

ANM_Lipid_Differences

Asger Wretlind

2024-04-23

```
#Load libraries  
library(tidyverse)
```

```
## Warning: pakke 'ggplot2' blev bygget under R version 4.3.1
```

```
## Warning: pakke 'purrr' blev bygget under R version 4.3.1
```

```
## Warning: pakke 'dplyr' blev bygget under R version 4.3.1
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
```

```
## v dplyr      1.1.3      v readr      2.1.4
```

```
## v forcats   1.0.0      v stringr   1.5.0
```

```
## v ggplot2   3.4.3      v tibble    3.2.1
```

```
## v lubridate 1.9.2      v tidyr     1.3.0
```

```
## v purrr     1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(here)
```

```
## here() starts at H:/Desktop/ANM_PRS/ANM_Data_Analysis
```

```
library(vroom)
```

```
## Warning: pakke 'vroom' blev bygget under R version 4.3.1
```

```
##
```

```
## Vedhæfter pakke: 'vroom'
```

```
##
```

```
## De følgende objekter er maskerede fra 'package:readr':
```

```
##
```

```
## as.col_spec, col_character, col_date, col_datetime, col_double,
```

```
## col_factor, col_guess, col_integer, col_logical, col_number,
```

```
## col_skip, col_time, cols, cols_condense, cols_only, date_names,
```

```
## date_names_lang, date_names_langs, default_locale, fwf_cols,
```

```
## fwf_empty, fwf_positions, fwf_widths, locale, output_column,
```

```
## problems, spec
```

```

library(tableone)

#Set color palette
color_palette <- c("#11A1B7", "#FF660C", "#OCA61E", "#FE3C1A",
                  "#9966CC", "#4DDF2C", "#FE5387", "#85D0AB",
                  "#18548A", "#FCBB0B", "#FD908F", "#DF56BD", "#F0E4AD")

color_palette2 <- c("Cer" = "#935116", "DG" = "#b7950b", "LacCer" = "#ec407a",
                  "LPC" = "#48c9b0", "PA" = "#ec7063", "PC" = "#2874a6",
                  "PE" = "#7d3c98", "PI" = "#f8c471", "SM" = "#1e8449",
                  "TG" = "#C71B42", "Not significant" = "#808080")

#Load data
data <- vroom(here("data/ANM_Lipid_Preprocessed_v4.csv"))

## Rows: 841 Columns: 293
## -- Column specification -----
## Delimiter: "\t"
## chr (8): ID, Site, Date, Status, Sex, DOB, Accommodation, Marital_Status
## dbl (285): Visit, Order, Label, Age, Fulltime_Education_Years, apoe, e4_p, e...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

#Load list of selected lipids
selected_lipids <- vroom(here("data/ANM_module_selected_lipids_res_v1.1.csv"))

## Rows: 47 Columns: 3
## -- Column specification -----
## Delimiter: "\t"
## chr (3): Module, Lipid, Color
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

#Characteristics of of interest
tmp_vars <- c("Age", "Sex", "APOE-e4", "Total_Cholesterol", "LDL", "HDL", "Total_Triglyceride", "ApoB",
             "MMSE_Score")

table1 <- data %>%
  mutate(e4_c = factor(e4_c, levels = c(0, 1, 2),
                      labels = c("Absent", "Heterozygote", "Homozygote"))) %>%
  rename("APOE-e4" = e4_c) %>%
  rename("Total_Cholesterol" = Total_C) %>%
  rename("LDL" = LDL_C) %>%
  rename("HDL" = HDL_C) %>%
  rename("Total_Triglyceride" = Total_TG) %>%
  rename("MMSE_Score" = MMSE_Total) %>%
  CreateTableOne(vars = tmp_vars,
                strata = "Status",
                addOverall = TRUE)

table1

```

```
##
## Stratified by Status
## Overall ADC
## n 841 306
## Age (mean (SD)) 76.24 (6.76) 77.37 (6.63)
## Sex = Male (%) 350 (41.6) 123 (40.2)
## APOE-e4 (%)
## Absent 456 (60.4) 128 (43.2)
## Heterozygote 245 (32.5) 129 (43.6)
## Homozygote 54 ( 7.2) 39 (13.2)
## Total_Cholesterol (mean (SD)) 5.40 (1.22) 5.52 (1.17)
## LDL (mean (SD)) 2.05 (0.58) 2.10 (0.57)
## HDL (mean (SD)) 1.60 (0.38) 1.63 (0.37)
## Total_Triglyceride (mean (SD)) 1.42 (0.64) 1.39 (0.61)
## ApoB (mean (SD)) 0.97 (0.26) 0.99 (0.25)
## MMSE_Score (mean (SD)) 25.46 (4.99) 20.87 (5.07)
## Fulltime_Education_Years (mean (SD)) 9.95 (4.32) 8.86 (4.09)
## Marital_Status (%)
## Divorced 31 ( 4.7) 5 ( 1.9)
## Married 394 (59.2) 144 (55.6)
## Single 41 ( 6.2) 12 ( 4.6)
## Widowed 200 (30.0) 98 (37.8)
##
## Stratified by Status
## CTL MCI p test
## n 370 165
## Age (mean (SD)) 75.43 (6.79) 75.95 (6.68) 0.001
## Sex = Male (%) 150 (40.5) 77 (46.7) 0.339
## APOE-e4 (%) <0.001
## Absent 236 (73.8) 92 (66.2)
## Heterozygote 75 (23.4) 41 (29.5)
## Homozygote 9 ( 2.8) 6 ( 4.3)
## Total_Cholesterol (mean (SD)) 5.24 (1.20) 5.52 (1.31) 0.005
## LDL (mean (SD)) 1.97 (0.57) 2.11 (0.60) 0.004
## HDL (mean (SD)) 1.59 (0.39) 1.59 (0.40) 0.369
## Total_Triglyceride (mean (SD)) 1.41 (0.63) 1.48 (0.72) 0.336
## ApoB (mean (SD)) 0.93 (0.25) 1.00 (0.28) 0.003
## MMSE_Score (mean (SD)) 28.47 (2.67) 26.80 (2.06) <0.001
## Fulltime_Education_Years (mean (SD)) 11.19 (4.34) 9.49 (4.05) <0.001
## Marital_Status (%) <0.001
## Divorced 20 ( 7.7) 6 ( 4.1)
## Married 157 (60.2) 93 (63.7)
## Single 23 ( 8.8) 6 ( 4.1)
## Widowed 61 (23.4) 41 (28.1)
```

```
#write.csv(print(table1, printToggle = FALSE), here("data/table1_v1.1.csv"))
```

```
#Average AD disease duration
print(paste("Average AD disease duration",
  data %>%
  filter(Status == "ADC") %>%
  pull(Disease_Duration) %>%
  mean(., na.rm = TRUE) %>%
  round(., digits = 2),
  "years"))
```

```
## [1] "Average AD disease duration 3.74 years"
```

```
#Average age at AD onset
print(paste("Average age at AD onset",
  data %>%
  filter(Status == "ADC") %>%
  pull(Onset_Age) %>%
  mean(., na.rm = TRUE) %>%
  round(., digits = 2),
  "years"))
```

```
## [1] "Average age at AD onset 73.04 years"
```

```
##Participants characteristics stratified by sex
```

```
#Swap sex with AD status
tmp_vars[which(tmp_vars == "Sex")] <- "Status"

table2 <- data %>%
  mutate(e4_c = factor(e4_c, levels = c(0, 1, 2),
    labels = c("Absent", "Heterozygote", "Homozygote"))) %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
    labels = c("Ctrl", "MCI", "ADC"))) %>%
  rename("APOE-e4" = e4_c) %>%
  rename("Total_Cholesterol" = Total_C) %>%
  rename("LDL" = LDL_C) %>%
  rename("HDL" = HDL_C) %>%
  rename("Total_Triglyceride" = Total_TG) %>%
  rename("MMSE_Score" = MMSE_Total) %>%
  CreateTableOne(vars = tmp_vars,
    strata = "Sex",
    addOverall = TRUE)
```

```
table2
```

