

Heterogeneity study

Clémence Métayer metayer.clemence@yahoo.com (main contributor)
Marta Avalos Fernandez marta.avalos-fernandez@u-bordeaux.fr

2022-07-19

Data import

```
library(FactoMineR)
library("factoextra")

data_study <- read.csv("studydata_bin_reduced.csv", sep = ";",
  dec = ".")

data_study$Simpson_control_mean_healthy <-
  ↪ as.numeric(data_study$Simpson_control_mean_healthy)
data_study$Simpson_control_mean_moderate <-
  ↪ as.numeric(data_study$Simpson_control_mean_moderate)
data_study$Simpson_case_exa <- as.numeric(data_study$Simpson_case_exa)
data_study$Simpson_case_diseased <- as.numeric(data_study$Simpson_case_diseased)
data_study$Difference_simpson_healthy_stable <-
  ↪ as.numeric(data_study$Difference_simpson_healthy_stable)
data_study$Difference_simpson_healthy_exa <-
  ↪ as.numeric(data_study$Difference_simpson_healthy_exa)
data_study$Difference_simpson_healthy_diseased <-
  ↪ as.numeric(data_study$Difference_simpson_healthy_diseased)
data_study$Difference_simpson_stable_exa <-
  ↪ as.numeric(data_study$Difference_simpson_stable_exa)
data_study$Chao1_control_mean_healthy <-
  ↪ as.numeric(data_study$Chao1_control_mean_healthy)
data_study$Chao1_control_mean_moderate_COPD <-
  ↪ as.numeric(data_study$Chao1_control_mean_moderate_COPD)
data_study$Chao1_control_mean_moderate_other <-
  ↪ as.numeric(data_study$Chao1_control_mean_moderate_other)
data_study$Chao1_case_group_mean_exa <-
  ↪ as.numeric(data_study$Chao1_case_group_mean_exa)
data_study$Difference_Chao1_healthy_stable_COPD <-
  ↪ as.numeric(data_study$Difference_Chao1_healthy_stable_COPD)
data_study$Difference_Chao1_healthy_stable_other <-
  ↪ as.numeric(data_study$Difference_Chao1_healthy_stable_other)
data_study$Difference_Chao1_healthy_exa <-
  ↪ as.numeric(data_study$Difference_Chao1_healthy_exa)
data_study$Difference_Chao1_stable_exa <-
  ↪ as.numeric(data_study$Difference_Chao1_stable_exa)
```

```

data_study$Shannon_control_mean_healthy <-
  ↳ as.numeric(data_study$Shannon_control_mean_healthy)
data_study$Shannon_control_mean_moderate_COPD <-
  ↳ as.numeric(data_study$Shannon_control_mean_moderate_COPD)
data_study$Shannon_control_mean_moderate_other <-
  ↳ as.numeric(data_study$Shannon_control_mean_moderate_other)
data_study$Shannon_exa_group_mean <- as.numeric(data_study$Shannon_exa_group_mean)
data_study$Shannon_case_group_mean <- as.numeric(data_study$Shannon_case_group_mean)
data_study$Difference_Shannon_healthy_stable_COPD <-
  ↳ as.numeric(data_study$Difference_Shannon_healthy_stable_COPD)
data_study$Difference_Shannon_healthy_stable_other <-
  ↳ as.numeric(data_study$Difference_Shannon_healthy_stable_other)
data_study$Difference_Shannon_healthy_exa <-
  ↳ as.numeric(data_study$Difference_Shannon_healthy_exa)
data_study$Difference_Shannon_healthy_diseased <-
  ↳ as.numeric(data_study$Difference_Shannon_healthy_diseased)
data_study$Difference_Shannon_stable_exa <-
  ↳ as.numeric(data_study$Difference_Shannon_stable_exa)
data_study$Somme_expertise <- as.factor(data_study$Somme_expertise)

summary(data_study)

```

```

rownames(data_study) <- c("Byu17", "Car13", "Cob15", "Den16",
  "Ein16", "Erb11", "Fei17", "Fil12", "Gol13", "Hua20", "Kim17",
  "Lee18", "Li17", "Liu20", "Mar13", "Mil15", "Mun16", "Nar17",
  "Par14", "Ple19", "Pra12", "Sor20", "Sve17", "Wan20")

```

```

colnames(data_study)

```

In the **data_study** dataset, not all variables will be used in the FAMD. We only take on : Continent, Sample, Sequencing, Rarefaction, Taxonomic_study_level, ASV_method, Case_group_size, Control_group_size.

