

Untitled

Stephanie Schultz

9/1/2024

Load libraries

```
library (ggplot2); library(dplyr); library(tidyr); library(ggrain); library(ggbreak);library(stringr);library(ggpubr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
## Registered S3 methods overwritten by 'ggpp':
```

```
## method from
```

```
## heightDetails.titleGrob ggplot2
```

```
## widthDetails.titleGrob ggplot2
```

```
## ggbreak v0.1.2
```

```
##
```

```
## If you use ggbreak in published research, please cite the following
```

```
## paper:
```

```
##
```

```
## S Xu, M Chen, T Feng, L Zhan, L Zhou, G Yu. Use ggbreak to effectively
```

```
## utilize plotting space to deal with large datasets and outliers.
```

```
## Frontiers in Genetics. 2021, 12:774846. doi: 10.3389/fgene.2021.774846
```

```
##
```

```
## Attaching package: 'ggpubr'
```

```
## The following object is masked from 'package:cowplot':
```

```
##
```

```
## get_legend
```

Load data

```

# Load raw data; for "homolog" variable: 0= no PSEN1/2 homolog, 1 = Yes PSEN1/2 homolog, 3= WT/na
data <- read.csv("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/PSEN1_2_homo

# Load gsm data
gsm_data <- read.csv("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/PSEN1_2_1

#merge
data_merge<-full_join( data,gsm_data,by=c("Orig_Variant", "PSEN"))

# load PSEN2 pathogenicity data
pathogenicity_psen2 <- read.csv("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homo

#merge
data_merge<-left_join(data, pathogenicity_psen2,by=c("Orig_Variant", "PSEN"))

regexp <- "[[:digit:]]+"
data_merge$variant_codon_PSEN2 <- as.numeric(str_extract(data_merge$PSEN2_homolog.y,regexp))

regexp <- "[[:digit:]]+"
data_merge$variant_codon_PSEN_orig <- as.numeric(str_extract(data_merge$Orig_Variant,regexp))

#convert long to wide
data_merge_wide <- data_merge %>%
  mutate(PSEN = ifelse(PSEN == 1, "PSEN1", "PSEN2")) %>%
  reshape( idvar = "PSEN2_homolog.x", timevar = "PSEN", direction = "wide")

#remove WT rows from long and wide df
data_merge_rmWT <- data_merge %>% filter (homolog != "3")
data_merge_wide_rmWT <- data_merge_wide %>% filter (homolog.PSEN2 != "3")

```

Supplemental Table 1 and Figure 1 Analyses

```

#force variable into correct type
data_merge_rmWT$PSEN <- as.factor(data_merge_rmWT$PSEN)
data_merge_rmWT$PSEN2_homolog.x <- as.factor(data_merge_rmWT$PSEN2_homolog.x)
data_merge_rmWT$homolog <- as.factor(data_merge_rmWT$homolog)
data_merge_rmWT$Pathogenicity <- as.factor(data_merge_rmWT$Pathogenicity)
data_merge_rmWT$AD_y_n <- as.factor(data_merge_rmWT$AD_y_n)
data_merge_rmWT$AD_path <- as.character(data_merge_rmWT$AD_path)
data_merge_rmWT$AD_path2 <- as.character(data_merge_rmWT$AD_path2)
data_merge_rmWT$AD_y_n2 <- ifelse(data_merge_rmWT$AD_path2 == "1", "1", "0")

# Supplemental Figure 1 ##
#Make a variable with PSEN1, PSEN2 non-homolog, and PSEN2 homolog groups
data_merge_rmWT$PSEN_homolog_cat <- ifelse(data_merge_rmWT$PSEN == "1", "0", ifelse(data_merge_rmWT$homolog == "3", "1", "0"))

#make a df for barchart with all data for selected variables to plot
barchart <- data_merge_rmWT %>% dplyr::select(AD_path2, PSEN_homolog_cat )
#get the frequencies of Pathogenicity classifications in homolog and non-homolog groups and make a df

```

```
print(freq_path_ingrps <- as.data.frame(barchart %>% dplyr::count(barchart$AD_path2, group = PSEN_homo
```

```
##   barchart$AD_path2 group  n
## 1                1     0 14
## 2                1     1  4
## 3                1     2  3
## 4                2     1 23
## 5                2     2  4
## 6                3     0  7
## 7                3     1 26
## 8                3     2 14
```