	Stratified by Sex	
	Overall	Female
n	841	491
Age (mean (SD))	76.24 (6.76)	75.88 (7.19)
Status (%)		
Ctrl	370 (44.0)	220 (44.8)
MCI	165 (19.6)	88 (17.9)
ADC	306 (36.4)	183 (37.3)
APOE-e4 (%)		
Absent	456 (60.4)	252 (57.0)
Heterozygote	245 (32.5)	155 (35.1)
Homozygote	54 (7.2)	35 (7.9)
Total_Cholesterol (mean (SD))	5.40 (1.22)	5.70 (1.20)
LDL (mean (SD))	2.05 (0.58)	2.14 (0.59)
HDL (mean (SD))	1.60 (0.38)	1.71 (0.38)
Total_Triglyceride (mean (SD))	1.42 (0.64)	1.44 (0.65)
ApoB (mean (SD))	0.97 (0.26)	1.01 (0.26)
MMSE_Score (mean (SD))	25.46 (4.99)	25.27 (5.25)

```
## Fulltime_Education_Years (mean (SD)) 9.95 (4.32) 9.83 (4.37)
## Marital_Status (%)
## Divorced 31 ( 4.7) 20 ( 5.1)
## Married 394 (59.2) 167 (42.9)
## Single 41 ( 6.2) 27 ( 6.9)
## Widowed 200 (30.0) 175 (45.0)
##
## Stratified by Sex
## Male p test
## n 350
## Age (mean (SD)) 76.74 (6.07) 0.072
## Status (%) 0.339
## Ctrl 150 (42.9)
## MCI 77 (22.0)
## ADC 123 (35.1)
## APOE-e4 (%) 0.076
## Absent 204 (65.2)
## Heterozygote 90 (28.8)
## Homozygote 19 ( 6.1)
## Total_Cholesterol (mean (SD)) 4.98 (1.11) <0.001
## LDL (mean (SD)) 1.91 (0.54) <0.001
## HDL (mean (SD)) 1.45 (0.34) <0.001
## Total_Triglyceride (mean (SD)) 1.38 (0.63) 0.128
## ApoB (mean (SD)) 0.91 (0.24) <0.001
## MMSE_Score (mean (SD)) 25.73 (4.59) 0.189
## Fulltime_Education_Years (mean (SD)) 10.13 (4.24) 0.334
## Marital_Status (%) <0.001
## Divorced 11 ( 4.0)
## Married 227 (81.9)
## Single 14 ( 5.1)
## Widowed 25 ( 9.0)
```

```
#write.csv(print(table2, printToggle = FALSE), here("data/table2_v1.0.csv"))
```

```
rm(tmp_vars, table1, table2)
```

```
#Function that takes a continuous variable and normalizes so that the maximum value is 1 and the minimum
Mm_normalize <- function(x){
  return((x- min(x)) /(max(x)-min(x)))
}
```

```
#List of lipids
lipids <- colnames(data)[which(colnames(data) == "Cer(d42:0)":length(data))]

#data is Min-max normalized to get consistent direction of fold change
data_norm <- data %>%
  mutate(across(all_of(lipids), ~ Mm_normalize(.)))

#Create subsets of the data stratified by sex and AD status
data_female_AD <- data_norm %>%
  filter(Sex == "Female") %>%
  filter(Status == "ADC")

data_female_ctrl <- data_norm %>%
```

```

filter(Sex == "Female") %>%
filter(Status == "CTL")

data_male_AD <- data_norm %>%
  filter(Sex == "Male") %>%
  filter(Status == "ADC")

data_male_ctrl <- data_norm %>%
  filter(Sex == "Male") %>%
  filter(Status == "CTL")

```

```

#Function that takes in two data sets and a list of lipids and outputs a dataframe of the mean differen
Ttest_extract <- function(data_1, data_2, list_of_lipids){
  output_dataframe <- data.frame("mean_diff" = c(),
                                "fold_change" = c(),
                                "p_value" = c())

  for(i in list_of_lipids){
    tmp_model <- t.test(data_1[i], data_2[i])
    tmp_mean_diff <- tmp_model$estimate[[1]]-tmp_model$estimate[[2]]
    tmp_fold_change <- tmp_model$estimate[[1]]/tmp_model$estimate[[2]]
    tmp_pval <- tmp_model$p.value

    output_dataframe[i, "mean_diff"] <- tmp_mean_diff
    output_dataframe[i, "fold_change"] <- tmp_fold_change
    output_dataframe[i, "p_value"] <- tmp_pval

  }

  output_dataframe$Lipid <- rownames(output_dataframe)
  output_dataframe$FDR <- p.adjust(output_dataframe$p_value, method = "fdr")
  output_dataframe <- output_dataframe[, c("Lipid", "mean_diff", "fold_change",
                                           "p_value", "FDR")]

  return(output_dataframe)
}

```

```

library(ggrepel)

#In women
ttest_female_AD <- Ttest_extract(data_1 = data_female_AD,
                                data_2 = data_female_ctrl,
                                list_of_lipids = lipids)

# Save figure
pdf(here("figures/Volcano_female_AD-ctrl_v1.0.pdf"), width = 6, height = 5)

#Volcano plot

```

```

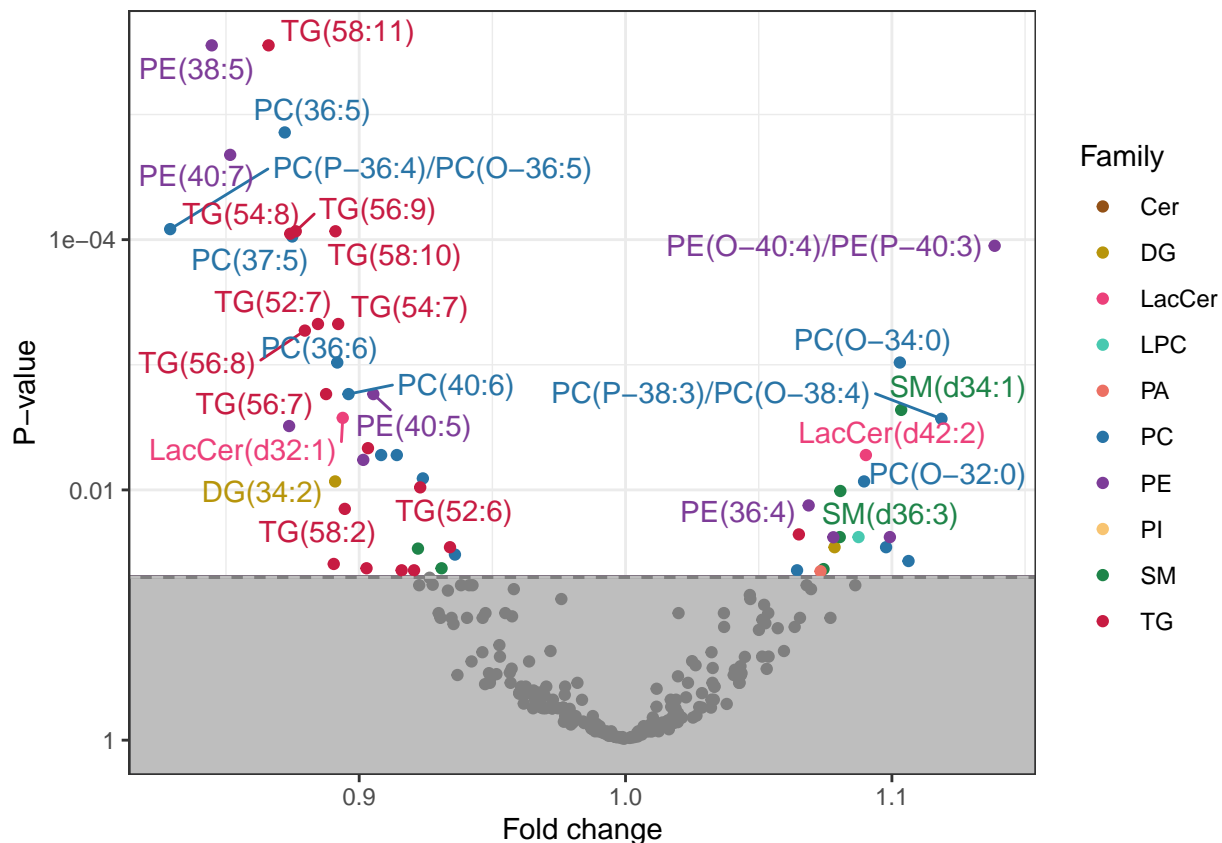
tibble(ttest_female_AD) %>%
  left_join(., selected_lipids) %>%
  mutate(tmp_lipid = Lipid) %>%
  mutate(Lipid = gsub("_A", "", Lipid)) %>%
  mutate(Lipid = gsub("_B", "", Lipid)) %>%
  separate(tmp_lipid, sep = "\\(", c("Family", "Chain1")) %>%
  mutate(Family = ifelse(FDR > 0.05, "Not significant", Family)) %>%
  ggplot(aes(x = fold_change, y = -log10(FDR), color = Family))+
  geom_rect(aes(xmin = -Inf, xmax = Inf,
                ymin = -log10(0.05), ymax = -Inf),
            fill="grey", alpha=0.05, show.legend = FALSE)+
  geom_point(alpha = 1)+
  scale_color_manual(values = color_palette2,
                    limits = names(color_palette2)[-11])+
  geom_hline(yintercept = -log10(0.05), linetype = "dashed", colour = "grey")+
  geom_text_repel(aes(label = Lipid), show.legend = FALSE) +
  scale_y_continuous(limits = c(0,NA), labels = function(i) 10^-i)+
  ylab(label = "P-value")+
  xlab(label = "Fold change")+
  theme_bw()+
  theme()

```

```
## Joining with 'by = join_by(Lipid)'
```

```
## Warning: Expected 2 pieces. Additional pieces discarded in 34 rows [69, 70, 71, 72, 73,
## 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, ...].
```

```
## Warning: ggrepel: 240 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



```
#legend.position = "none")

# dev.off()

print(paste0("Number of lipids with a p-value < 0.05: ", sum(ttest_female_AD$FDR < 0.05)))

## [1] "Number of lipids with a p-value < 0.05: 53"

#In men
ttest_male_AD <- Ttest_extract(data_1 = data_male_AD,
                              data_2 = data_male_ctrl,
                              list_of_lipids = lipids)

