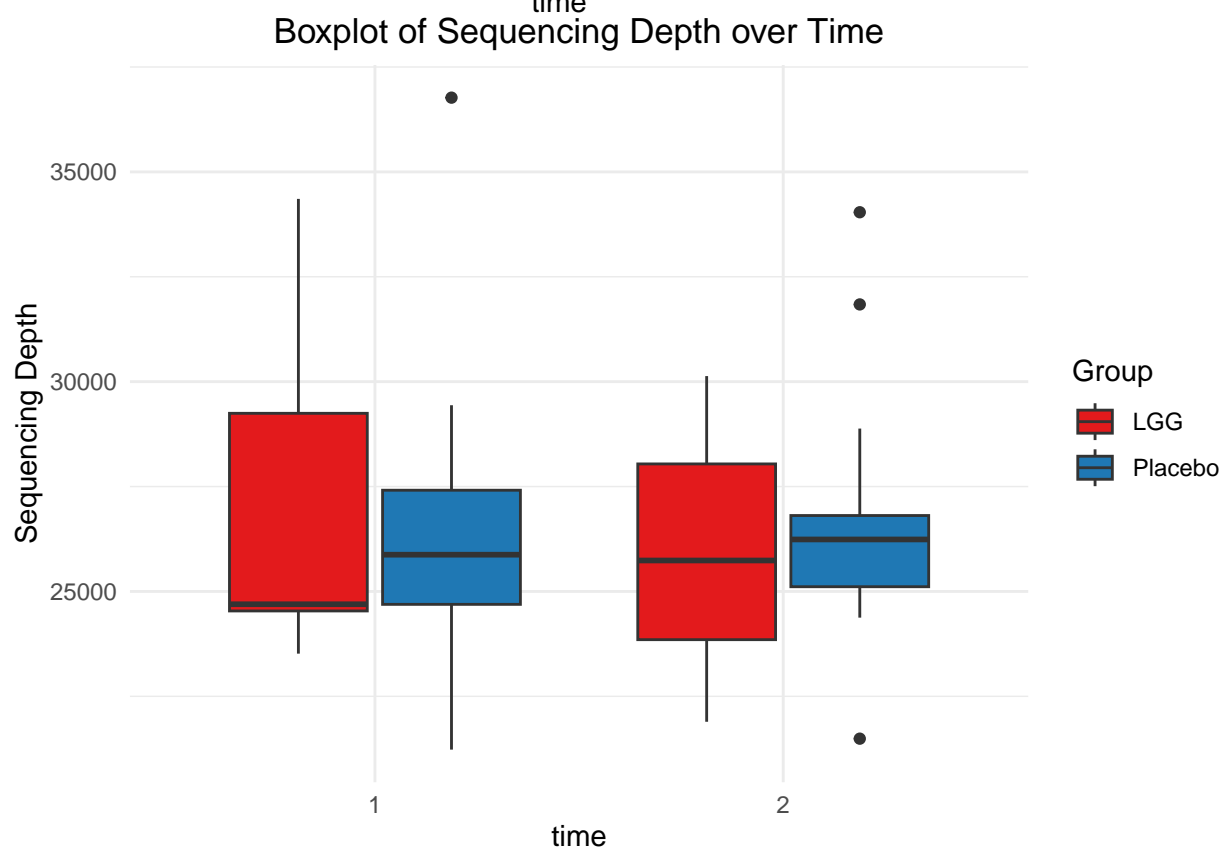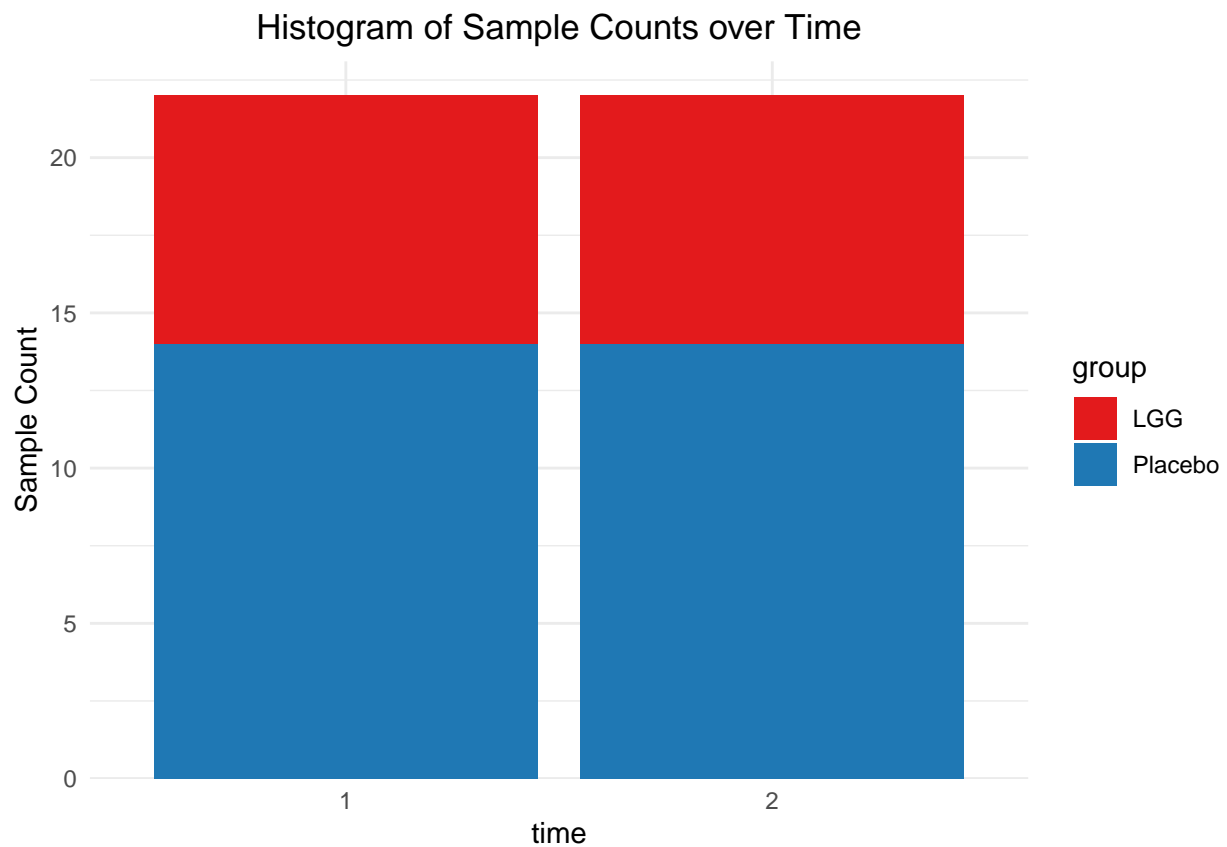
Category	Variable	Value
Basic Statistics	Number of samples	44
Basic Statistics	Number of features	130
Basic Statistics	Min. reads per sample	21230
Basic Statistics	Max. reads per sample	36769
Basic Statistics	Total reads across all samples	1167517
Basic Statistics	Average reads per sample	8980.9
Basic Statistics	Median reads per sample	25833
Basic Statistics	Proportion of zero counts	0.223
Basic Statistics	Count of features that only appear once	0
Metadata	Number of metadata variables	4
Feature Annotations	Proportion of missing annotations in Phylum	0
Feature Annotations	Proportion of missing annotations in Family	0
Feature Annotations	Proportion of missing annotations in Genus	0
Phylogenetic Tree	Exists in the dataset	No
Time-Series Information	Number of unique time points	2
Time-Series Information	Sample count at time point: 1	22
Time-Series Information	Sample count at time point: 2	22

No rarefaction depth is specified. The minimum depth, 21230, is used as the rarefaction depth. After rarefaction, 44 samples remain in the analysis. alpha.obj is calculated based on the rarefied data.obj. dist.obj is calculated based on the rarefied data.obj. pc.obj is calculated based on the dist.obj using multi-dimensional scaling.

1.3 Data visualization (overall)

Rarefaction has been enabled for feature-level analysis and visualization.

Reason: The observed abundance of rare/low-abundance features can be strongly influenced by the sequencing depth. Rarefaction is an effective method to control the effect of sequencing depth variation. By employing rarefaction, presence/absence status of the features are more comparable and we can potentially increase the power of detecting those rare/low-abundance features, even though it introduces some variation due to under-sampling. In essence, this step improves comparability across samples across samples with varying sequencing depth.

If you do not wish to perform rarefaction during feature-level analysis, please turn feature.analysis.rarafy to FALSE.

1.3.1 Feature heatmap

```
taxa_heatmap_pair_results <- generate_taxa_heatmap_pair(
  data.obj = data.obj,
  subject.var = subject.var,
  time.var = time.var,
  group.var = group.var,
  strata.var = strata.var,
  feature.level = vis.feature.level,
  feature.dat.type = feature.dat.type,
  features.plot = NULL,
  top.k.plot = heatmap.feature.no,
  top.k.func = 'mean',
  prev.filter = prev.filter,
  abund.filter = abund.filter,
```

```
  base.size = base.size,
  palette = palette,
  cluster.rows = NULL,
  cluster.cols = NULL,
  pdf = pdf,
  file.ann = file.ann,
  pdf.wid = pdf.wid,
  pdf.hei = pdf.hei
)
```

The following plots display the average proportions for each time point, group, and stratum.
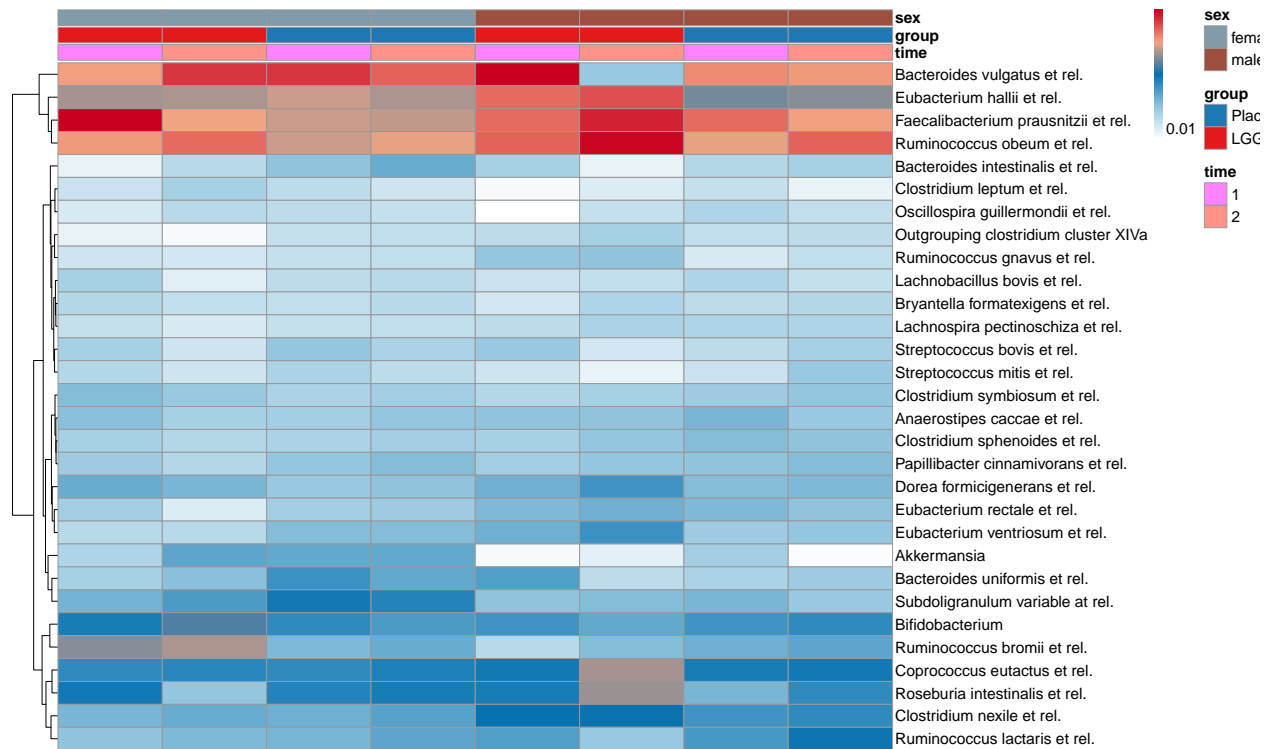
$Genus

The following plots display the individual proportions for each sample.