FAMD

```

dta_study <- data_study[, c(8, 9, 10, 12, 13, 14, 16, 17)]
dta_study$Case_group_size[19] = 35 # Number of cases in Park et al. 2014
dta_study$Case_group_size <- as.numeric(dta_study$Case_group_size)
dta_study$Control_group_size <- as.numeric(dta_study$Control_group_size)
summary(dta_study)
afdm <- FAMD(dta_study, graph = FALSE)

# Variables
pdf(file = "var.pdf", width = 7, height = 7)
variables <- fviz_famd_var(afdm, repel = TRUE) + ggtitle("Variables")
variables
dev.off()

```

```

# Variables modalities
pdf(file = "var_quali.pdf", width = 8, height = 8)
variables.mod <- fviz_famd_var(afdm, "quali.var", repel = TRUE) +
  ggtitle("Variables' modalities")
variables.mod
dev.off()

# Quantitative variables
pdf(file = "var_quant.pdf", width = 7, height = 7)
quanti.var <- fviz_famd_var(afdm, "quanti.var", repel = TRUE,
  col.var = "black") + ggtitle("Quantitative variables")
quanti.var
dev.off()

# Modalities + Individuals
Y_control <- quanti.var$data[1, 3]
X_control <- quanti.var$data[1, 2]
Y_case <- quanti.var$data[2, 3]
X_case <- quanti.var$data[2, 2]
pdf(file = "ind_var.pdf", width = 12, height = 12)
iv <- fviz_famd_ind(afdm, repel = TRUE) + ggtitle("Individuals, qualitative and
  ↪ quantitative variables' modalities") +
  geom_segment(aes(x = 0, y = 0, xend = X_case, yend = Y_case,
    col = "red"), arrow = arrow(length = unit(0.015, "npc"))) +
  annotate(geom = "text", x = 0.6, y = 0.1, label = "Case group size",
    fontface = "plain", col = "red") + geom_segment(aes(x = 0,
    y = 0, xend = X_control, yend = Y_control, col = "red"),
    arrow = arrow(length = unit(0.015, "npc"))) + annotate(geom = "text",
    x = 0.5, y = -0.25, label = "Control group size", color = "red",
    fontface = "plain")
iv
dev.off()

# Individuals
pdf(file = "ind.pdf", width = 7, height = 7)
individuals <- fviz_famd_ind(afdm, repel = TRUE, invisible = "quali.var") +
  ggtitle("Individuals")
individuals
dev.off()

# Individuals colored by sample
library("ggplot2")
pdf(file = "sample.pdf", width = 7, height = 7)
ind.sample <- fviz_mfa_ind(afdm, habillage = "Sample", addEllipses = FALSE,
  repel = TRUE, mean.point = FALSE, geom = c("point", "text")) +
  ggtitle("Individuals colored by the type of sample") + guides(colour =
    ↪ guide_legend(override.aes = aes(label = "")))
ind.sample
dev.off()