# #Save figure
# pdf(here("figures/Volcano_male_AD-ctrl_v1.0.pdf"), width = 5, height = 5)

#Volcano plot
tibble(ttest_male_AD) %>%
  left_join(., selected_lipids) %>%
  mutate(tmp_lipid = Lipid) %>%
  mutate(Lipid = gsub("_A", "", Lipid)) %>%
  mutate(Lipid = gsub("_B", "", Lipid)) %>%
  separate(tmp_lipid, sep = "\\(", c("Family", "Chain1")) %>%
  mutate(Family = ifelse(FDR > 0.05, "Not significant", Family)) %>%
  ggplot(aes(x = fold_change, y = -log10(FDR), color = Family))+
  geom_rect(aes(xmin = -Inf, xmax = Inf,
```



```

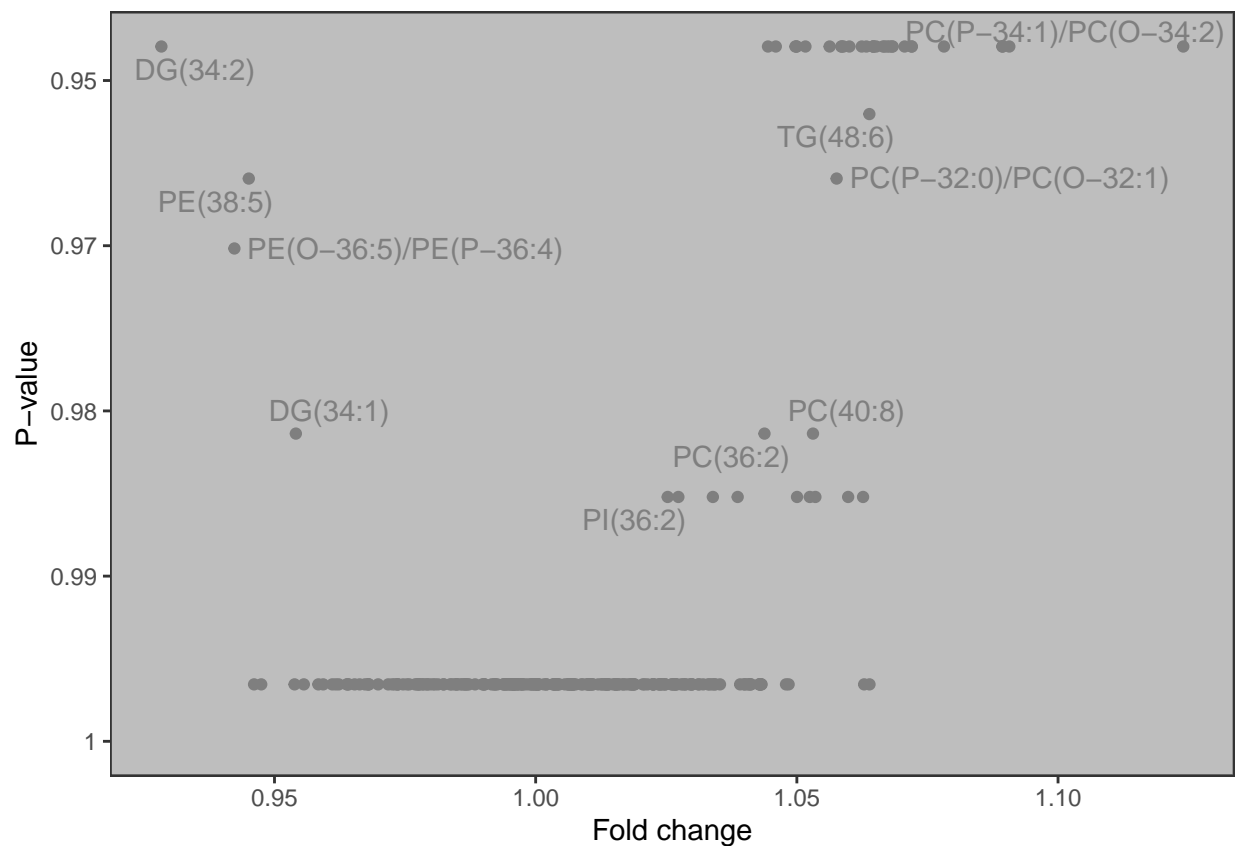
        ymin = Inf, ymax = -Inf),
        fill="grey", alpha=0.05, show.legend = FALSE)+
geom_point(alpha = 1)+
scale_color_manual(values = color_palette2,
                    limits = names(color_palette2)[-11])+
#geom_hline(yintercept = -log10(0.05), linetype = "dashed", colour = "grey")+
geom_text_repel(aes(label = Lipid), show.legend = FALSE)+
scale_y_continuous(limits = c(0,NA), labels = function(i) round(10^-i, 2))+
ylab(label = "P-value")+
xlab(label = "Fold change")+
theme_bw()+
theme(legend.position = "none")

```

```
## Joining with 'by = join_by(Lipid)'
```

```
## Warning: Expected 2 pieces. Additional pieces discarded in 34 rows [69, 70, 71, 72, 73,
## 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, ...].
```

```
## Warning: ggrepel: 258 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



```

# dev.off()

print(paste0("Number of lipids with a p-value < 0.05: ", sum(ttest_male_AD$FDR < 0.05)))

```

```
## [1] "Number of lipids with a p-value < 0.05: 0"
```

```
#Add Sex variable before merging
ttest_female_AD$Sex <- "Female"

#Facet male and female ttest_AD
fig3_A_volcano <- ttest_male_AD %>%
  mutate(Sex = "Male") %>%
  rbind(., ttest_female_AD) %>%
  tibble() %>%
  left_join(., selected_lipids) %>%
  mutate(tmp_lipid = Lipid) %>%
  mutate(Lipid = gsub("_A", "", Lipid)) %>%
  mutate(Lipid = gsub("_B", "", Lipid)) %>%
  separate(tmp_lipid, sep = "\\(", c("Family", "Chain1")) %>%
  mutate(Family = ifelse(FDR > 0.05, "Not significant", Family)) %>%
  ggplot(aes(x = fold_change, y = -log10(FDR), color = Family))+
  geom_rect(aes(xmin = -Inf, xmax = Inf, ymin = -log10(0.05), ymax = -Inf),
    fill="grey", alpha=0.05, show.legend = FALSE)+
  geom_point(alpha = 1)+
  scale_color_manual(values = color_palette2,
    limits = names(color_palette2)[-11])+
  geom_hline(yintercept = -log10(0.05), linetype = "dashed", colour = "grey")+
  geom_text_repel(aes(label = Lipid), size = 3,
    max.overlaps = 8,
    show.legend = FALSE)+
  scale_y_continuous(limits = c(0,NA), labels = function(i) 10^-i)+
  ylab(label = "P-value")+
  xlab(label = "Fold change")+
  labs(color = "Lipid\nFamily")+
  facet_grid(.~Sex)+
  theme_bw()+
  theme(strip.background = element_blank(),
    strip.text = element_text(face = "bold", size = 12),
    legend.text = element_text(face = "bold"),
    axis.title = element_text(face = "bold"),
    legend.title = element_text(face = "bold"),
    legend.position="bottom")
```

```
## Joining with 'by = join_by(Lipid)'
```

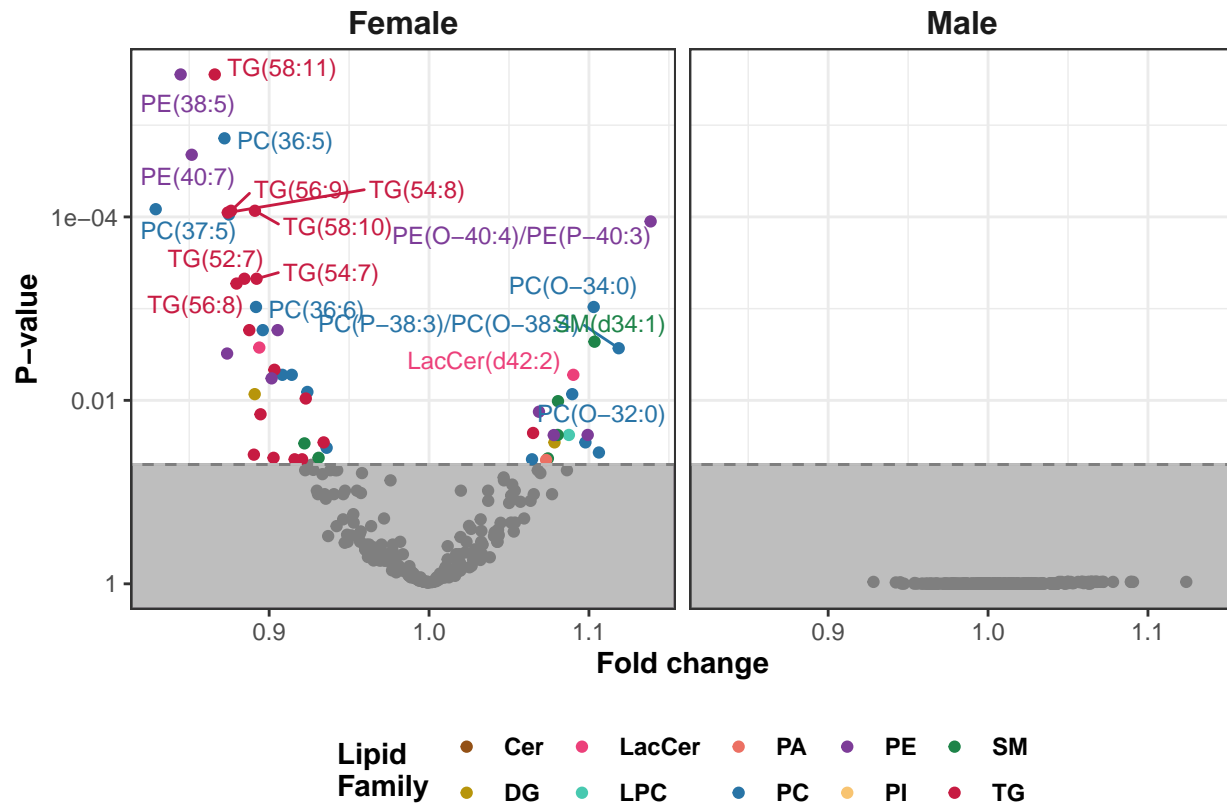
```
## Warning: Expected 2 pieces. Additional pieces discarded in 68 rows [69, 70, 71, 72, 73,
## 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, ...].
```

```
# Save figure
# pdf(here("figures/Volcano_sex_AD-ctrl_v1.4.pdf"), width = 8, height = 5)

fig3_A_volcano
```

```
## Warning: ggrepel: 250 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