$Genus
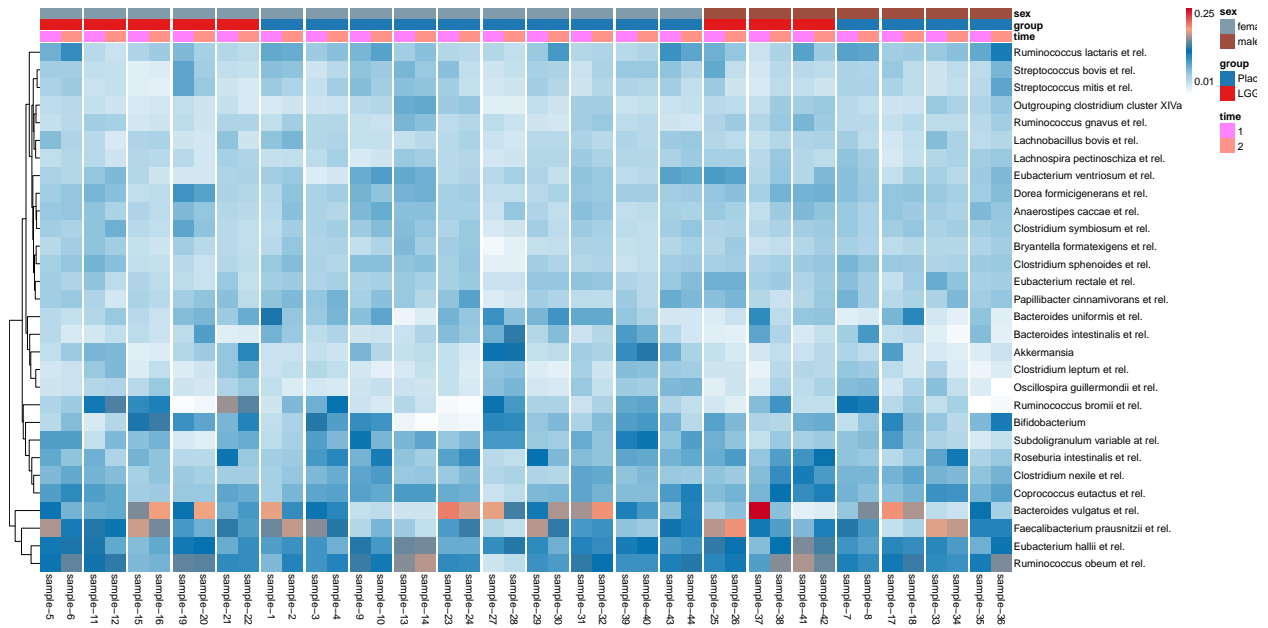
1.3.2 Feature barplot

```
taxa_barplot_pair_results <- generate_taxa_barplot_pair(
  data.obj = data.obj,
  subject.var = subject.var,
  time.var = time.var,
  group.var = group.var,
  strata.var = strata.var,
  feature.level = vis.feature.level,
  feature.dat.type = feature.dat.type,
  feature.number = bar.area.feature.no,
  base.size = base.size,
  theme.choice = theme.choice,
  custom.theme = custom.theme,
  palette = palette,
  pdf = pdf,
  file.ann = file.ann,
  pdf.wid = pdf.wid,
  pdf.hei = pdf.hei
)
```

The following plots display the average proportions for each time point, group, and stratum.

$Genus

The following plots display the individual proportions for each sample.

$Genus
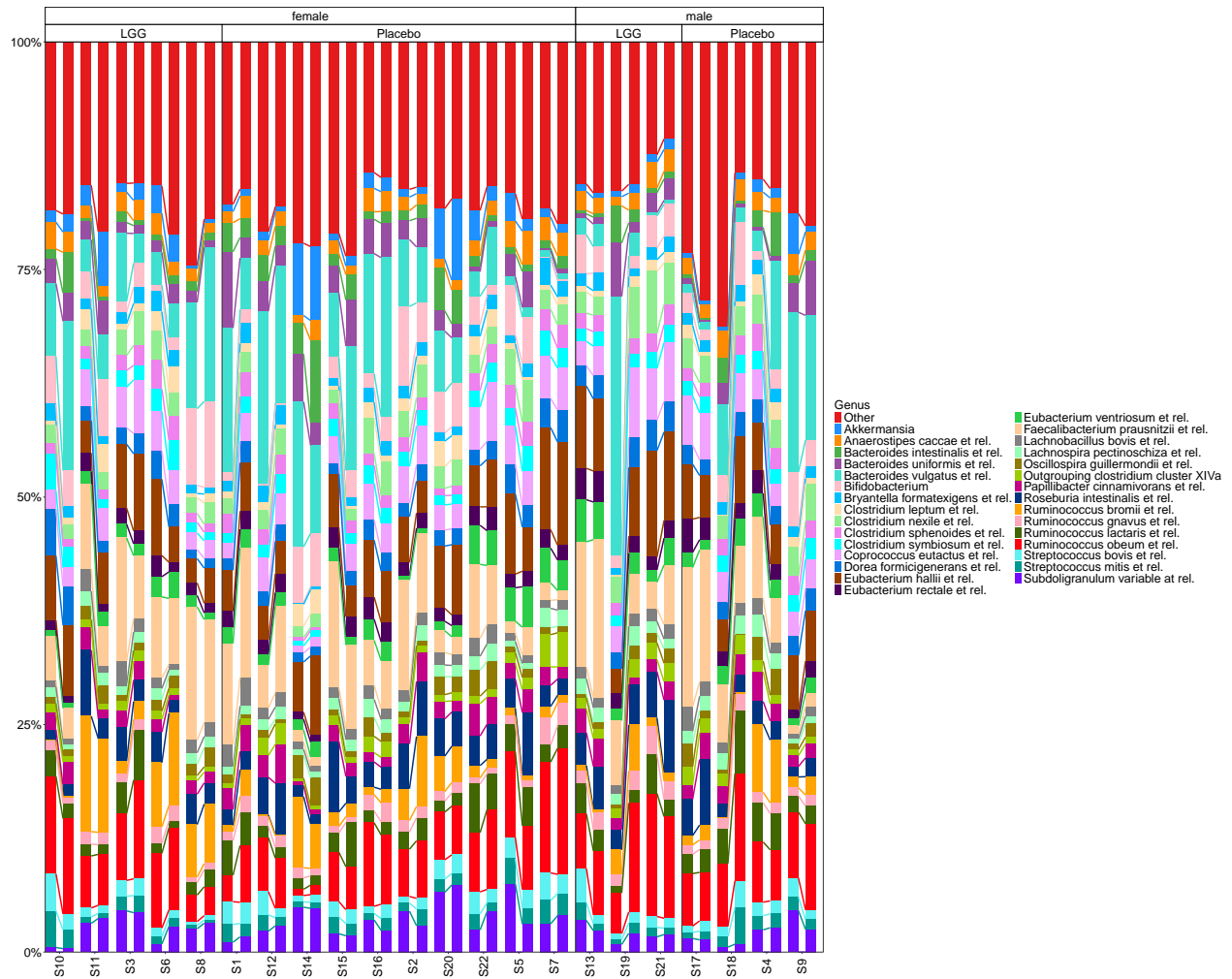
1.3.3 Feature dotplot

```
taxa_dotplot_results <- generate_taxa_dotplot_pair(
                                        data.obj = data.obj,
                                        subject.var = subject.var,
                                        time.var = time.var,
                                        group.var = group.var,
                                        strata.var = strata.var,
                                        feature.level = vis.feature.level,
                                        feature.dat.type = feature.dat.type,
                                        features.plot = NULL,
                                        top.k.plot = dotplot.feature.no,
                                        top.k.func = 'mean',
                                        prev.filter = prev.filter,
                                        abund.filter = abund.filter,
                                        base.size = base.size,
                                        theme.choice = theme.choice,
                                        custom.theme = custom.theme,
                                        palette = palette,
                                        pdf = pdf,
                                        file.ann = file.ann,
```

```
                                           pdf.wid = pdf.wid,
                                           pdf.hei = pdf.hei)
```

The following plots display the average proportions for each time point, group, and stratum.

$Genus
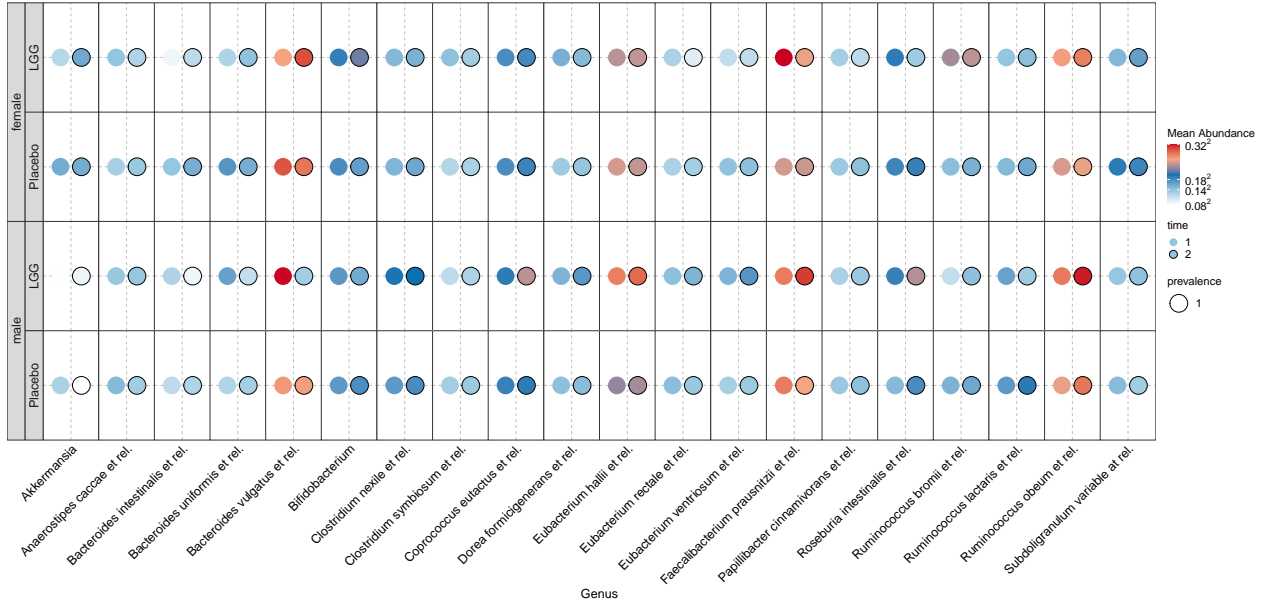
1.3.4 Feature change heatmap

```
taxa_change_heatmap_results <- generate_taxa_change_heatmap_pair(
                                    data.obj = data.obj,
                                    subject.var = subject.var,
                                    time.var = time.var,
                                    change.base = change.base,
                                    group.var = group.var,
                                    strata.var = strata.var,
                                    feature.level = vis.feature.level,
                                    feature.change.func = feature.change.func,
                                    feature.dat.type = feature.dat.type,
                                    features.plot = NULL,
                                    top.k.plot = heatmap.feature.no,
                                    top.k.func = 'mean',
                                    prev.filter = prev.filter,
                                    abund.filter = abund.filter,
                                    base.size = base.size,
                                    palette = palette,
                                    cluster.cols = NULL,
                                    cluster.rows = NULL,
                                    pdf = pdf,
                                    file.ann = file.ann,
                                    pdf.wid = pdf.wid,
                                    pdf.hei = pdf.hei)
```

The changes were relative changes, which were computed as (after.abund - before.abund) / (after.abund + before.abund) so the values lie between [-1, 1]. The following plots display the change for each subject.

$Genus
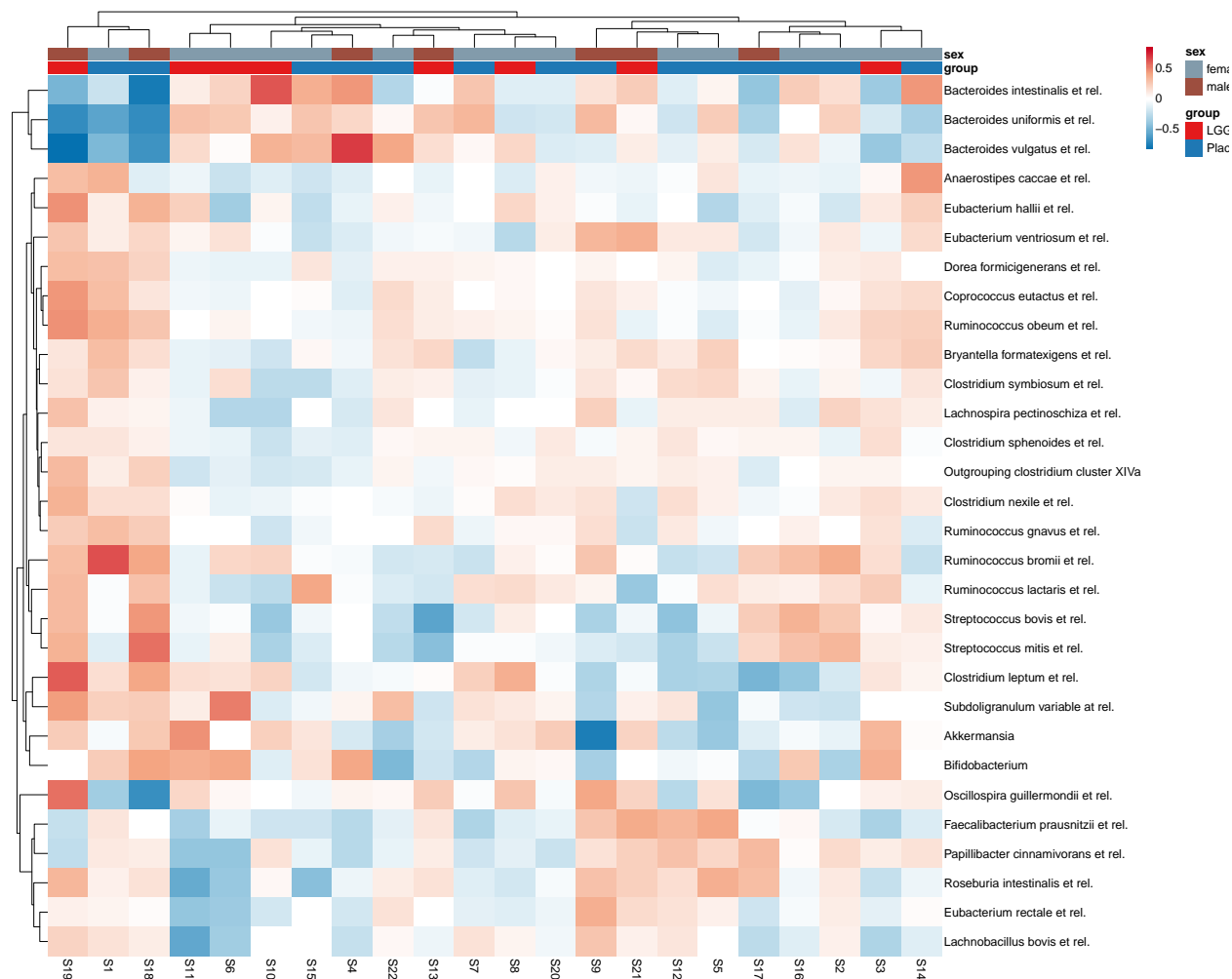
The following plots display the average change for each time point, group, and stratum.