```

```

# Individuals colored by rarefaction
pdf(file = "rarefaction.pdf", width = 7, height = 7)
ind.rarefaction <- fviz_mfa_ind(afdm, habillage = "Rarefaction",
  palette = c("#00AFBB", "#E7B800", "#FC4E07"), addEllipses = TRUE,
  ellipse.type = "confidence", repel = TRUE) + ggtitle("Individuals colored by the
  ↪ rarefaction") +
  guides(colour = guide_legend(override.aes = aes(label = "")))
ind.rarefaction
dev.off()

# Individuals colored by the sample size my.cont.var <-
# dta_study$Total_sample_size ind.size <-
# fviz_mca_ind(afdm, col.ind = my.cont.var, gradient.cols =
# c('#00AFBB', '#E7B800', '#FC4E07'), legend.title = 'Size
# of the sample', repel=TRUE)+ ggtitle('Individuals colored
# by the size of the sample')

# Individuals colored by the control group size
pdf(file = "control.pdf", width = 7, height = 7)
my.cont.var <- dta_study$Control_group_size
size_control <- fviz_mca_ind(afdm, col.ind = my.cont.var, gradient.cols = c("#00AFBB",
  "#E7B800", "#FC4E07"), legend.title = "Size of the control group",
  repel = TRUE) + ggtitle("Individuals colored by the size of the control group")
size_control
dev.off()

# Individuals colored by the case group size
pdf(file = "case.pdf", width = 7, height = 7)
my.cont.var <- dta_study$Case_group_size
size_case <- fviz_mca_ind(afdm, col.ind = my.cont.var, gradient.cols = c("#00AFBB",
  "#E7B800", "#FC4E07"), legend.title = "Size of the case group",
  repel = TRUE) + ggtitle("Individuals colored by the size of the case group")
size_case
dev.off()

# Individuals colored by the sequencing method
pdf(file = "sequencing.pdf", width = 7, height = 7)
ind.sequencing <- fviz_mfa_ind(afdm, habillage = "Sequencing",
  palette = c("#00AFBB", "#E7B800", "#FC4E07", "#BA3B70", "#25B800"),
  addEllipses = TRUE, ellipse.type = "confidence", repel = TRUE) +
  ggtitle("Individuals colored by the sequencing") + guides(colour =
  ↪ guide_legend(override.aes = aes(label = "")))
ind.sequencing
dev.off()

# Individuals colored by the continent
pdf(file = "continent.pdf", width = 7, height = 7)
ind.continent <- fviz_mfa_ind(afdm, habillage = "Continent",

```

```

    palette = c("#00AFBB", "#E7B800", "#FC4E07"), addEllipses = TRUE,
    ellipse.type = "confidence", repel = TRUE) + ggtitle("Individuals colored by the
    ↪ continent") +
    guides(colour = guide_legend(override.aes = aes(label = "")))
ind.continent
dev.off()

# Individuals colored by the taxonomic level
pdf(file = "taxonomy.pdf", width = 7, height = 7)
ind.taxonomic <- fviz_mfa_ind(afdm, habillage = "Taxonomic.study.level",
    palette = c("#00AFBB", "#E7B800", "#FC4E07"), addEllipses = TRUE,
    ellipse.type = "confidence", repel = TRUE) + ggtitle("Individuals colored by the
    ↪ taxonomic level") +
    guides(colour = guide_legend(override.aes = aes(label = "")))
ind.taxonomic
dev.off()

# Individuals colored by the ASV method
pdf(file = "asv.pdf", width = 7, height = 7)
ind.asv <- fviz_mfa_ind(afdm, habillage = "ASV_method", palette = c("#00AFBB",
    "#E7B800", "#FC4E07"), addEllipses = TRUE, ellipse.type = "confidence",
    repel = TRUE) + ggtitle("Individuals colored by the ASV method") +
    guides(colour = guide_legend(override.aes = aes(label = "")))
ind.asv
dev.off()

# Individuals colored by the type of journal
pdf(file = "journal.pdf", width = 7, height = 7)
ind.colored.journal <- fviz_mfa_ind(afdm, col.ind = data_study$Journal,
    palette = c("#00AFBB", "#E7B800", "#FC4E07"), legend.title = "Journal",
    mean.point = FALSE, repel = TRUE) + ggtitle("Individuals colored by journal") +
    guides(colour = guide_legend(override.aes = aes(label = "")))
ind.colored.journal
dev.off()

# Individuals colored by disease
pdf(file = "disease.pdf", width = 7, height = 7)
p <- fviz_mfa_ind(afdm, col.ind = data_study$Disease, mean.point = FALSE,
    legend.title = "Disease", repel = TRUE)
ind.colored.disease <- p + scale_color_discrete(limits = c("Asthma",
    "Asthma, COPD", "COPD", "Cystic fibrosis", "Bronchiectasis")) +
    ggtitle("Individuals colored by disease") + guides(colour =
    ↪ guide_legend(override.aes = aes(label = "")))
ind.colored.disease
dev.off()

# Individuals colored by the sum of expert assessments
pdf(file = "expert.pdf", width = 7, height = 7)
ind.colored.expert <- fviz_mfa_ind(afdm, col.ind = data_study$Somme_expertise,

```

```

    legend.title = "Sum of the expert assessments", mean.point = FALSE,
    repel = TRUE) + ggtitle("Individuals colored by the sum of expert assessments") +
    guides(colour = guide_legend(override.aes = aes(label = "")))
ind.colored.expert
dev.off()

```

Individuals graph colored by mean indexes

The graph of individuals can also be colored by the diversity index of the microbiota.