```
## Warning: ggrepel: 268 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



```
# dev.off()

#Supplementary table
tmp_table <- ttest_male_AD %>%
  mutate(Sex = "Male") %>%
  rbind(., ttest_female_AD) %>%
  tibble()

# #Export supplementary table
# vroom_write(tmp_table, here("data/sup_table_ttest_v1.1.csv"))

#clean
rm(ttest_male_AD, tmp_table)

#difference between women and men with AD
ttest_sex_AD <- Ttest_extract(data_1 = data_female_AD,
                             data_2 = data_male_AD,
                             list_of_lipids = lipids)

#difference between healthy women and men
ttest_sex_ctrl <- Ttest_extract(data_1 = data_female_ctrl,
                               data_2 = data_male_ctrl,
                               list_of_lipids = lipids)

#Add Status variable before merging
ttest_sex_ctrl$Status <- "Healthy male vs. healthy female"
```

```

#Facet male and female ttest_AD
fig4_A_volcano <- ttest_sex_AD %>%
  mutate(Status = "Male with AD vs. female with AD") %>%
  rbind(ttest_sex_ctrl, .) %>%
  tibble() %>%
  left_join(., selected_lipids) %>%
  mutate(tmp_lipid = Lipid) %>%
  mutate(Lipid = gsub("_A", "", Lipid)) %>%
  mutate(Lipid = gsub("_B", "", Lipid)) %>%
  separate(tmp_lipid, sep = "\\(", c("Family", "Chain1")) %>%
  mutate(Family = ifelse(FDR > 0.05, "Not significant", Family)) %>%
  ggplot(aes(x = fold_change, y = -log10(FDR), color = Family))+
  geom_rect(aes(xmin = -Inf, xmax = Inf, ymin = -log10(0.05), ymax = -Inf),
    fill="grey", alpha=0.05, show.legend = FALSE)+
  geom_point(alpha = 1)+
  scale_color_manual(values = color_palette2,
    limits = names(color_palette2)[-11])+
  geom_hline(yintercept = -log10(0.05), colour = "#808080")+
  geom_hline(yintercept = -log10(0.05), linetype = "dashed", colour = "grey")+
  geom_text_repel(aes(label = Lipid), size = 3,
    max.overlaps = 8,
    show.legend = FALSE)+
  scale_y_continuous(limits = c(0,NA), labels = function(i) 10^-i)+
  ylab(label = "P-value")+
  xlab(label = "Fold change")+
  labs(color = "Lipid\nFamily")+
  facet_grid(.~Status)+
  theme_bw()+
  theme(strip.background = element_blank(),
    strip.text = element_text(face = "bold", size = 12),
    legend.text = element_text(face = "bold"),
    axis.title = element_text(face = "bold"),
    legend.title = element_text(face = "bold"))

```

```
## Joining with 'by = join_by(Lipid)'
```

```
## Warning: Expected 2 pieces. Additional pieces discarded in 68 rows [69, 70, 71, 72, 73,
## 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, ...].
```

```

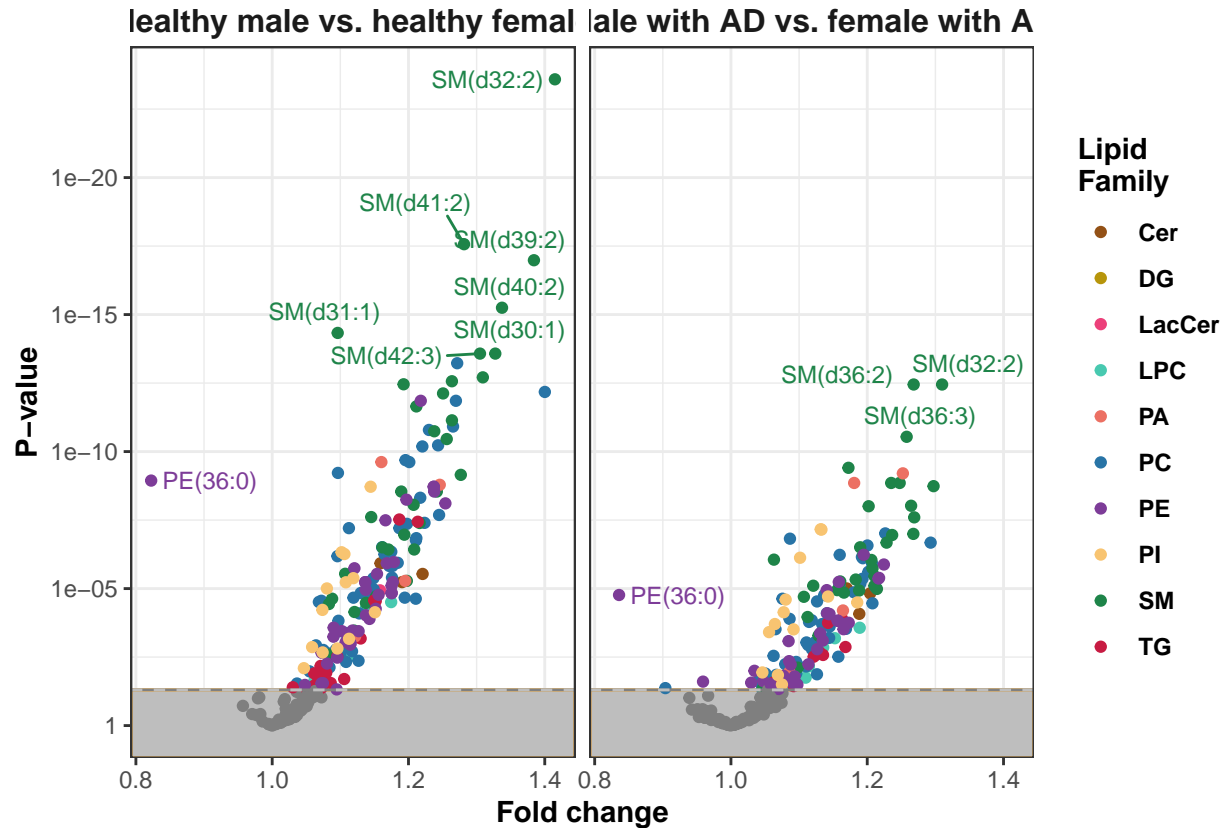
# #Save figure
# pdf(here("figures/Supplementary_volcano_Female_vs_male_v1.3.pdf"), width = 8, height = 5)

fig4_A_volcano

```

```
## Warning: ggrepel: 260 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

```
## Warning: ggrepel: 264 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



```
# dev.off()

rm(data_female_AD, data_female_ctrl, data_male_AD, data_male_ctrl)

library(eulerr)

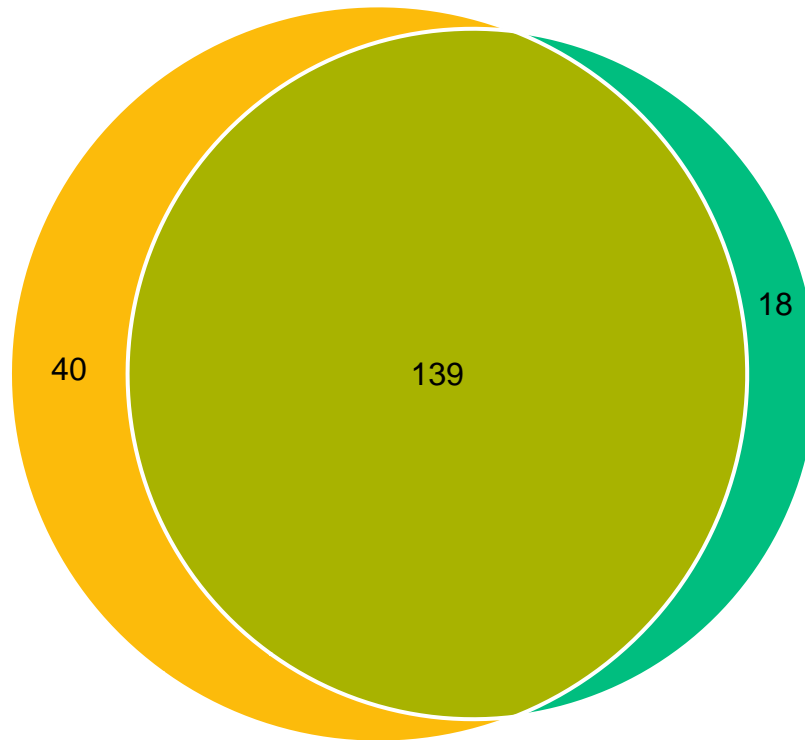
## Warning: pakke 'eulerr' blev bygget under R version 4.3.3

#Create list for plotting
venn_list <- list(
  "Healthy male vs. healthy female" = ttest_sex_ctrl$Lipid[ttest_sex_ctrl$FDR < 0.05],
  "Male with AD vs. female with AD" = ttest_sex_AD$Lipid[ttest_sex_AD$FDR < 0.05])

#Plot
fig4_B_venn <- plot(euler(venn_list),
  quantities = TRUE,
  fills = c("#FCBB0B", "#00BE7F", "#A8B300"),
  list(col = "white", lex = 2),
  legend = list(side = "bottom",
    cex = 0.8))

# #Save figure
# pdf(here("figures/Supplementary_venn_lipid_difference_v1.1.pdf"), width = 5, height = 3)

fig4_B_venn
```