$Genus

1.3.5 Feature change dotplot

```r
taxa_change_dotplot_results <- generate_taxa_change_dotplot_pair(
                                    data.obj = data.obj,
                                    subject.var = subject.var,
                                    time.var = time.var,
                                    change.base = change.base,
                                    feature.change.func = feature.change.func,
                                    group.var = group.var,
                                    strata.var = strata.var,
                                    feature.level = vis.feature.level,
                                    feature.dat.type = feature.dat.type,
                                    features.plot = NULL,
                                    top.k.plot = dotplot.feature.no,
                                    top.k.func = 'mean',
                                    prev.filter = prev.filter,
                                    abund.filter = abund.filter,
                                    base.size = base.size,
                                    theme.choice = theme.choice,
                                    custom.theme = custom.theme,
                                    palette = palette,
```
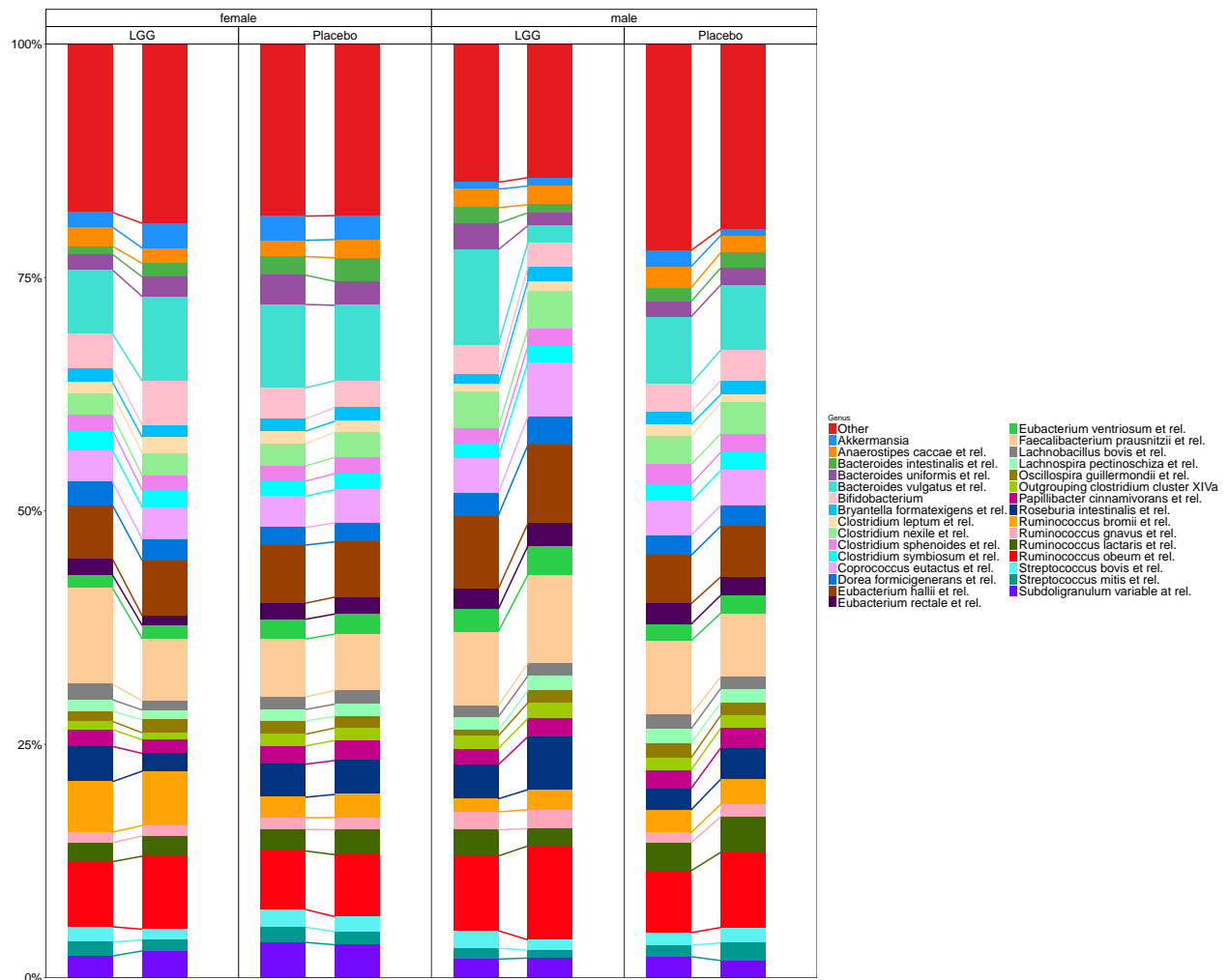
```
                                             pdf = pdf,
                                             file.ann = file.ann,
                                             pdf.wid = pdf.wid,
                                             pdf.hei = pdf.hei)
```

The changes were relative changes, which were computed as (after.abund - before.abund) / (after.abund + before.abund) so the values lie between [-1, 1]. The following plots display the average change for each group, and stratum.

$Genus
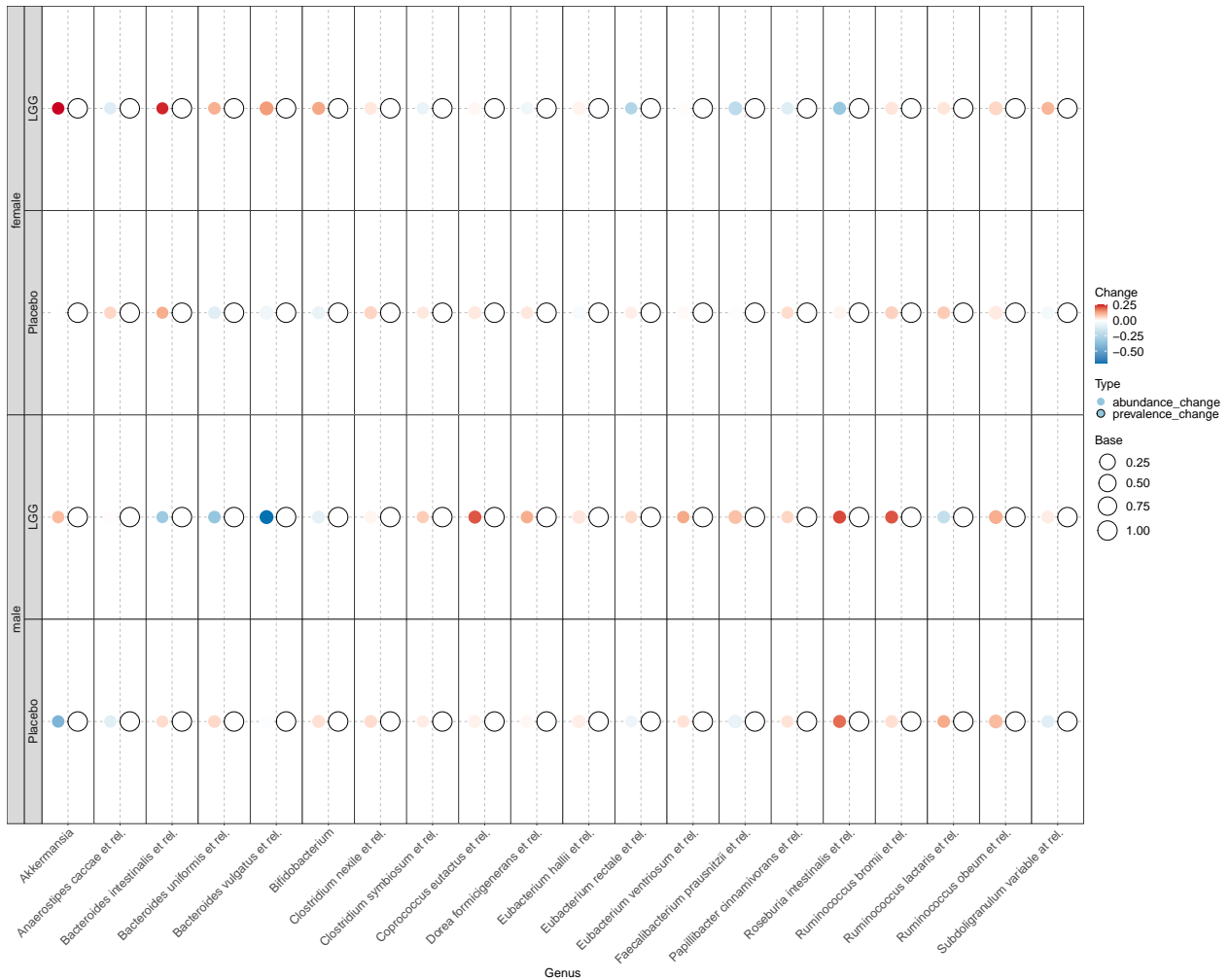
2. Alpha diversity analysis

2.1 Data visualization

2.1.1 Alpha diversity boxplot

```
alpha_boxplot_results <- generate_alpha_boxplot_long(
                                             data.obj = data.obj,
                                             alpha.obj = alpha.obj,
                                             alpha.name = alpha.name,
                                             depth = depth,
```

```
                                       subject.var = subject.var,
                                       time.var = time.var,
                                       t0.level = change.base,
                                       ts.levels = NULL,
                                       group.var = group.var,
                                       strata.var = strata.var,
                                       base.size = base.size,
                                       theme.choice = theme.choice,
                                       custom.theme = custom.theme,
                                       palette = palette,
                                       pdf = pdf,
                                       file.ann = file.ann,
                                       pdf.wid = pdf.wid,
                                       pdf.hei = pdf.hei)
alpha_boxplot_results
```
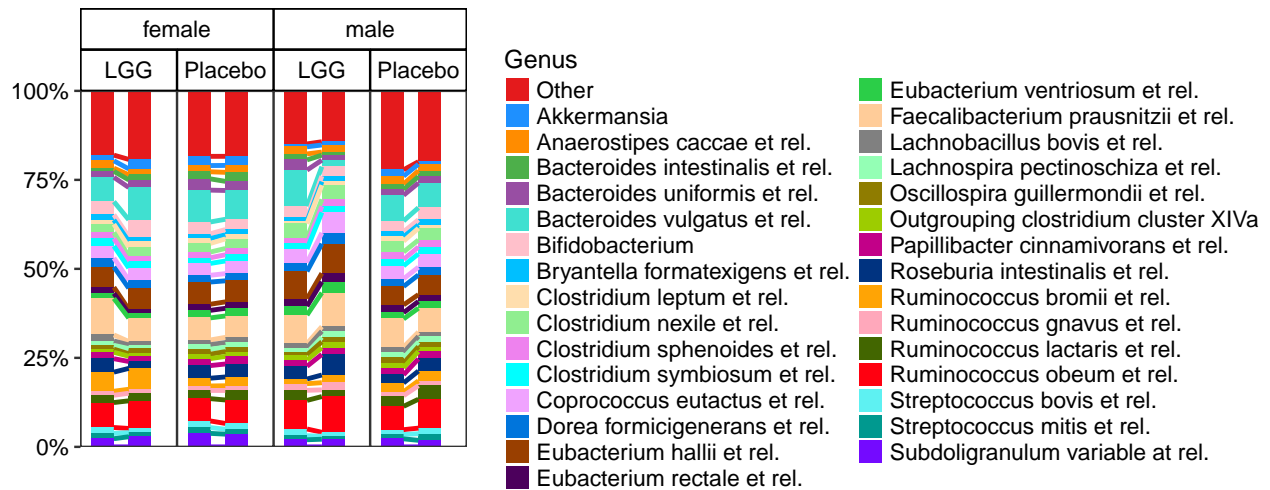