Chao1

```

require(gridExtra)
library("cowplot")
name <- list(name = c("Byu17", "Car13", "Cob15", "Den16", "Ein16",
  "Erb11", "Fei17", "Fil12", "Gol13", "Hua20", "Kim17", "Lee18",
  "Li17", "Liu20", "Mar13", "Mil15", "Mun16", "Nar17", "Par14",
  "Ple19", "Pra12", "Sor20", "Sve17", "Wan20"))

# Healthy vs Stable
pdf(file = "diff_chao1_healthy_stable.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Chao1_healthy_stable_COPD
name <- name <- list(name = c("Gol13", "Lee18", "Li17", "Par14"))
diff.chao1.healthy.stable.copd <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
  "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
  ↪ Chao1 index",
  repel = TRUE) + ggtitle("Difference - Healthy/Stable") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
  limits = c(-0.56, 1.99))
diff.chao1.healthy.stable.copd
dev.off()

# Healthy vs Stable
pdf(file = "diff_chao1_healthy_stable_other.pdf", width = 6,
  height = 6)
my.cont.var <- data_study$Difference_Chao1_healthy_stable_other
name <- name <- list(name = c("Gol13", "Lee18", "Li17", "Par14"))
diff.chao1.healthy.stable.other <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
  "red"), pointsize = data_study$Disease_Other, legend.title = "Difference of
  ↪ mean Chao1 index",
  repel = TRUE) + ggtitle("Difference - Healthy/Stable") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
  limits = c(-0.56, 1.99))
diff.chao1.healthy.stable.other
dev.off()

```

```

# Healthy vs Diseased
pdf(file = "diff_chao1_healthy_diseased.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Chao1_healthy_diseased
name <- list(name = c("Den16", "Hua20", "Nar17"))
diff.chao1.healthy.diseased <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_Other, legend.title = "Difference of
    ↪ mean Chao1 index",
  repel = TRUE) + ggtitle("Difference - Healthy/Diseased") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-0.56, 1.99))
diff.chao1.healthy.diseased
dev.off()

# Healthy vs Exacerbated
pdf(file = "diff_chao1_healthy_exa.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Chao1_healthy_exa
name <- list(name = c("Li17", "Sor20"))
diff.chao1.healthy.exa <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_Other, legend.title = "Difference of
    ↪ mean Chao1 index",
  repel = TRUE) + ggtitle("Difference - Healthy/Exacerbated") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-0.56, 1.99))
diff.chao1.healthy.exa
dev.off()

# Stable vs Exacerbated
pdf(file = "diff_chao1_stable_exa.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Chao1_stable_exa
name <- list(name = c("Li17", "Mil15"))
diff.chao1.stable.exa <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Chao1 index",
  repel = TRUE) + ggtitle("Difference - Stable/Exacerbated") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-0.56, 1.99))
diff.chao1.stable.exa
dev.off()

```

Shannon

```

require(gridExtra)
library("cowplot")
name <- list(name = c("Byu17", "Car13", "Cob15", "Den16", "Ein16",
  "Erb11", "Fei17", "Fil12", "Gol13", "Hua20", "Kim17", "Lee18",

```



```

    "Li17", "Liu20", "Mar13", "Mil15", "Mun16", "Nar17", "Par14",
    "Ple19", "Pra12", "Sor20", "Sve17", "Wan20"))

# Healthy vs Stable
pdf(file = "diff_shannon_healthy_stable.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Shannon_healthy_stable_COPD
name <- name <- list(name = c("Ein16", "Erb11", "Gol13", "Lee18",
    "Li17", "Liu20", "Mar13", "Par14", "Pra12", "Wan20", "Wil09"))
diff.shannon.healthy.stable.copd <- fviz_mca_ind(afdm, select.ind = name,
    col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Shannon index",
    repel = TRUE) + ggtitle("Difference - Healthy/Stable") +
    scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.948, 1.66))
diff.shannon.healthy.stable.copd
dev.off()

# Healthy vs Stable
pdf(file = "diff_shannon_healthy_stable_other.pdf", width = 6,
    height = 6)
my.cont.var <- data_study$Difference_Shannon_healthy_stable_other
name <- name <- list(name = c("Ein16", "Erb11", "Gol13", "Lee18",
    "Li17", "Liu20", "Mar13", "Par14", "Pra12", "Wan20", "Wil09"))
diff.shannon.healthy.stable.other <- fviz_mca_ind(afdm, select.ind = name,
    col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_Other, legend.title = "Difference of
    ↪ mean Shannon index",
    repel = TRUE) + ggtitle("Difference - Healthy/Stable") +
    scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.948, 1.66))
diff.shannon.healthy.stable.other
dev.off()

# Healthy vs Exacerbated
pdf(file = "diff_shannon_healthy_exa.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Shannon_healthy_exa
name <- list(name = c("Li17", "Sor20"))
diff.shannon.healthy.exa <- fviz_mca_ind(afdm, select.ind = name,
    col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_Other, legend.title = "Difference of
    ↪ mean Shannon index",
    repel = TRUE) + ggtitle("Difference - Healthy/Exacerbated") +
    scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.948, 1.66))
diff.shannon.healthy.exa
dev.off()

# Healthy vs Diseased