- Healthy male vs. healthy female
- Male with AD vs. female with AD

```
# dev.off()
```

```
rm(venn_list)
```

```
#Subset AD vs. AD only
```

```
AD_vs_AD <- ttest_sex_AD[ttest_sex_AD$FDR < 0.05,][!ttest_sex_AD$Lipid[ttest_sex_AD$FDR < 0.05] %in% tt
```

```
#Plot
```

```
fig4_C_barplot <-
```

```
  AD_vs_AD %>%
```

```
  arrange(fold_change) %>%
```

```
  mutate(lipid_family = str_extract(Lipid, ".*?(?=\\()")) %>%
```

```
  mutate(Lipid = gsub("_A", "", Lipid)) %>%
```

```
  mutate(Lipid = gsub("_B", "", Lipid)) %>%
```

```
  mutate(Lipid = factor(Lipid, levels = c(rev(Lipid)))) %>%
```

```
  ggplot(aes(x = fold_change, y = Lipid, fill = lipid_family))+
```

```
  geom_bar(stat = "identity")+
```

```
  xlab(label = "Fold change")+
```

```
  ggtitle("Lipid differences between\\nmales and females with AD")+
```

```
  scale_fill_manual(values = color_palette2)+
```

```
  geom_vline(xintercept = 1, color = "black", linetype = "dashed")+
```

```
  coord_cartesian(xlim = c(0.8, 1.2))+
```

```
  theme_bw()+
```

```
  theme(axis.title.y = element_blank(),
```

```
        axis.title.x = element_text(face = "bold"),
```

```
        plot.title = element_text(size = 11, face = "bold"),
```

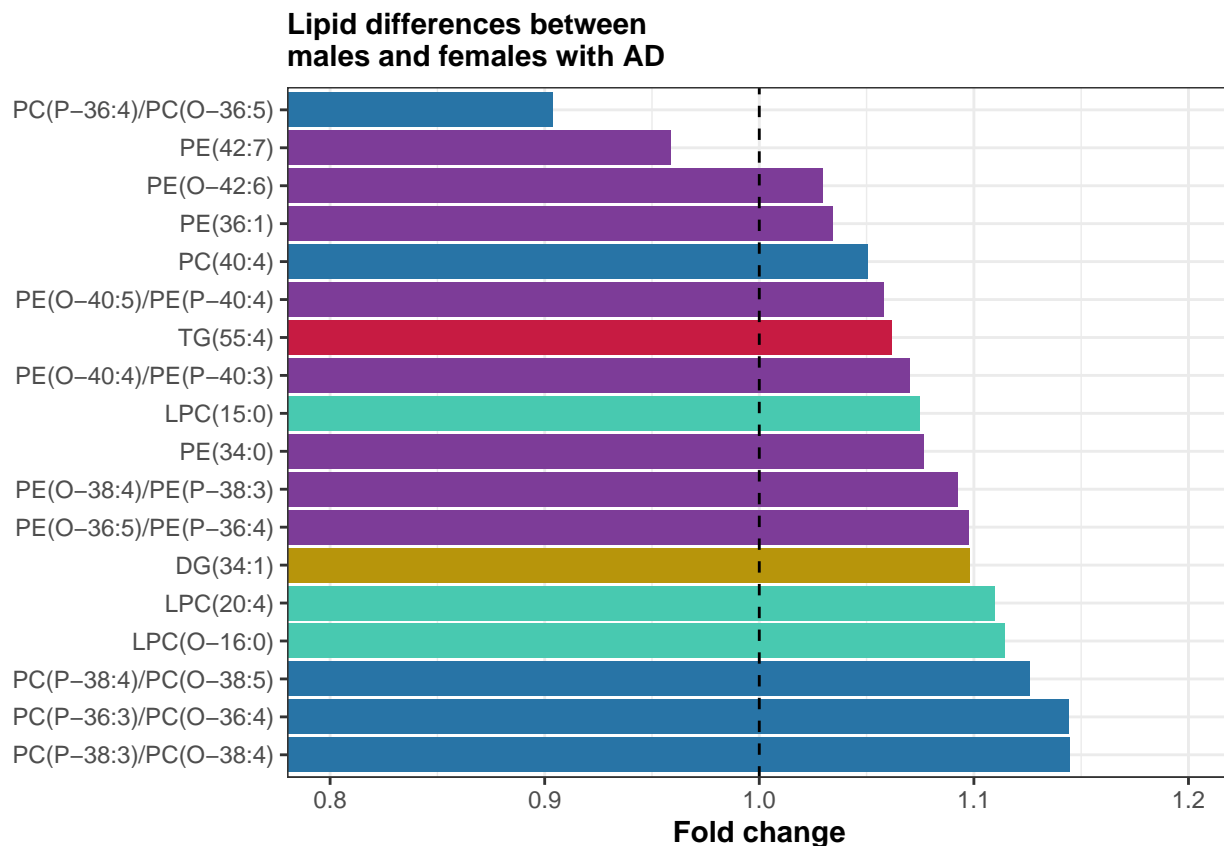
```

legend.position = "none")

# #Save figure
# pdf(here("figures/Supplementary_bar_lipid_difference_v1.3.pdf"), width = 6, height = 4)

fig4_C_barplot

```



```

# dev.off()

rm(ttest_sex_AD, ttest_sex_ctrl, AD_vs_AD)

library(patchwork)

#Merge plots for Figure 2
Figure4 <- fig4_A_volcano / (wrap_elements(fig4_B_venn, clip = FALSE) | fig4_C_barplot) +
  plot_annotation(tag_levels = 'A')

# #Save figure
# pdf(here("figures/Figure_4_lipid_differences_women_vs_men_v1.0.pdf"),
#     width = 8, height = 7)

Figure4

```

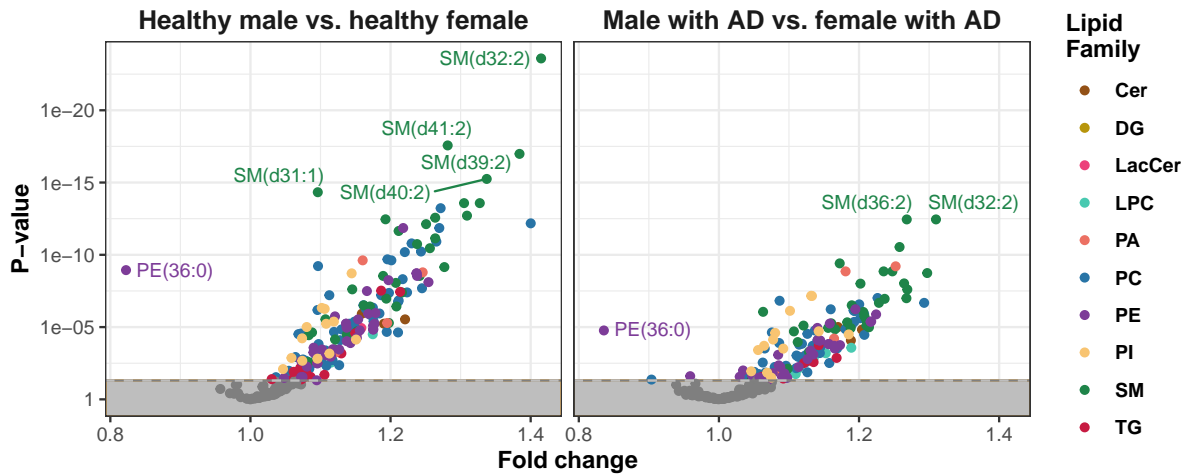
```

## Warning: ggrepel: 262 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps

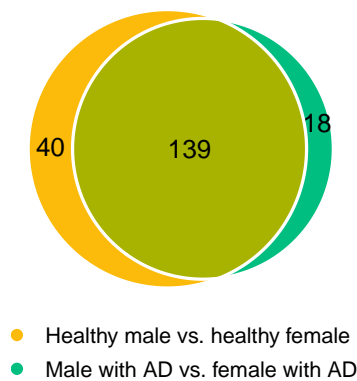
```

```
## Warning: ggrepel: 265 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

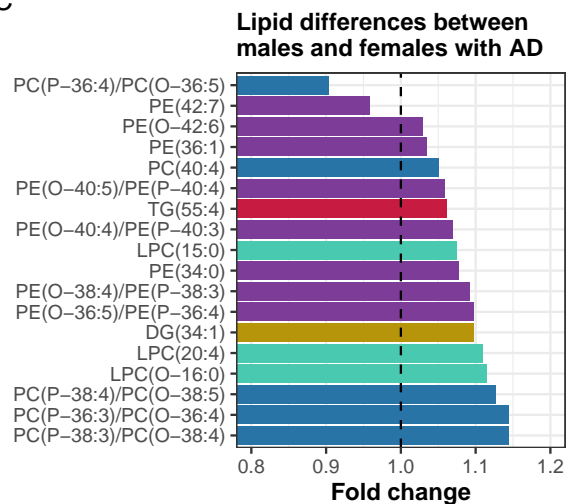
A



B



C