$shannon
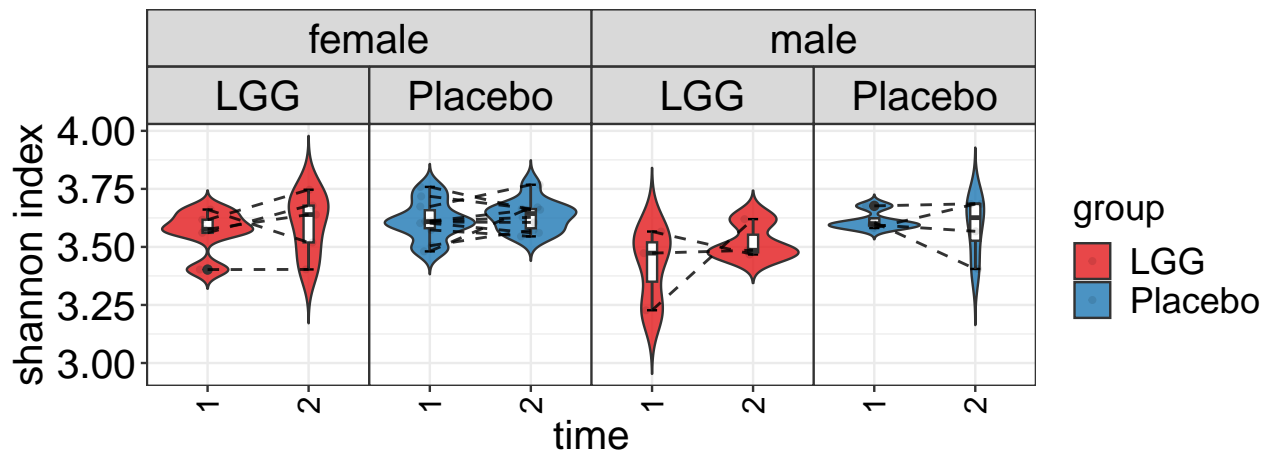
$observed_species
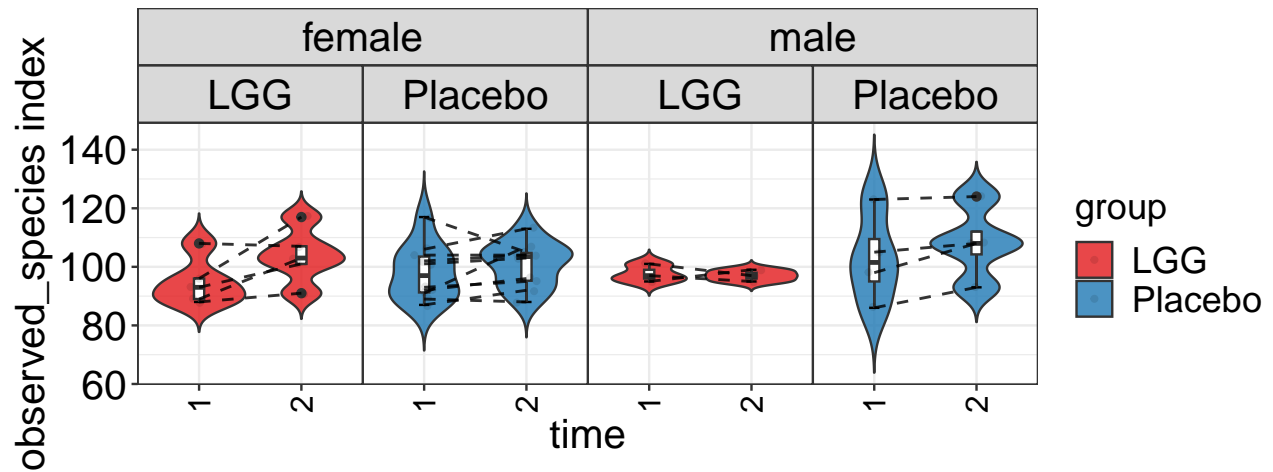
2.1.2 Alpha diversity change boxplot

```
alpha_change_boxplot_results <- generate_alpha_change_boxplot_pair(
                                    data.obj = data.obj,
                                    alpha.obj = alpha.obj,
                                    alpha.name = alpha.name,
                                    depth = depth,
                                    subject.var = subject.var,
                                    time.var = time.var,
                                    change.base = change.base,
                                    alpha.change.func = alpha.change.func,
                                    group.var = group.var,
                                    strata.var = strata.var,
                                    base.size = base.size,
                                    theme.choice = theme.choice,
                                    custom.theme = custom.theme,
                                    palette = palette,
                                    pdf = pdf,
                                    file.ann = file.ann,
                                    pdf.wid = pdf.wid,
                                    pdf.hei = pdf.hei)
```

The changes from change.base were computed as the log2 fold change of alpha diversity at the current timepoint versus change.base.

$shannon

$observed_species

2.2 Alpha diversity association test based on LMM

```
alpha_test_results <- generate_alpha_test_pair(data.obj = data.obj,
                                               alpha.obj = alpha.obj,
                                               alpha.name = alpha.name,
                                               depth = depth,
                                               time.var = time.var,
                                               subject.var = subject.var,
                                               group.var = group.var,
                                               adj.vars = test.adj.vars)
```
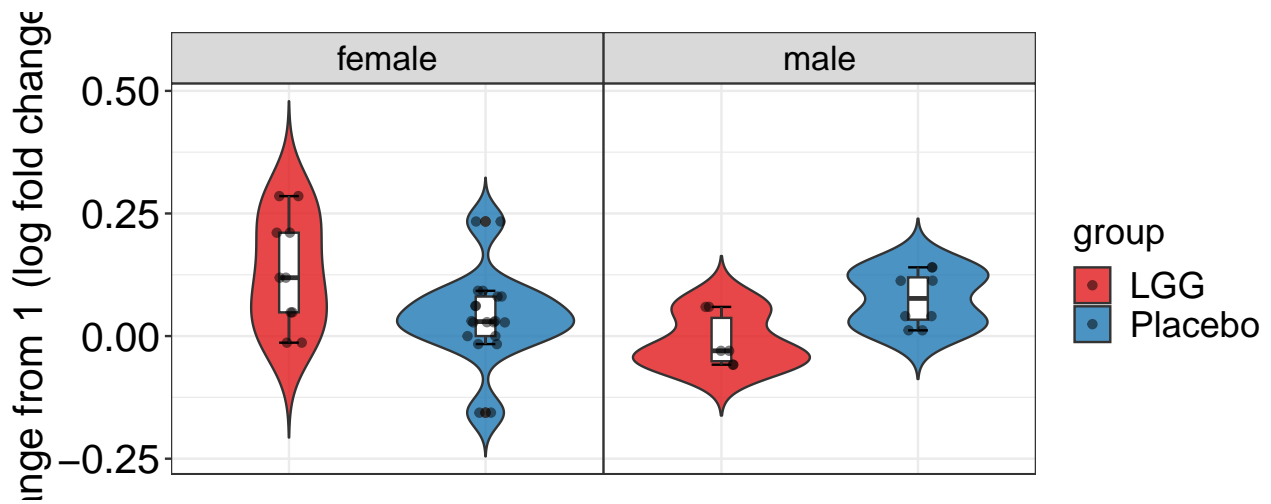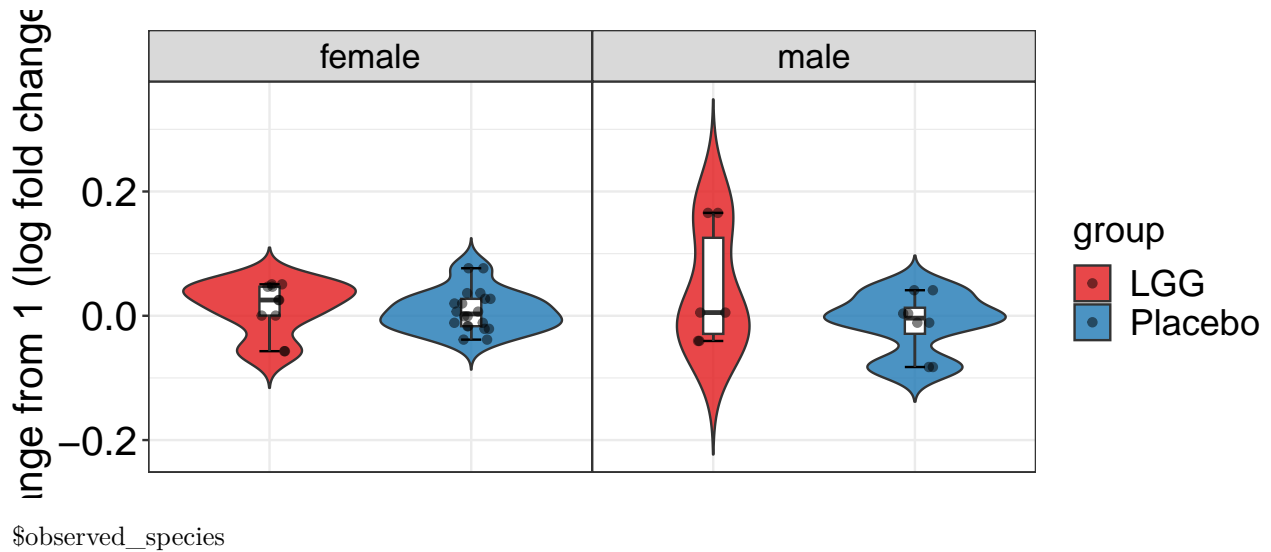
In this analysis, we utilized a linear mixed effects model with a random intercept and possibly a random slope for time to investigate a potential difference in alpha diversity across different levels of group. The model includes an interaction term between time and group. Specifically, we included time and group as covariates.

2.2.1 Shannon index

Based on the linear mixed effects model, a significant group difference was observed between LGG and Placebo of the variable group, with a p-value of 0.026. Based on the linear mixed effects model, no significant

trend difference was detected between LGG and Placebo of the variable group, with a p-value of 0.340.

Term	Estimate	Std.Error	Statistic	P.Value
(Intercept)	3.511	0.03564	98.51	3.325e-47
groupPlacebo	0.1037	0.04467	2.322	0.02574
time2	0.05937	0.04393	1.351	0.1916
groupPlacebo:time2	-0.05383	0.05507	-0.9775	0.34

2.2.2 Observed species index

Based on the linear mixed effects model, no significant group difference was detected between LGG and Placebo of the variable group, with a p-value of 0.380. Based on the linear mixed effects model, no significant trend difference was detected between LGG and Placebo of the variable group, with a p-value of 0.518.

Term	Estimate	Std.Error	Statistic	P.Value
(Intercept)	95.87	3.306	29	1.039e-21
groupPlacebo	3.696	4.144	0.8919	0.3804
time2	5.375	2.533	2.122	0.04654
groupPlacebo:time2	-2.089	3.176	-0.6579	0.5181

2.3 Alpha diversity association test based on changes

```
alpha_change_test_results <- generate_alpha_change_test_pair(
                                    data.obj = data.obj,
                                    time.var = time.var,
                                    change.base = change.base,
                                    subject.var = subject.var,
                                    group.var = group.var,
                                    adj.vars = test.adj.vars,
                                    alpha.obj = alpha.obj,
                                    alpha.name = alpha.name,
                                    depth = depth,
                                    alpha.change.func = alpha.change.func)
```

In this analysis, we utilized a general linear model to examine the influence of the variable group on change. The alpha diversity change is calculated by taking the logarithm of the fold change between consecutive time points.

2.3.1 Shannon index

Based on the general linear model, the level Placebo of the variable group did not significantly differ from level LGG, with a p-value of 0.333.

Term	Estimate	Std.Error	Statistic	P.Value
(Intercept)	0.02453	0.01806	1.359	0.1894
groupPlacebo	-0.02243	0.02264	-0.9911	0.3335

2.3.2 Observed species index

Based on the general linear model, the level Placebo of the variable group did not significantly differ from level LGG, with a p-value of 0.530.

Term	Estimate	Std.Error	Statistic	P.Value
(Intercept)	0.07766	0.03563	2.18	0.04139
groupPlacebo	-0.02853	0.04466	-0.6387	0.5302

3. Beta diversity analysis

3.1 Data visualization
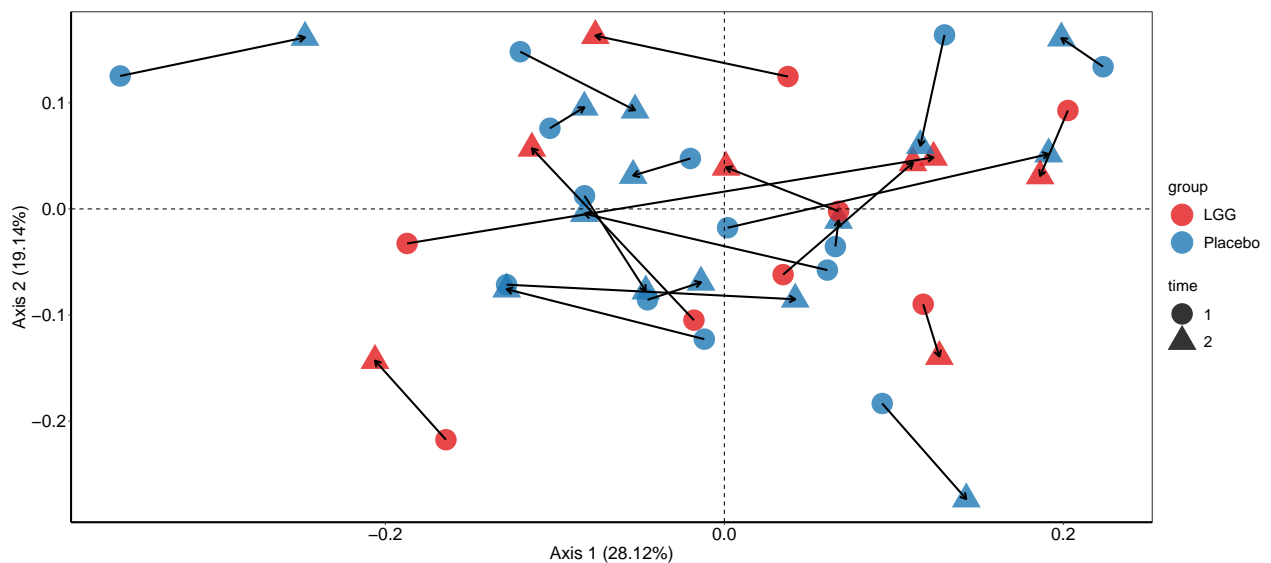
3.1.1 Beta diversity ordinationplot

```r
beta_ordination_results <- generate_beta_ordination_pair(
                                          data.obj = data.obj,
                                          dist.obj = dist.obj,
                                          pc.obj = pc.obj,
                                          subject.var = subject.var,
                                          time.var = time.var,
                                          group.var = group.var,
                                          strata.var = NULL,
                                          dist.name = dist.name,
                                          base.size = base.size,
                                          theme.choice = theme.choice,
                                          custom.theme = custom.theme,
                                          palette = palette,
                                          pdf = pdf,
                                          file.ann = file.ann,
                                          pdf.wid = pdf.wid,
                                          pdf.hei = pdf.hei)
beta_ordination_results
```