```



```
pdf(file = "diff_shannon_healthy_diseased.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Shannon_heathy_diseased
name <- list(name = c("Den16", "Fei17", "Hua20", "Kim17", "Mun16",
  "Nar17", "Ple19", "Sve17"))
diff.shannon.healthy.diseased <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Shannon index",
  repel = TRUE) + ggtitle("Difference - Healthy/Diseased") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.948, 1.66))
diff.shannon.healthy.diseased
dev.off()

# Stable vs Exacerbated
pdf(file = "diff_shannon_stable_exa.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_Shannon_stable_exa
name <- list(name = c("Byu17", "Car13", "Cob15", "Li17", "Mil15"))
diff.shannon.stable.exa <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Shannon index",
  repel = TRUE) + ggtitle("Difference - Stable/Exacerbated") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.948, 1.66))
diff.shannon.stable.exa
dev.off()
```

Simpson

```
require(gridExtra)
library("cowplot")
name <- list(name = c("Byu17", "Car13", "Cob15", "Den16", "Ein16",
  "Erb11", "Fei17", "Fil12", "Gol13", "Hua20", "Kim17", "Lee18",
  "Li17", "Liu20", "Mar13", "Mil15", "Mun16", "Nar17", "Par14",
  "Ple19", "Pra12", "Sor20", "Sve17", "Wan20"))

# Healthy vs Stable
pdf(file = "diff_simpson_healthy_stable.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_simpson_healthy_stable
name <- name <- list(name = c("Li17", "Pra12"))
diff.simpson.healthy.stable <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Simpson index",
  repel = TRUE) + ggtitle("Difference - Healthy/Stable") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.271, 0.674))
```

```

diff.simpson.healthy.stable
dev.off()

# Healthy vs Exacerbated
pdf(file = "diff_simpson_healthy_exa.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_simpson_healthy_exa
name <- list(name = c("Li17", "Sor20"))
diff.simpson.healthy.exa <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_Other, legend.title = "Difference of
    ↪ mean Simpson index",
  repel = TRUE) + ggtitle("Difference - Healthy/Exacerbated") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.271, 0.674))
diff.simpson.healthy.exa
dev.off()

# Healthy vs Diseased
pdf(file = "diff_simpson_healthy_diseased.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_simpson_healthy_diseased
name <- list(name = c("Den16", "Hua20", "Kim17"))
diff.simpson.healthy.diseased <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Simpson index",
  repel = TRUE) + ggtitle("Difference - Healthy/Diseased") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.271, 0.674))
diff.simpson.healthy.diseased
dev.off()

# Stable vs Exacerbated
pdf(file = "diff_simpson_stable_exa.pdf", width = 6, height = 6)
my.cont.var <- data_study$Difference_simpson_stable_exa
name <- list(name = c("Byu17", "Fil12", "Li17"))
diff.simpson.stable.exa <- fviz_mca_ind(afdm, select.ind = name,
  col.ind = my.cont.var, gradient.cols = c("blue", "yellow",
    "red"), pointsize = data_study$Disease_COPD, legend.title = "Difference of mean
    ↪ Simpson index",
  repel = TRUE) + ggtitle("Difference - Stable/Exacerbated") +
  scale_color_gradient2(low = "blue", mid = "yellow", high = "red",
    limits = c(-1.271, 0.674))
diff.simpson.stable.exa
dev.off()

```