```
# dev.off()

rm(fig4_A_volcano, fig4_B_venn, fig4_C_barplot, Figure4)
```

```
#Boxplot of PE(38:5)
Box_PE <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
    labels = c("Ctrl", "MCI", "AD"))) %>%
  ggplot(aes(x = Status, y = `PE(38:5)`, fill = Status))+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(alpha = 0.1)+
  scale_fill_manual(values = c("#7d3c98",
    "#a17ead",
    "#ffe8ff"))+
  facet_grid(. ~ Sex)+
```



```

labs(title = "PE(38:5)")+
theme_bw()+
theme(axis.title.x = element_blank(),
      #axis.title.y = element_blank(),
      legend.position = "none",
      strip.background = element_blank(),
      strip.text = element_text(face = "bold", size = 12),
      axis.text.x = element_text(face = "bold"),
      plot.title = element_text(hjust = 0.5))

#Boxplot of TG(58:11)
Box_TG <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                        labels = c("Ctrl", "MCI", "AD"))) %>%
  ggplot(aes(x = Status, y = `TG(58:11)`, fill = Status))+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(alpha = 0.1)+
  scale_fill_manual(values = c("#c71b42",
                              "#ff6678",
                              "#ffa7b4"))+

  facet_grid(. ~ Sex)+
  labs(title = "TG(58:11)")+
  theme_bw()+
  theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.position = "none",
        strip.background = element_blank(),
        strip.text = element_text(face = "bold", size = 12),
        axis.text.x = element_text(face = "bold"),
        plot.title = element_text(hjust = 0.5))

#Boxplot of PC(0-34:0)
Box_PC <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                        labels = c("Ctrl", "MCI", "AD"))) %>%
  ggplot(aes(x = Status, y = `PC(0-34:0)`, fill = Status))+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(alpha = 0.1)+
  scale_fill_manual(values = c("#2874a6",
                              "#67a8dd",
                              "#d4f4ff"))+

  facet_grid(. ~ Sex)+
  labs(title = "PC(0-34:0)")+
  theme_bw()+
  theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.position = "none",
        strip.background = element_blank(),
        strip.text = element_text(face = "bold", size = 12),
        axis.text.x = element_text(face = "bold"),
        plot.title = element_text(hjust = 0.5))

#Boxplot of SM(d34:1)

```

```

Box_SM <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                          labels = c("Ctrl", "MCI", "AD"))) %>%
  ggplot(aes(x = Status, y = `SM(d34:1)`, fill = Status))+
  #geom_hline(yintercept = 0, color = "black")+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(alpha = 0.1)+
  scale_fill_manual(values = c("#1e8449",
                                "#5bba7a",
                                "#d4fadd"))+

  facet_grid(. ~ Sex)+
  labs(title = "SM(d34:1)")+
  theme_bw()+
  theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.position = "none",
        strip.background = element_blank(),
        strip.text = element_text(face = "bold", size = 12),
        axis.text.x = element_text(face = "bold"),
        plot.title = element_text(hjust = 0.5))

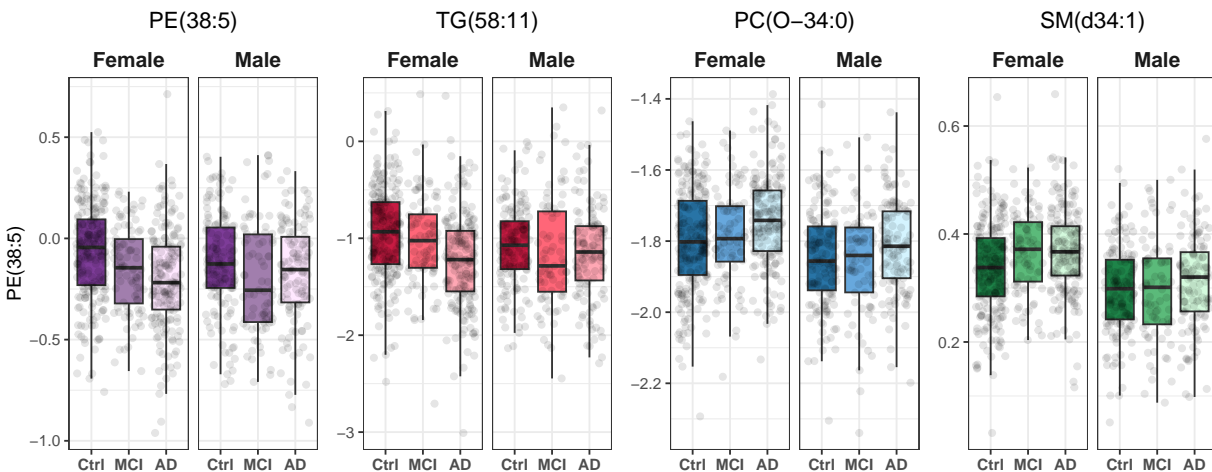
#Merge Plots
library(patchwork)
sup_fig2_boxplot <- plot_annotation(Box_PE | Box_TG | Box_PC | Box_SM)

#Save figure
# pdf(here("figures/Box_plot_individual_lipids_Ctrl_MCI_AD_v1.1.pdf"), width = 10, height = 4)

sup_fig2_boxplot

```

```
## $title
```



```
##
## $subtitle
## NULL
##
```

```
## $caption
## NULL
##
## $tag_levels
## NULL
##
## $tag_prefix
## NULL
##
## $tag_suffix
## NULL
##
## $tag_sep
## NULL
##
## $theme
## Named list()
## - attr(*, "class")= chr [1:2] "theme" "gg"
## - attr(*, "complete")= logi FALSE
## - attr(*, "validate")= logi TRUE
##
## attr("class")
## [1] "plot_annotation"
```

```
# dev.off()
```

```
#clean
```

```
#rm(Box_PC, Box_PE, Box_SM, Box_TG)
```

```
# library(patchwork)
```

```
#
```

```
# #Load heatmap from "ANM_Lipid_Regression.Rmd"
```

```
# fig2_heatmap <- readRDS(file = here("data/Heatmap_lipid_regression_v1.0.rds"))
```

```
#
```

```
# #Shuffling legends around for final figure 2 composistion
```

```
# fig2_volcano2 <- fig4_A_volcano+theme(legend.position = "bottom")
```

```
#
```

```
# fig2_heatmap2 <- fig2_heatmap+theme(legend.position = "none")
```

```
#
```

```
# Box_PE2 <- Box_PE+
```

```
#   ylab(label = "Abundance")+
```

```
#   theme(axis.title.y = element_text(face = "bold"),
```

```
#         axis.text.x = element_text(angle = 45))
```

```
#
```

```
# Box_TG2 <- Box_TG+
```

```
#   theme(axis.text.x = element_text(angle = 45))
```

```
#
```

```
# #Merge plots for Figure 2
```

```
# Figure2 <- (fig2_heatmap2 | (fig2_volcano2 / (Box_PE2 | Box_TG2 ))) +
```

```
#   plot_annotation(tag_levels = "A")
```

```
#
```

```
# # #Save figure
```

```
# # pdf(here("figures/Figue_2_individual_lipids_Analysis_v1.1.pdf"), width = 10, height = 9)
```

```
#
```

```

# Figure2
#
# # dev.off()

#Create subsets of the data stratified by sex and AD status
data_FAD <- data_norm %>%
  filter(Sex == "Female") %>%
  filter(Status == "ADC") %>%
  filter(!is.na(e4_c))

data_FCTRL <- data_norm %>%
  filter(Sex == "Female") %>%
  filter(Status == "CTL") %>%
  filter(!is.na(e4_c))

data_MAD <- data_norm %>%
  filter(Sex == "Male") %>%
  filter(Status == "ADC") %>%
  filter(!is.na(e4_c))

data_MCTRL <- data_norm %>%
  filter(Sex == "Male") %>%
  filter(Status == "CTL") %>%
  filter(!is.na(e4_c))