$BC

$Jaccard

```
if (!is.null(strata.var)){
beta_ordination_stratified_results <- generate_beta_ordination_pair(
                                          data.obj = data.obj,
                                          dist.obj = dist.obj,
                                          pc.obj = pc.obj,
                                          subject.var = subject.var,
                                          time.var = time.var,
                                          group.var = group.var,
                                          strata.var = strata.var,
                                          dist.name = dist.name,
                                          base.size = base.size,
                                          theme.choice = theme.choice,
                                          custom.theme = custom.theme,
                                          palette = palette,
                                          pdf = pdf,
                                          file.ann = file.ann,
                                          pdf.wid = pdf.wid,
                                          pdf.hei = pdf.hei)
beta_ordination_stratified_results
}
```
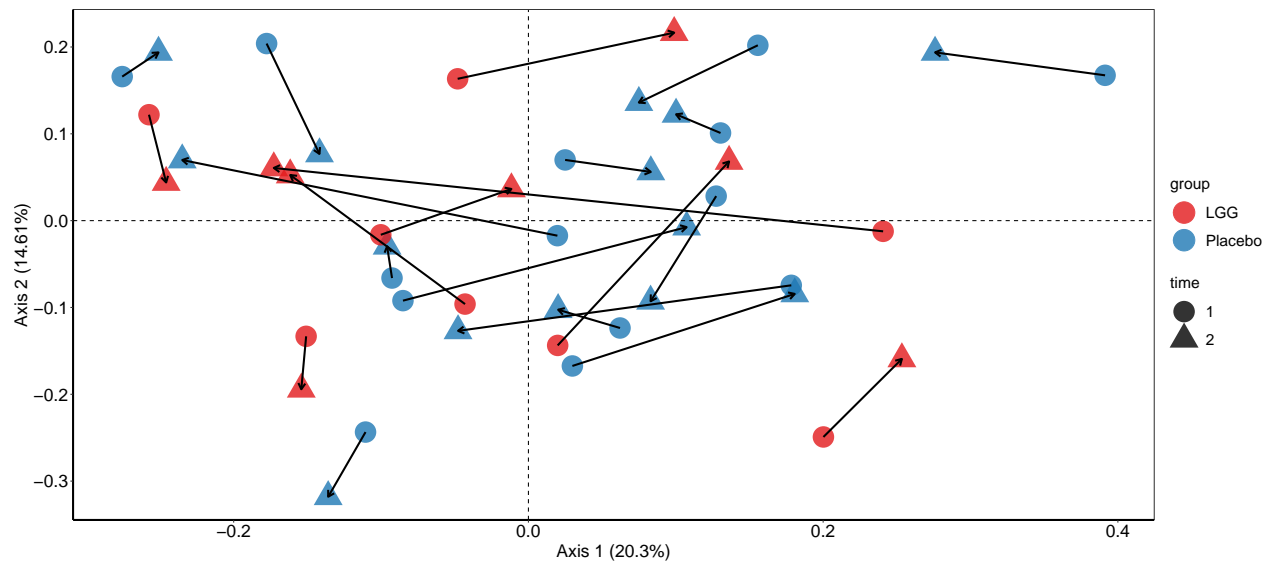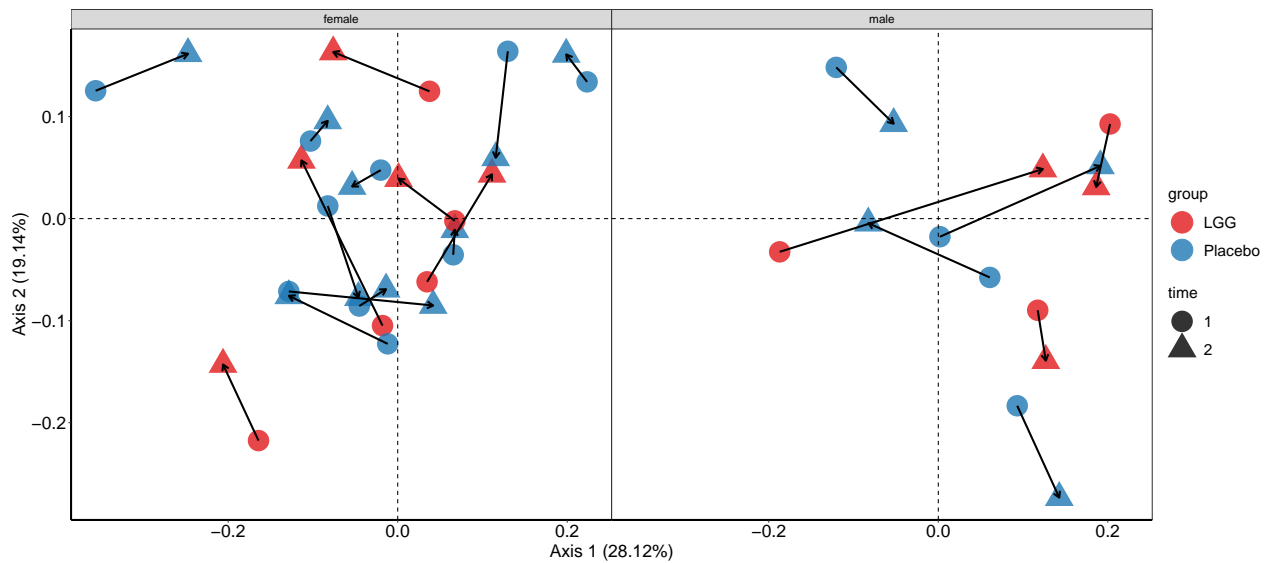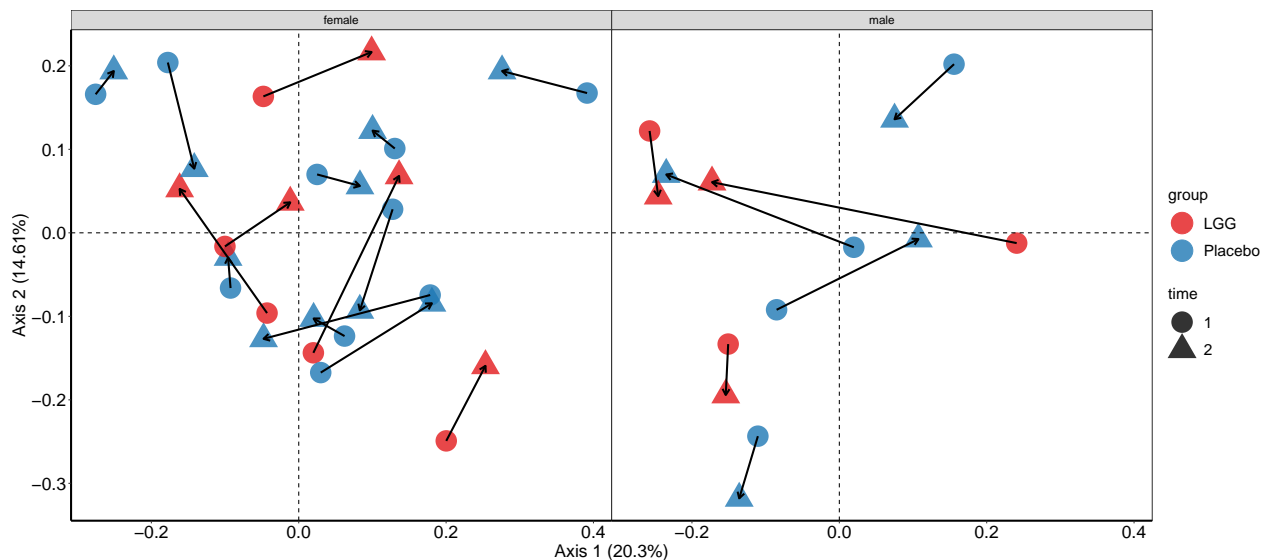
$BC

$Jaccard

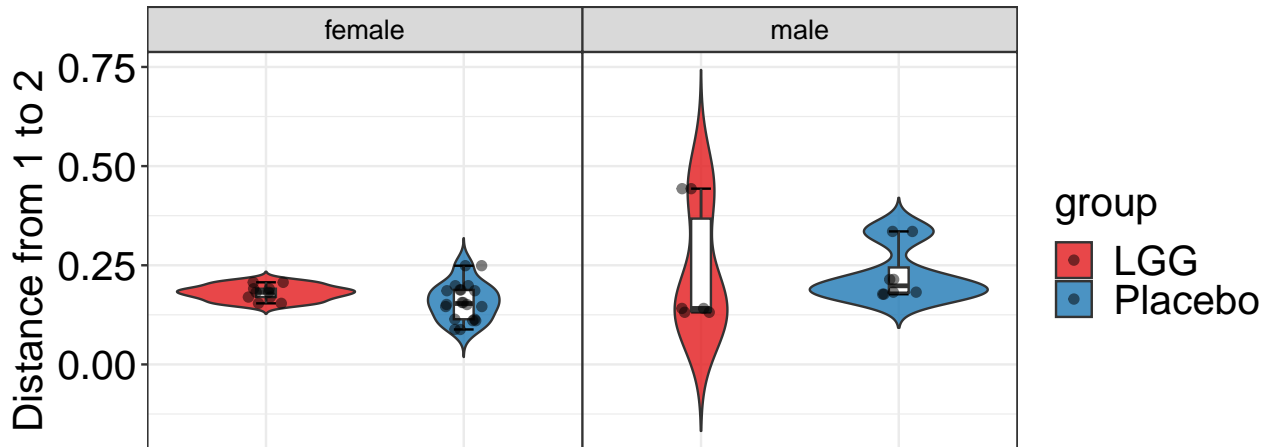### 3.1.2 Beta diversity change boxplot

```
beta_change_boxplot_results <- generate_beta_change_boxplot_pair(
                                    data.obj = data.obj,
                                    dist.obj = dist.obj,
                                    subject.var = subject.var,
                                    time.var = time.var,
                                    group.var = group.var,
                                    strata.var = strata.var,
                                    change.base = change.base,
                                    dist.name = dist.name,
                                    base.size = base.size,
                                    theme.choice = theme.choice,
                                    custom.theme = custom.theme,
                                    palette = palette,
                                    pdf = pdf,
```

```
                                        file.ann = file.ann,
                                        pdf.wid = pdf.wid,
                                        pdf.hei = pdf.hei)
```
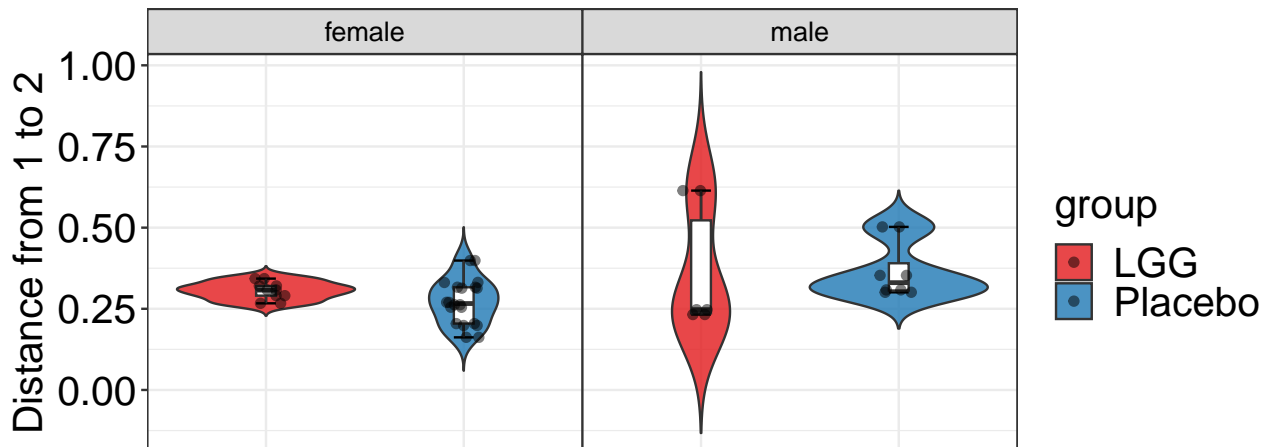
Beta change represents the distance of each subject from their change.base.

$BC

$Jaccard

3.1.3 Beta diversity PC boxplot

```
pc_boxplot_longitudinal_results <- generate_beta_pc_boxplot_long(
  data.obj = data.obj,
  dist.obj = dist.obj,
  pc.obj = pc.obj,
  pc.ind = c(1, 2),
  subject.var = subject.var,
  time.var = time.var,
  t0.level = change.base,
  ts.levels = NULL,
  group.var = group.var,
  strata.var = strata.var,
```
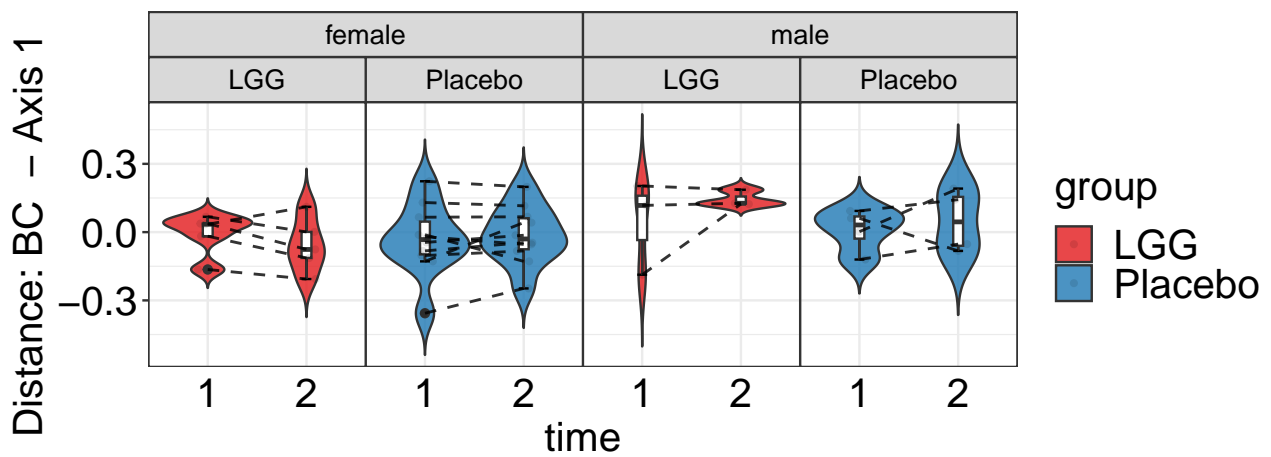
```
    dist.name = dist.name,
    base.size = base.size,
    theme.choice = theme.choice,
    custom.theme = custom.theme,
    palette = palette,
    pdf = pdf,
    file.ann = file.ann,
    pdf.wid = pdf.wid,
    pdf.hei = pdf.hei
)

pc_boxplot_longitudinal_results
```
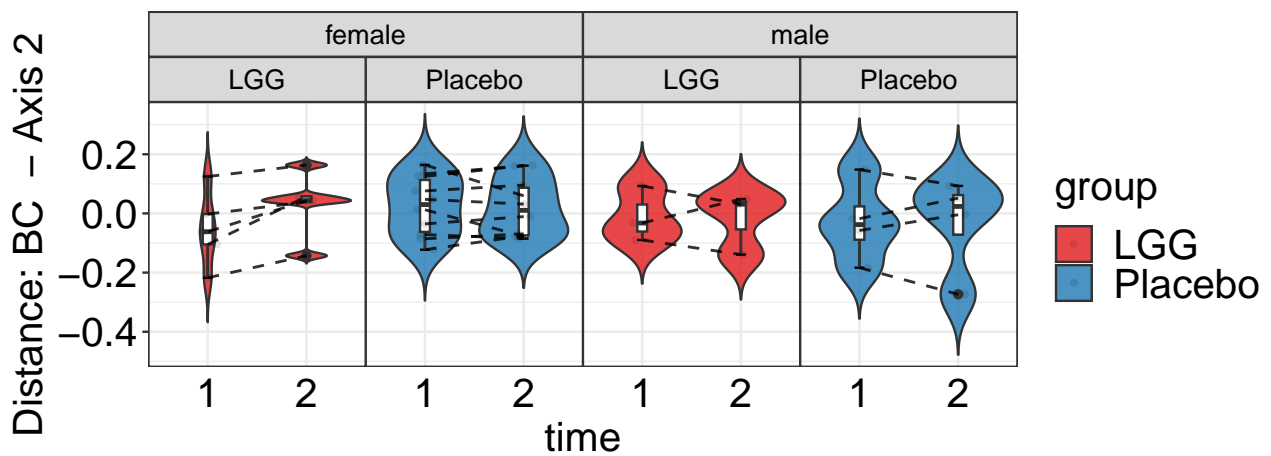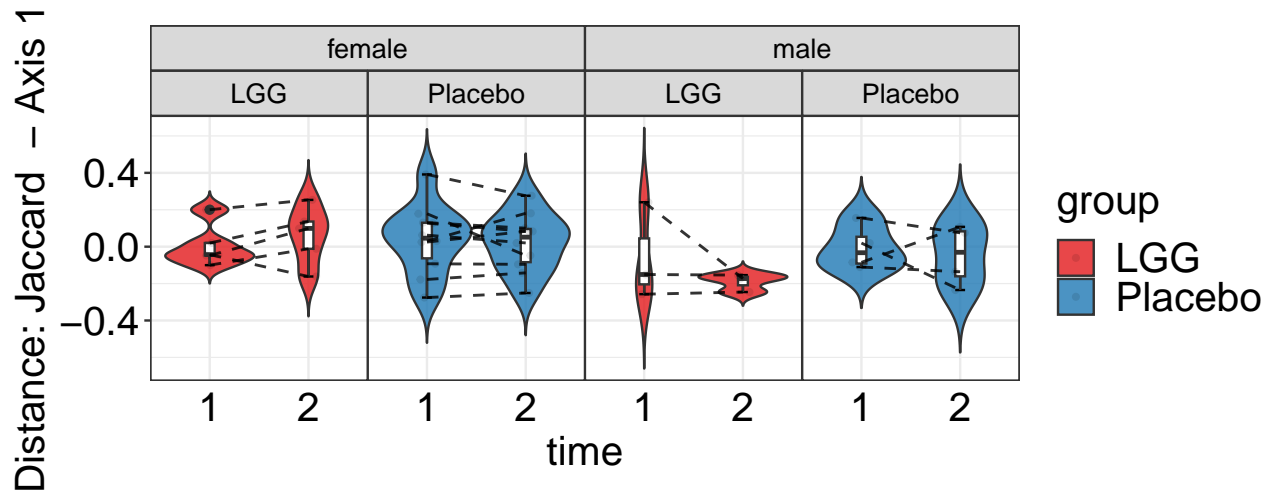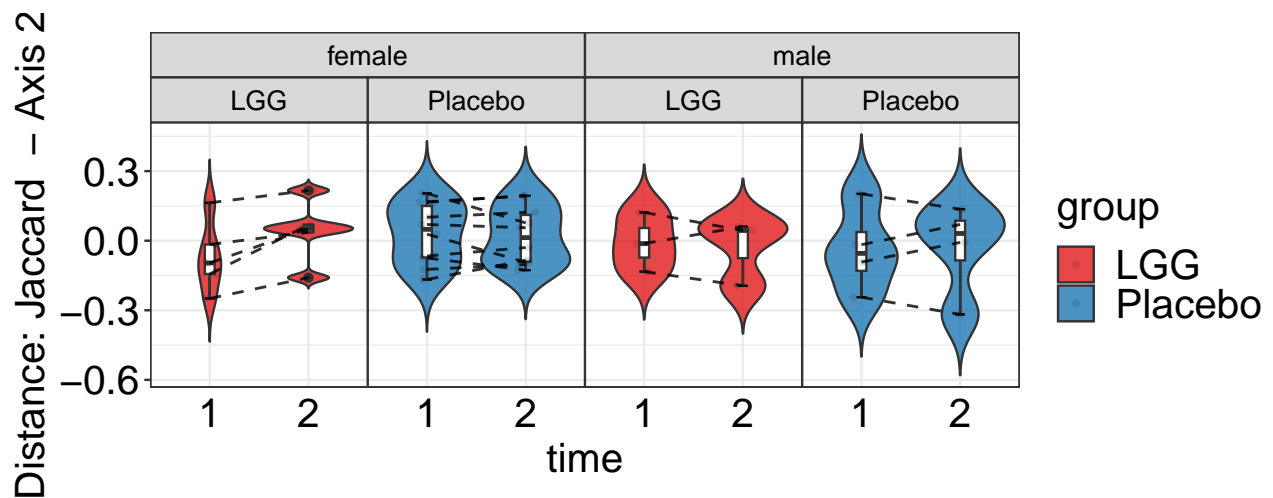
$BC BCPC1