#t-tests
ttest_M0 <- Ttest_extract(data_1 = data_MAD[data_MAD$e4_c == 0,],
  data_2 = data_MCTRL[data_MCTRL$e4_c == 0,],
  list_of_lipids = lipids)

ttest_M1 <- Ttest_extract(data_1 = data_MAD[data_MAD$e4_c == 1,],
  data_2 = data_MCTRL[data_MCTRL$e4_c == 1,],
  list_of_lipids = lipids)

ttest_M2 <- Ttest_extract(data_1 = data_MAD[data_MAD$e4_c == 2,],
  data_2 = data_MCTRL[data_MCTRL$e4_c == 2,],
  list_of_lipids = lipids)

ttest_F0 <- Ttest_extract(data_1 = data_FAD[data_FAD$e4_c == 0,],
  data_2 = data_FCTRL[data_FCTRL$e4_c == 0,],
  list_of_lipids = lipids)

ttest_F1 <- Ttest_extract(data_1 = data_FAD[data_FAD$e4_c == 1,],
  data_2 = data_FCTRL[data_FCTRL$e4_c == 1,],
  list_of_lipids = lipids)

ttest_F2 <- Ttest_extract(data_1 = data_FAD[data_FAD$e4_c == 2,],
  data_2 = data_FCTRL[data_FCTRL$e4_c == 2,],
  list_of_lipids = lipids)

#Create APOE variable
ttest_M0$APOE <- ttest_F0$APOE <- "No APOE4"
ttest_M1$APOE <- ttest_F1$APOE <- "APOE4 Heterozygote"
ttest_M2$APOE <- ttest_F2$APOE <- "APOE4 Homozygote"

```

```

#Create sex variable
ttest_F0$Sex <- ttest_F1$Sex <- ttest_F2$Sex <- "Female"
ttest_M0$Sex <- ttest_M1$Sex <- ttest_M2$Sex <- "Male"

#Merge and plot as volcano plot
sup_volcano <- ttest_M0 %>%
  rbind(., ttest_M1, ttest_M2, ttest_F0, ttest_F1, ttest_F2) %>%
  tibble() %>%
  mutate(APOE = factor(APOE, levels = c("No APOE4",
                                         "APOE4 Heterozygote",
                                         "APOE4 Homozygote"))) %>%

  left_join(., selected_lipids) %>%
  mutate(tmp_lipid = Lipid) %>%
  separate(tmp_lipid, sep = "\\(", c("Family", "Chain1")) %>%
  mutate(Family = ifelse(FDR > 0.05, "Not significant", Family)) %>%
  ggplot(aes(x = fold_change, y = -log10(FDR), color = Family))+
  geom_rect(aes(xmin = -Inf, xmax = Inf, ymin = -log10(0.05), ymax = -Inf),
            fill="grey", alpha=0.05, show.legend = FALSE)+
  geom_point(alpha = 1)+
  scale_color_manual(values = color_palette2,
                    limits = names(color_palette2)[-11])+
  geom_hline(yintercept = -log10(0.05), linetype = "dashed", colour = "grey")+
  geom_text_repel(aes(label = Lipid), size = 2,
                 max.overlaps = 11,
                 show.legend = FALSE)+

  scale_y_continuous(limits = c(0,NA), labels = function(i) round(10^-i, 2))+
  ylab(label = "P-value")+
  xlab(label = "Fold change")+
  labs(color = "Lipid\nFamily")+
  facet_grid(APOE ~ Sex)+
  theme_bw()+
  theme(strip.background = element_blank(),
        strip.text = element_text(face = "bold", size = 12),
        legend.text = element_text(face = "bold"),
        axis.title = element_text(face = "bold"),
        legend.title = element_text(face = "bold"))

```

```
## Joining with 'by = join_by(Lipid)'
```

```
## Warning: Expected 2 pieces. Additional pieces discarded in 204 rows [69, 70, 71, 72, 73,
## 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, ...].
```