BCPC2

$Jaccard $Jaccard$PC1

*Jaccard*PC2

3.1.4 Beta diversity PC change boxplot

```r
pc_change_boxplot_pairs <- generate_beta_pc_change_boxplot_pair(
  data.obj = data.obj,
  dist.obj = dist.obj,
  pc.obj = pc.obj,
  pc.ind = c(1, 2),
  subject.var = subject.var,
  time.var = time.var,
  group.var = group.var,
  strata.var = strata.var,
  change.base = change.base,
  change.func = 'absolute change',
  dist.name = dist.name,
  base.size = base.size,
  theme.choice = theme.choice,
  custom.theme = custom.theme,
  palette = palette,
  pdf = pdf,
```
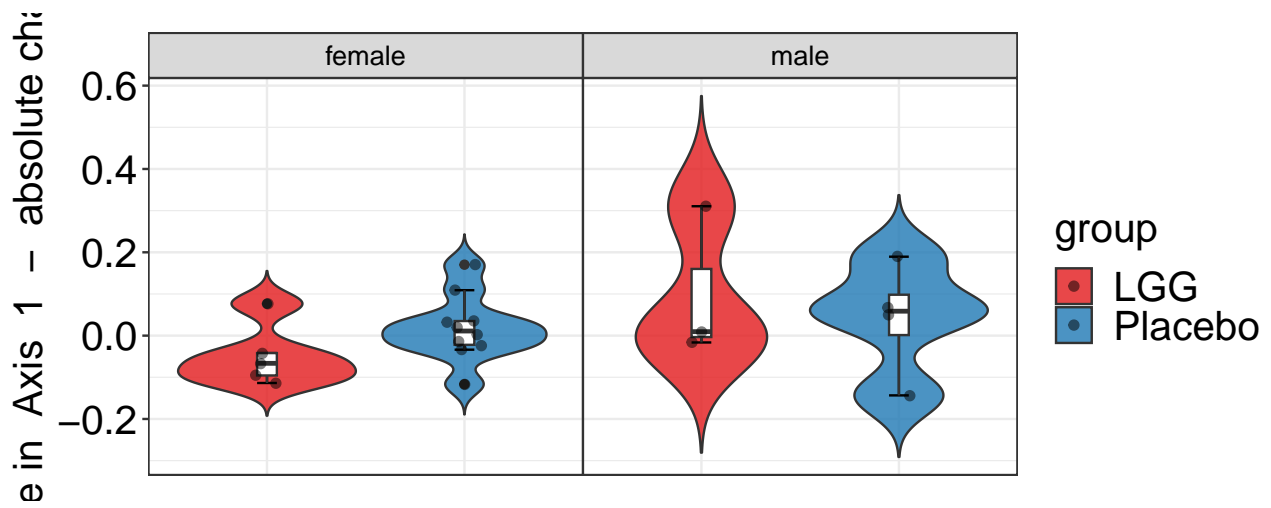
```
    file.ann = file.ann,
    pdf.wid = pdf.wid,
    pdf.hei = pdf.hei
)

pc_change_boxplot_pairs
```
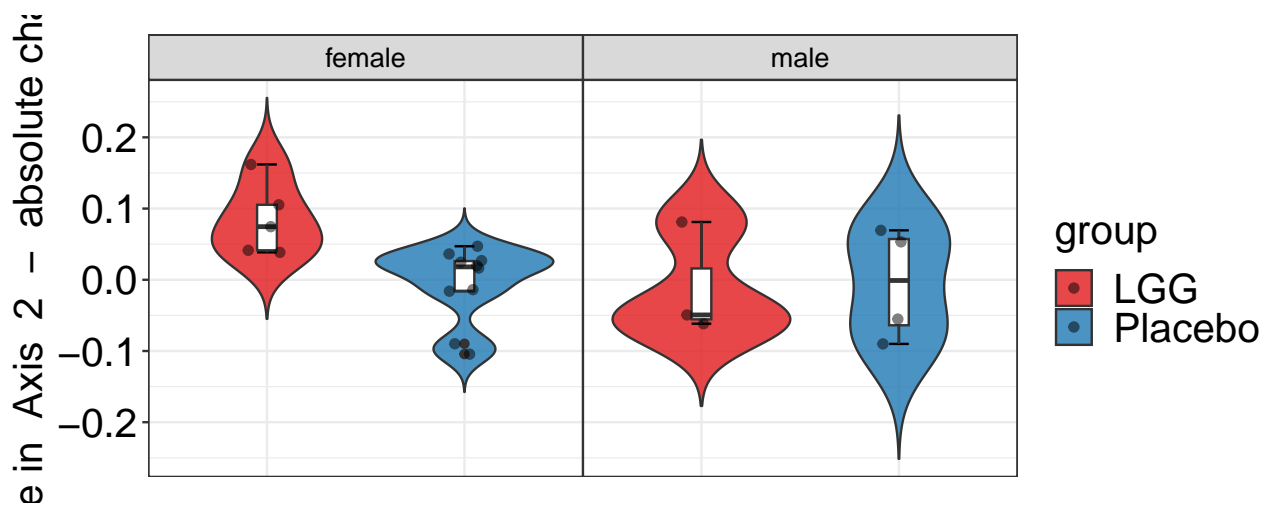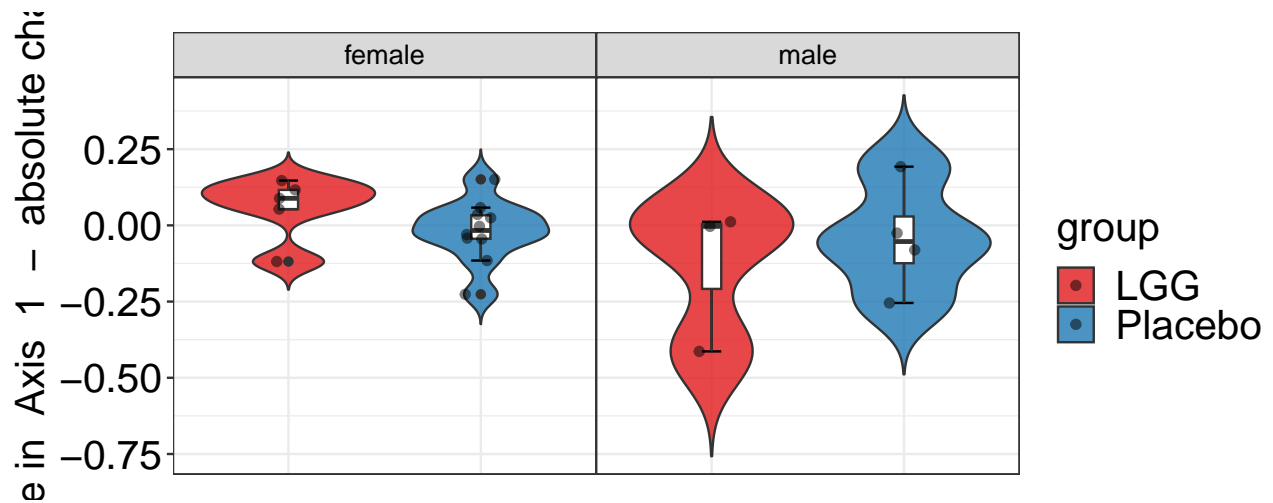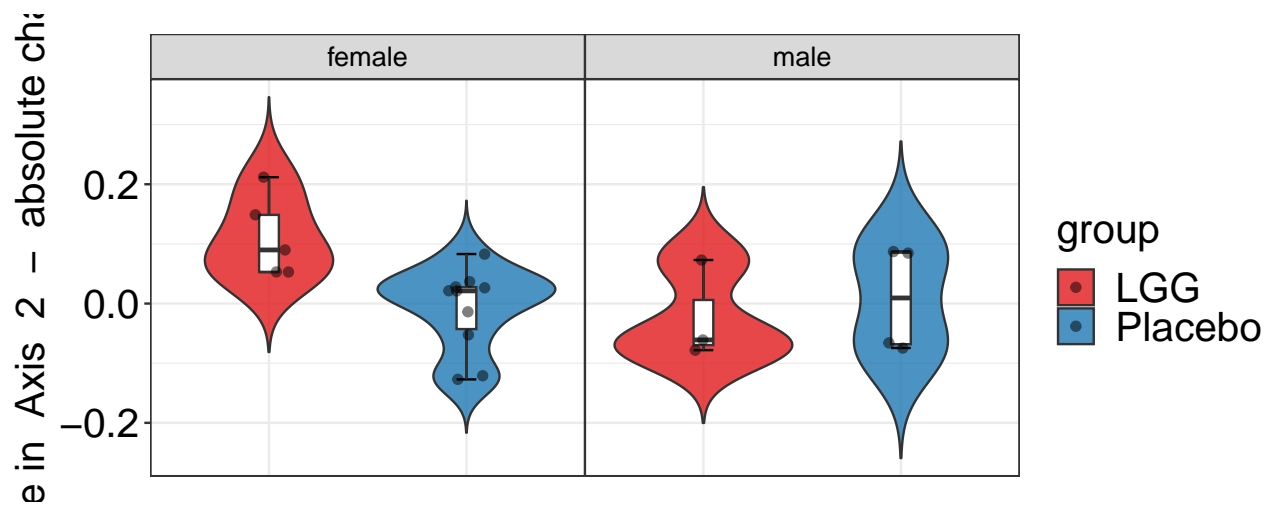
$BC $ BCPC1

BCPC2

$Jaccard $ $Jaccard$PC1

*Jaccard*PC2

3.2 Beta diversity association test based on changes

```
beta_change_test_results <- generate_beta_change_test_pair(
                                data.obj = data.obj,
                                dist.obj = dist.obj,
                                subject.var = subject.var,
                                time.var = time.var,
                                group.var = group.var,
                                adj.vars = test.adj.vars,
                                dist.name = dist.name,
                                change.base = change.base)
```

In this analysis, we utilized a general linear model to examine the influence of the variable group on beta diversity change. Beta change represents the distance of each subject from their change.base.

3.3.1 Bray-Curtis distance

Based on the general linear model, the level Placebo of the variable group did not significantly differ from level LGG, with a p-value of 0.316.

Term	Estimate	Std.Error	Statistic	P.Value
(Intercept)	0.2024	0.01901	10.65	1.67e-13
groupPlacebo	-0.02416	0.02383	-1.014	0.3164

3.3.2 Jaccard distance

Based on the general linear model, the level Placebo of the variable group did not significantly differ from level LGG, with a p-value of 0.342.

Term	Estimate	Std.Error	Statistic	P.Value
(Intercept)	0.3278	0.02446	13.4	9.188e-17
groupPlacebo	-0.02949	0.03066	-0.9617	0.3417

4. Feature-level Analysis

Rarefaction has been enabled for feature-level analysis and visualization.

Reason: The observed abundance of rare/low-abundance features can be strongly influenced by the sequencing depth. Rarefaction is an effective method to control the effect of sequencing depth variation. By employing rarefaction, presence/absence status of the features are more comparable and we can potentially increase the power of detecting those rare/low-abundance features, even though it introduces some variation due to under-sampling. In essence, this step improves comparability across samples across samples with varying sequencing depth.

If you do not wish to perform rarefaction during feature-level analysis, please turn feature.analysis.rarefy to FALSE.

4.1 Feature-level association test based on LinDA-LMM

```
taxa_test_results <- generate_taxa_test_pair(data.obj = data.obj,
                                             subject.var = subject.var,
                                             time.var = time.var,
                                             group.var = group.var,
                                             adj.vars = test.adj.vars,
                                             prev.filter = prev.filter,
                                             abund.filter = abund.filter,
                                             feature.level = test.feature.level,
                                             feature.dat.type = feature.dat.type)
```

$Genus *Genus*Placebo vs LGG (Reference) [Main Effect]

Placebo vs LGG (Reference) [Main Effect]

$Genus$Placebo vs LGG (Reference) [Interaction]
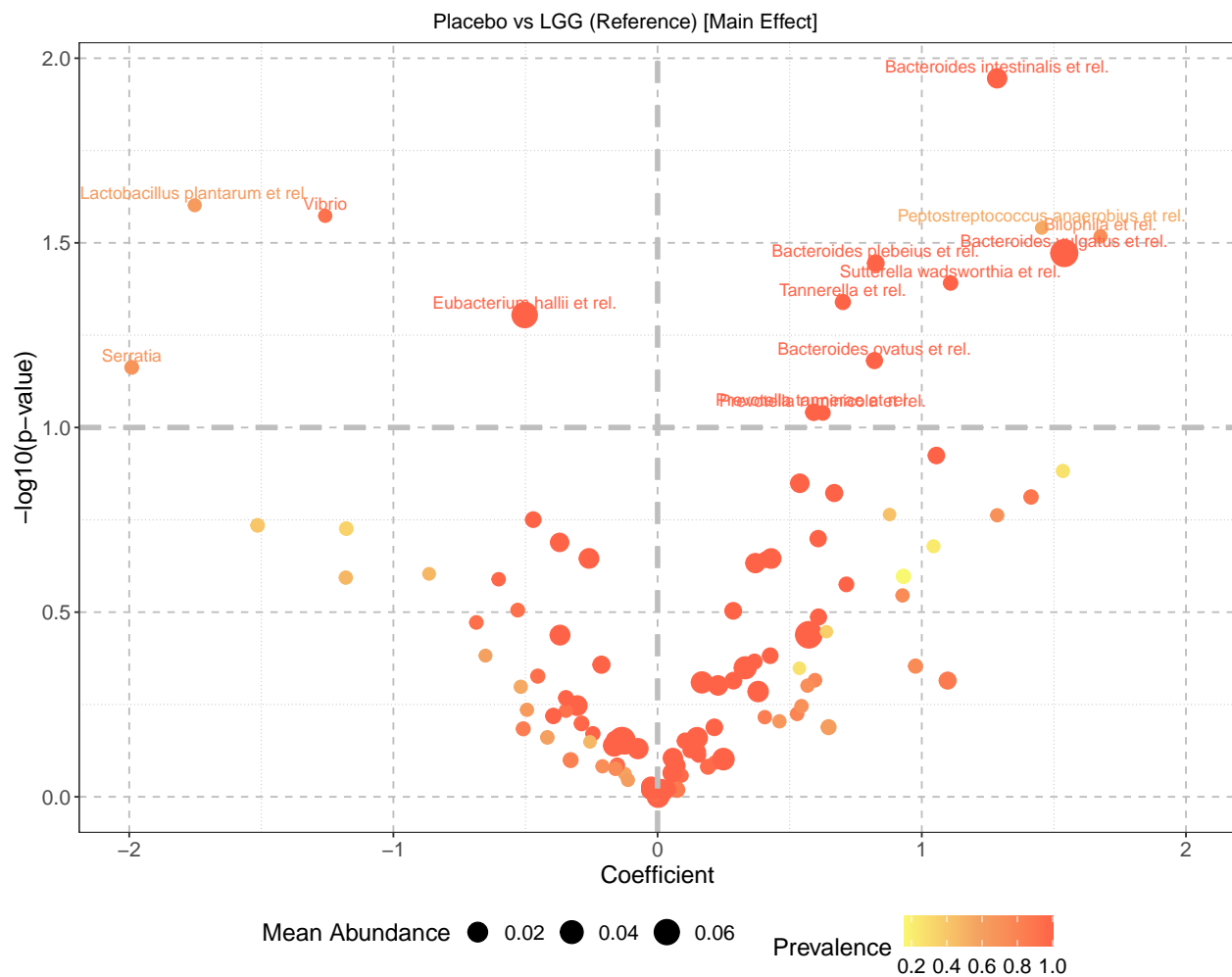
Placebo vs LGG (Reference) [Interaction]

In this analysis, we utilized the LinDA linear model to investigate potential differences in trend. Specifically, we tested the effect of the variable group and the interaction between group and time for different taxa, while adjusting for other covariates.

For the taxon Genus in comparison Placebo vs LGG (Reference) [Main Effect], significant results were identified using the none method for p-value adjustment, based on a threshold of 0.1.

Table 9: Table continues below

Variable	Coefficient	SE	P.Value
Bacteroides intestinalis et rel.	1.285	0.4841	0.01133
Bacteroides ovatus et rel.	0.8208	0.434	0.06588
Bacteroides plebeius et rel.	0.8257	0.3801	0.03591
Bacteroides vulgatus et rel.	1.539	0.6959	0.03371
Bilophila et rel.	1.677	0.7452	0.03036
Eubacterium hallii et rel.	-0.5032	0.2485	0.04959
Lactobacillus plantarum et rel.	-1.752	0.7523	0.02501
Peptostreptococcus anaerobius et rel.	1.454	0.6413	0.02882
Prevotella ruminicola et rel.	0.6255	0.3607	0.0913
Prevotella tannerae et rel.	0.5911	0.3413	0.09095
Serratia	-1.991	1.061	0.06873
Sutterella wadsworthia et rel.	1.109	0.5231	0.04062
Tannerella et rel.	0.7013	0.3401	0.04574

Variable	Coefficient	SE	P.Value
Vibrio	-1.258	0.5469	0.02672

Adjusted.P.Value	Mean.Abundance	Prevalence
0.5653	0.01748	1
0.6529	0.005188	1
0.5653	0.006995	1
0.5653	0.07784	1
0.5653	0.0002901	0.7273
0.5653	0.06158	1
0.5653	0.0002419	0.6364
0.5653	0.0001381	0.5455
0.7434	0.00209	1
0.7434	0.006152	1
0.6529	0.0007579	0.6818
0.5653	0.001568	1
0.5653	0.002786	1
0.5653	0.000365	0.9091

For the taxon Genus in comparison Placebo vs LGG (Reference) [Interaction], significant results were identified using the none method for p-value adjustment, based on a threshold of 0.1.

Table 11: Table continues below

| Variable	Coefficient	SE	P.Value
Anaerostipes caccae et rel.	0.54	0.2842	0.07195
Anaerovorax odorimutans et rel.	1.077	0.343	0.005142
Aneurinibacillus	1.421	0.6495	0.04075
Bacillus	1.085	0.6258	0.09825
Bacteroides intestinalis et rel.	-2.298	0.6846	0.001741
Bacteroides plebeius et rel.	-0.9995	0.5047	0.06159
Bacteroides vulgatus et rel.	-2.204	0.7608	0.008922
Bifidobacterium	1.611	0.702	0.03271
Clostridium difficile et rel.	0.8469	0.4885	0.09838
Clostridium nexile et rel.	0.6215	0.2367	0.0162
Clostridium stercorarium et rel.	0.9303	0.3239	0.009424
Dorea formicigenerans et rel.	0.5027	0.2733	0.08081
Enterobacter aerogenes et rel.	1.388	0.7408	0.06828
Eubacterium cylindroides et rel.	3.265	1.165	0.01099
Eubacterium hallii et rel.	0.7504	0.3512	0.04514
Lactobacillus plantarum et rel.	3.437	1.064	0.002473
Peptostreptococcus anaerobius et rel.	-1.532	0.907	0.09895
Prevotella tannerae et rel.	-0.9759	0.4826	0.0499
Ruminococcus obeum et rel.	0.8641	0.4693	0.08046
Sutterella wadsworthia et rel.	-1.177	0.6425	0.08194
Tannerella et rel.	-0.9098	0.481	0.06581
Uncultured Bacteroidetes	-2.11	0.7837	0.014

Adjusted.P.Value	Mean.Abundance	Prevalence
0.4917	0.01904	1
0.1954	0.003981	1
0.4645	0.0007237	0.9318
0.5127	0.0004625	0.9091
0.1409	0.01748	1
0.4917	0.006995	1
0.2087	0.07784	1
0.4143	0.03325	1
0.5127	0.001831	1
0.2309	0.02822	1
0.2087	0.002191	1
0.4917	0.02203	1
0.4917	0.0005042	0.9773
0.2087	0.0006894	0.4091
0.4679	0.06158	1
0.1409	0.0002419	0.6364
0.5127	0.0001381	0.5455
0.4741	0.006152	1
0.4917	0.07179	1
0.4917	0.001568	1
0.4917	0.002786	1
0.228	0.00625	1

The results for features have been saved in the current working directory. Each taxa rank has its own file named in the format: taxa_test_results_ followed by the taxon rank, the comparison, and the file extension .csv. Please refer to these files for more detailed results.

4.2 Feature-level association test based on changes

```
taxa_change_test_results <- generate_taxa_change_test_pair(data.obj = data.obj,
                                              subject.var = subject.var,
                                              time.var = time.var,
                                              change.base = change.base,
                                              feature.change.func = feature.change.func,
                                              group.var = group.var,
                                              adj.vars = test.adj.vars,
                                              prev.filter = prev.filter,
                                              abund.filter = abund.filter,
                                              feature.level = test.feature.level,
                                              feature.dat.type = feature.dat.type)
```

$Genus *Genus*Placebo vs LGG (Reference)

Placebo vs LGG (Reference)
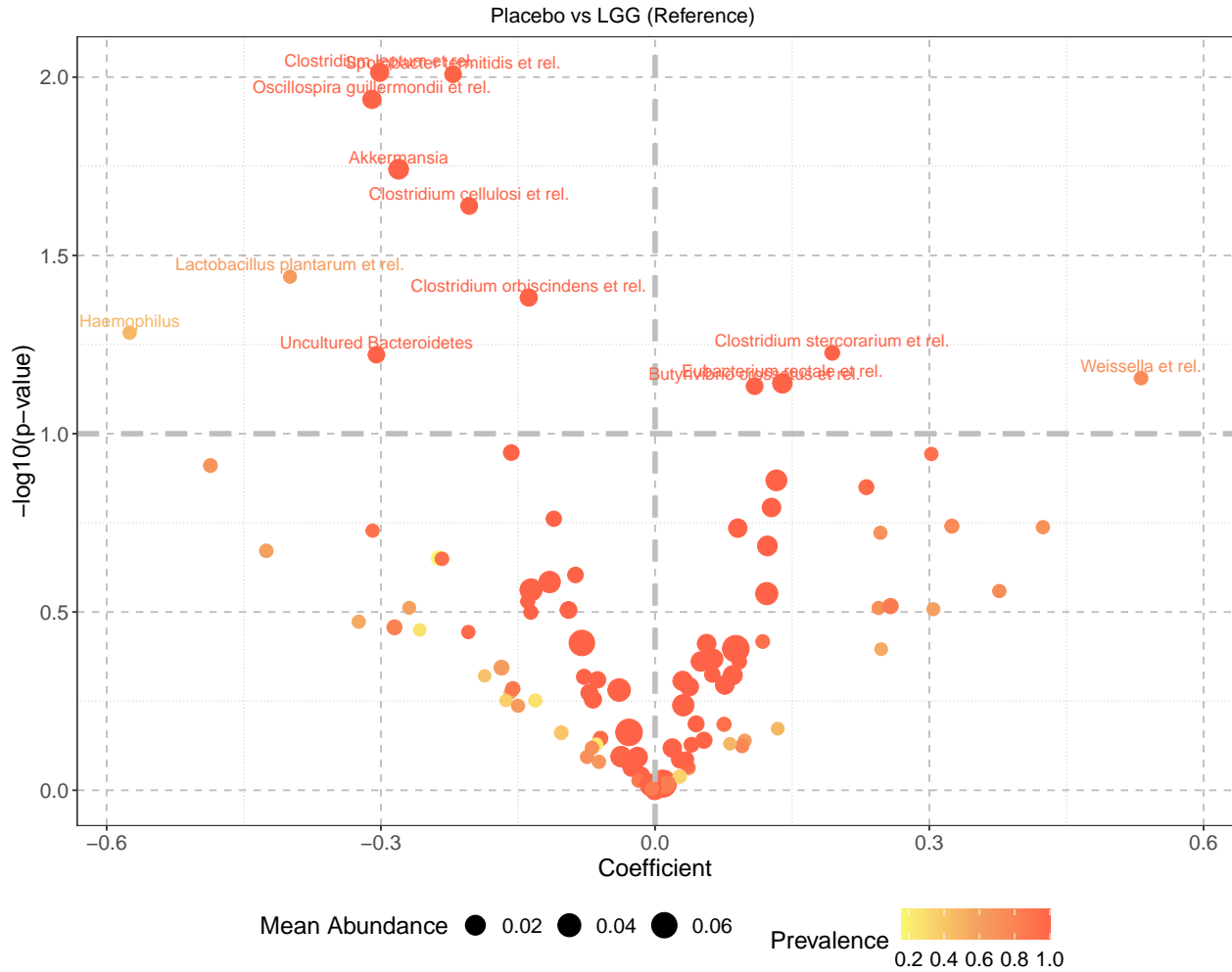
In this analysis, a general linear model was utilized to investigate the influence of the variable group on the change of various taxa abundances.The changes were relative changes, which were computed as (after.abund - before.abund) / (after.abund + before.abund) so the values lie between [-1, 1].For the taxon Genus in comparison Placebo vs LGG (Reference), significant results were identified using the none method for p-value adjustment, based on a threshold of 0.1.

Table 13: Table continues below

Variable	Coefficient	SE	P.Value
Akkermansia	-0.2806	0.109	0.01813
Butyrivibrio crossotus et rel.	0.1089	0.05767	0.07358
Clostridium cellulosi et rel.	-0.2034	0.08258	0.02298
Clostridium leptum et rel.	-0.3013	0.1054	0.009695
Clostridium orbiscindens et rel.	-0.1382	0.06345	0.04153
Clostridium stercorarium et rel.	0.1938	0.0969	0.05932
Eubacterium rectale et rel.	0.1393	0.07338	0.07224
Haemophilus	-0.5749	0.2784	0.05211
Lactobacillus plantarum et rel.	-0.3994	0.178	0.03628
Oscillospira guillermondii et rel.	-0.3098	0.1114	0.01154
Sporobacter termitidis et rel.	-0.221	0.07746	0.009818
Uncultured Bacteroidetes	-0.3049	0.1529	0.06005
Weissella et rel.	0.5316	0.2776	0.06986

Adjusted.P.Value	Mean.Abundance	Prevalence
0.5166	0.01987	1
0.6452	0.006837	1
0.5238	0.007213	1
0.4386	0.01213	1
0.6452	0.008349	1
0.6452	0.002191	1
0.6452	0.01805	1
0.6452	0.0003897	0.4773
0.6452	0.0002419	0.6364
0.4386	0.01277	1
0.4386	0.006457	1
0.6452	0.00625	1
0.6452	0.0005631	0.75

The change test results for individual feature have been saved in the current working directory. Each taxa rank and its corresponding comparison have their own file named with the prefix: taxa_change_test_results_ followed by the taxon rank, the comparison, and the file extension .csv. Please refer to these files for more detailed data.

4.3 Data visualization (significant features)

4.3.1 Significant features boxplot

```r
if (length(significant_vars) != 0){
taxa_indiv_boxplot_results <- generate_taxa_indiv_boxplot_long(
                       data.obj = data.obj,
                       subject.var = subject.var,
                       time.var = time.var,
                       t0.level = change.base,
                       ts.levels = NULL,
                       group.var = group.var,
                       strata.var = strata.var,
                       feature.level = test.feature.level,
                       features.plot = significant_vars,
                       transform = feature.box.axis.transform,
                       feature.dat.type = feature.dat.type,
                       top.k.plot = NULL,
                       top.k.func = NULL,
                       prev.filter = prev.filter,
                       abund.filter = abund.filter,
                       base.size = base.size,
                       theme.choice = theme.choice,
                       custom.theme = custom.theme,
                       palette = palette,
                       pdf = TRUE,
                       file.ann = file.ann,
                       pdf.wid = pdf.wid,
                       pdf.hei = pdf.hei)

taxa_indiv_boxplot_results
}
```

$Genus *Genus*`Anaerostipes caccae et rel.`

Genus`Anaerovorax odorimutans et rel.`

Genus`Aneurinibacillus`

Genus`Bacillus`

Genus`Bacteroides intestinalis et rel.`

Genus`Bacteroides ovatus et rel.`

Genus`Bacteroides plebeius et rel.`

Genus`Bacteroides vulgatus et rel.`

Genus`Bifidobacterium`

Genus`Bilophila et rel.`

Genus`Clostridium difficile et rel.`

Genus`Clostridium nexile et rel.`

Genus`Clostridium stercorarium et rel.`

Genus`Dorea formicigenerans et rel.`

Genus`Enterobacter aerogenes et rel.`

Genus`Eubacterium cylindroides et rel.`

Genus`Eubacterium hallii et rel.`

Genus`Lactobacillus plantarum et rel.`

Genus`Peptostreptococcus anaerobius et rel.`

Genus`Prevotella ruminicola et rel.`

Genus`Prevotella tannerae et rel.`

Genus`Ruminococcus obeum et rel.`

Genus`Serratia`

Genus`Sutterella wadsworthia et rel.`

Genus`Tannerella et rel.`

Genus`Uncultured Bacteroidetes`

Genus`Vibrio`

The boxplot results for individual features can be found in the current working directory. The relevant file is named: taxa_indiv_boxplot_long_subject_subject_time_time_feature_level_Genus_transform_identity_prev_filter_0.1_abund_filter_1e-04_group_group_strata_sex. Please refer to this file for more detailed visualizations.

4.3.2 Significant features boxplot (change)

```
if (length(significant_vars_change) != 0){
taxa_indiv_change_boxplot_results <- generate_taxa_indiv_change_boxplot_pair(
                                      data.obj = data.obj,
                                      subject.var = subject.var,
                                      time.var = time.var,
                                      group.var = group.var,
```

```
                                    strata.var = strata.var,
                                    change.base = change.base,
                                    feature.change.func = feature.change.func,
                                    feature.level = test.feature.level,
                                    features.plot = significant_vars_change,
                                    feature.dat.type = feature.dat.type,
                                    top.k.plot = NULL,
                                    top.k.func = NULL,
                                    prev.filter = prev.filter,
                                    abund.filter = abund.filter,
                                    base.size = base.size,
                                    theme.choice = theme.choice,
                                    custom.theme = custom.theme,
                                    palette = palette,
                                    pdf = pdf,
                                    file.ann = file.ann,
                                    pdf.wid = pdf.wid,
                                    pdf.hei = pdf.hei)
taxa_indiv_change_boxplot_results
}
```

$Genus *Genus*Akkermansia

*Genus*Butyrivibrio crossotus et rel.

*Genus*Clostridium cellulosi et rel.

*Genus*Clostridium leptum et rel.

*Genus*Clostridium orbiscindens et rel.

*Genus*Clostridium stercorarium et rel.

*Genus*Eubacterium rectale et rel.

*Genus*Haemophilus

*Genus*Lactobacillus plantarum et rel.

*Genus*Oscillospira guillermondii et rel.

*Genus*Sporobacter termitidis et rel.

*Genus*Uncultured Bacteroidetes

*Genus*Weissella et rel.

The change boxplot results for individual features can be found in the current working directory. The relevant file is named: taxa_indiv_change_boxplot_pair_subject_subject_time_time_change_base_1_feature_level_Genus_prev_filter_0.1_abund_filter_1e-04_group_group_strata_sex.pdf. Please refer to this file for more detailed visualizations.