```

# #Save figure
# pdf(here("figures/Supplementary_volcano_AD-ctrl_APOE_sex_v1.0.pdf"), width = 7, height = 7)

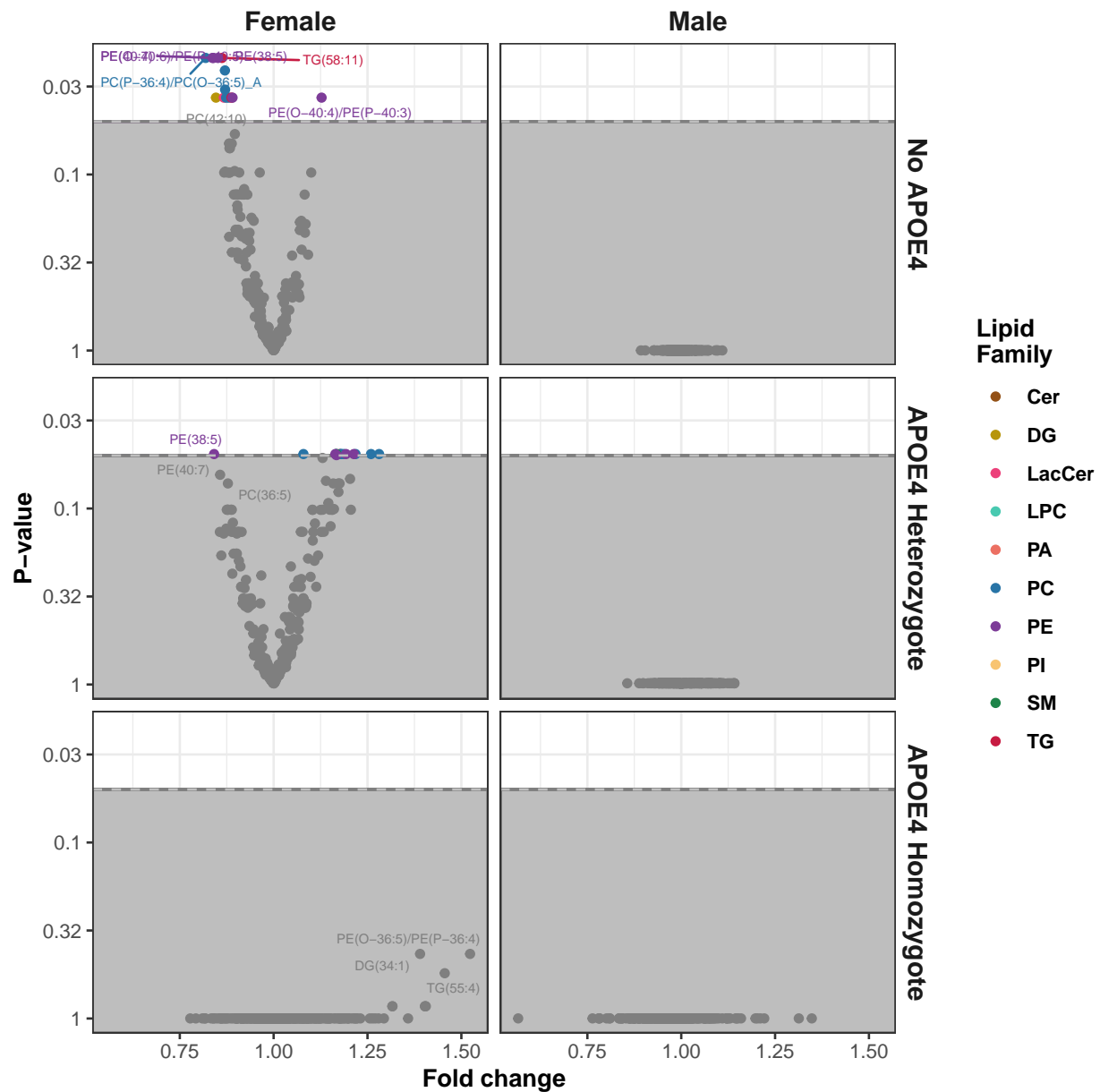
sup_volcano

```

```
## Warning: ggrepel: 261 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

```
## Warning: ggrepel: 265 unlabeled data points (too many overlaps). Consider increasing max.overlaps
## ggrepel: 265 unlabeled data points (too many overlaps). Consider increasing max.overlaps
```

```
## Warning: ggrepel: 268 unlabeled data points (too many overlaps). Consider increasing max.overlaps
## ggrepel: 268 unlabeled data points (too many overlaps). Consider increasing max.overlaps
## ggrepel: 268 unlabeled data points (too many overlaps). Consider increasing max.overlaps
```



```
# dev.off()
```

```
#clean
```

```
rm(data_FAD, data_FCTRL, data_MAD, data_MCTRL, sup_volcano,
    ttest_F0, ttest_F1, ttest_F2, ttest_M0, ttest_M1, ttest_M2,
    ttest_sex_AD, ttest_sex_ctrl)
```

```
## Warning in rm(data_FAD, data_FCTRL, data_MAD, data_MCTRL, sup_volcano,
## ttest_F0, : objekt 'ttest_sex_AD' blev ikke fundet
```

```
## Warning in rm(data_FAD, data_FCTRL, data_MAD, data_MCTRL, sup_volcano,
## ttest_F0, : objekt 'ttest_sex_ctrl' blev ikke fundet
```

```
rm(data_female_AD, data_female_ctrl, data_male_AD, data_male_ctrl, data_norm,
    Mm_normalize, Ttest_extract)
```

```
## Warning in rm(data_female_AD, data_female_ctrl, data_male_AD, data_male_ctrl, :
## objekt 'data_female_AD' blev ikke fundet
```

```
## Warning in rm(data_female_AD, data_female_ctrl, data_male_AD, data_male_ctrl, :
## objekt 'data_female_ctrl' blev ikke fundet
```

```
## Warning in rm(data_female_AD, data_female_ctrl, data_male_AD, data_male_ctrl, :
## objekt 'data_male_AD' blev ikke fundet
```

```
## Warning in rm(data_female_AD, data_female_ctrl, data_male_AD, data_male_ctrl, :
## objekt 'data_male_ctrl' blev ikke fundet
```

```
#Boxplot of PE(38:5)
Box_APOE_PE <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                          labels = c("Ctrl", "MCI", "AD"))) %>%
  mutate(APOE = factor(e4_c, levels = c(0, 1, 2),
                      labels = c("No APOE4", "APOE4 Heterozygote", "APOE4 Homozygote"))) %>%
  filter(!is.na(APOE)) %>%
  ggplot(aes(x = Status, y = `PE(38:5)`, fill = Status))+
  #geom_hline(yintercept = 0, color = "black")+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(alpha = 0.1)+
  scale_fill_manual(values = c("#7d3c98",
                              "#a17ead",
                              "#ffe8ff"))+

  facet_grid(APOE ~ Sex)+
  labs(title = "PE(38:5)")+
  theme_bw()+
  theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.position = "none",
        strip.background = element_blank(),
        strip.text.x.top = element_text(face = "bold", size = 12),
        strip.text.y.right = element_blank(),
        axis.text.x = element_text(face = "bold"),
        plot.title = element_text(hjust = 0.5))

#Boxplot of TG(58:11)
Box_APOE_TG <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                          labels = c("Ctrl", "MCI", "AD"))) %>%
  mutate(APOE = factor(e4_c, levels = c(0, 1, 2),
                      labels = c("No APOE4", "APOE4 Heterozygote", "APOE4 Homozygote"))) %>%
  filter(!is.na(APOE)) %>%
  ggplot(aes(x = Status, y = `TG(58:11)`, fill = Status))+
```

```

#geom_hline(yintercept = 0, color = "black")+
geom_boxplot(outlier.shape = NA)+
geom_jitter(alpha = 0.1)+
scale_fill_manual(values = c("#935116",
                              "#ce8347",
                              "#ffe9cf"))+

facet_grid(APOE ~ Sex)+
labs(title = "TG(58:11)")+
theme_bw()+
theme(axis.title.x = element_blank(),
      axis.title.y = element_blank(),
      legend.position = "none",
      strip.background = element_blank(),
      strip.text.x.top = element_text(face = "bold", size = 12),
      strip.text.y.right = element_blank(),
      axis.text.x = element_text(face = "bold"),
      plot.title = element_text(hjust = 0.5))

#Boxplot of PC(0-34:0)
Box_APOE_PC <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                        labels = c("Ctrl", "MCI", "AD"))) %>%
  mutate(APOE = factor(e4_c, levels = c(0, 1, 2),
                        labels = c("No APOE4", "APOE4 Heterozygote", "APOE4 Homozygote"))) %>%
  filter(!is.na(APOE)) %>%
  ggplot(aes(x = Status, y = `PC(0-34:0)`, fill = Status))+
  #geom_hline(yintercept = 0, color = "black")+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(alpha = 0.1)+
  scale_fill_manual(values = c("#2874a6",
                              "#67a8dd",
                              "#d4f4ff"))+

  facet_grid(APOE ~ Sex)+
  labs(title = "PC(0-34:0)")+
  theme_bw()+
  theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.position = "none",
        strip.background = element_blank(),
        strip.text.x.top = element_text(face = "bold", size = 12),
        strip.text.y.right = element_blank(),
        axis.text.x = element_text(face = "bold"),
        plot.title = element_text(hjust = 0.5))

#Boxplot of SM(d34:1)
Box_APOE_SM <- data %>%
  mutate(Status = factor(Status, levels = c("CTL", "MCI", "ADC"),
                        labels = c("Ctrl", "MCI", "AD"))) %>%
  mutate(APOE = factor(e4_c, levels = c(0, 1, 2),
                        labels = c("No APOE4", "APOE4 Heterozygote", "APOE4 Homozygote"))) %>%
  filter(!is.na(APOE)) %>%
  ggplot(aes(x = Status, y = `SM(d34:1)`, fill = Status))+
  #geom_hline(yintercept = 0, color = "black")+

```



```

geom_boxplot(outlier.shape = NA)+
geom_jitter(alpha = 0.1)+
scale_fill_manual(values = c("#1e8449",
                             "#5bba7a",
                             "#d4fadd"))+

facet_grid(APOE ~ Sex)+
labs(title = "SM(d34:1)")+
theme_bw()+
theme(axis.title.x = element_blank(),
      axis.title.y = element_blank(),
      legend.position = "none",
      strip.background = element_blank(),
      strip.text = element_text(face = "bold", size = 12),
      axis.text.x = element_text(face = "bold"),
      plot.title = element_text(hjust = 0.5))

```

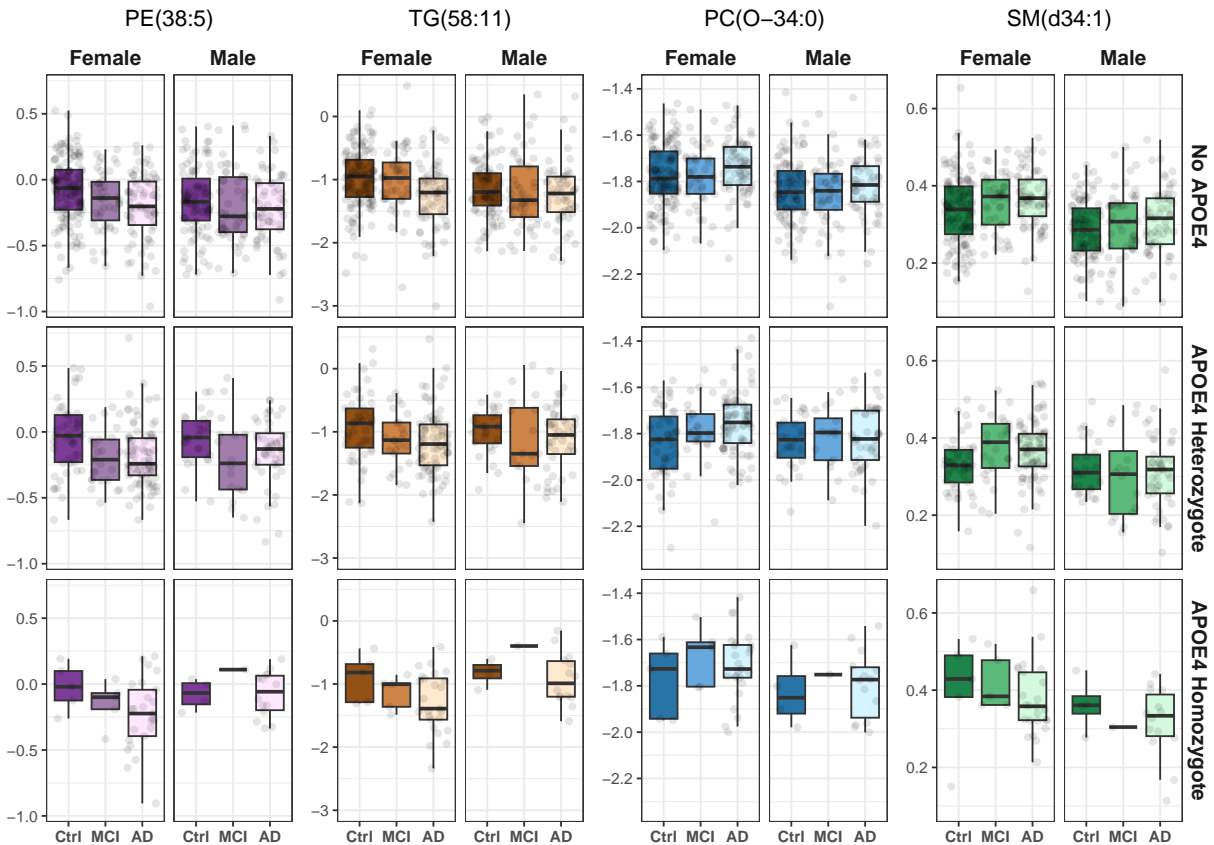
```
sup_boxplot_APOE <- plot_annotation(Box_APOE_PE | Box_APOE_TG | Box_APOE_PC | Box_APOE_SM)
```

```
# #Save figure
```

```
# pdf(here("figures/Supplementary_Box_plot_individual_lipids_APOE_Ctrl_MCI_AD_v1.0.pdf"), width = 10, h
```

```
sup_boxplot_APOE
```

```
## $title
```



```
##
## $subtitle
## NULL
##
## $caption
## NULL
##
## $tag_levels
## NULL
##
## $tag_prefix
## NULL
##
## $tag_suffix
## NULL
##
## $tag_sep
## NULL
##
## $theme
##   Named list()
##   - attr(*, "class")= chr [1:2] "theme" "gg"
##   - attr(*, "complete")= logi FALSE
##   - attr(*, "validate")= logi TRUE
##
## attr("class")
## [1] "plot_annotation"
```

```
# dev.off()
```

```
#clean
```

```
rm(Box_APOE_PE, Box_APOE_TG, Box_APOE_PC, Box_APOE_SM, sup_boxplot_APOE)
```