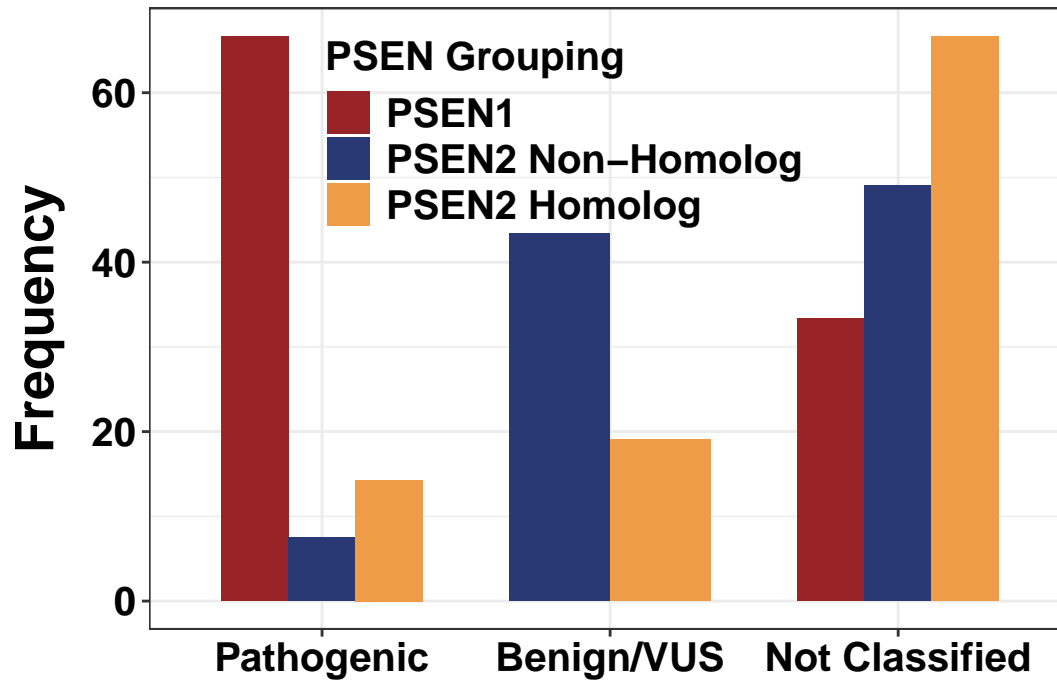
```
#make a new variable with these frequencies for plotting
freq_path_ingrps$Freq_grp <- ifelse(freq_path_ingrps$group == "1", freq_path_ingrps$n/53*100, freq_path_
#for the pathogenicity vairable into a normal variable
freq_path_ingrps$AD_path2 <- freq_path_ingrps$barchart$AD_path2`
```

```
#get a table with these freq
print(table <- freq_path_ingrps %>% arrange(group, desc(Freq_grp)))
```

```
##   barchart$AD_path2 group  n Freq_grp AD_path2
## 1                1     0 14 66.66667        1
## 2                3     0  7 33.33333        3
## 3                3     1 26 49.05660        3
## 4                2     1 23 43.39623        2
## 5                1     1  4  7.54717        1
## 6                3     2 14 66.66667        3
## 7                2     2  4 19.04762        2
## 8                1     2  3 14.28571        1
```

```
print(barchart <- ggplot(freq_path_ingrps, aes(fill=group, y=Freq_grp, x=as.factor(AD_path2))) +
  scale_x_discrete(labels=c("Pathogenic", "Benign/VUS", "Not Classified"))+
  scale_fill_manual(values = c( "#992428", "#283974", "#ee9c45"), labels = c("PSEN1", 'PSEN2 Non-Homolog
  geom_bar(position="dodge", stat="identity", show.legend=T, width =.7)+
  labs(x = "", y = "Frequency", fill = "PSEN Grouping")+
  theme_bw() +
  theme(axis.title.y = element_text(margin = margin(t = 0, r = 10, b = 0, l = 0), size =20, face=
  theme(axis.title.x = element_text(margin = margin(t = 10, r = 0, b = 0, l = 0), size =20, face=
  theme(legend.title = element_text(size=15, family="Helvetica", face="bold"))+
  theme(legend.background = element_rect(fill="transparent"), legend.key = element_rect(fill = "t
  theme(legend.position = c(0.45,.80))+
  theme(legend.text = element_text(size=15, family="Helvetica", face="bold"))+
  theme(axis.text = element_text( size = 15, face="bold", color = "black"), axis.title.x=element_l
  theme(plot.margin = unit(c(0.5, 0.5, 0.5, 0.5), "inches")))
```

```
## Warning: A numeric 'legend.position' argument in 'theme()' was deprecated in ggplot2
## 3.5.0.
## i Please use the 'legend.position.inside' argument of 'theme()' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```



Supplementary Table 1

#make a df for barchart with just psen2 data for selected variables to plot

```
barchart <- data_merge_rmWT %>% dplyr::select(Pathogenicity, PSEN_homolog_cat )
```

#get the frequencies of Pathogenicity classifications in homolog and non-homolog groups and make a df

```
print(freq_path_ingrps <- as.data.frame(barchart %>% dplyr::count(barchart$Pathogenicity, group = PSEN,
```

```
##          barchart$Pathogenicity group  n
## 1                AD : Benign        1 16
## 2                AD : Benign        2  1
## 3      AD : Benign, FTD : Not Classified 1  1
## 4                AD : Likely Benign 1  1
## 5                AD : Likely Pathogenic 0  3
## 6                AD : Likely Pathogenic 1  3
## 7                AD : Likely Pathogenic 2  1
## 8                AD : Not Classified  0  7
## 9                AD : Not Classified  1 18
## 10               AD : Not Classified  2 12
## 11      AD : Not Classified, DLB : Not Classified 2  1
## 12               AD : Pathogenic    0 11
## 13               AD : Pathogenic    1  1
## 14               AD : Pathogenic    2  2
## 15               AD : Uncertain Significance 1  4
## 16               AD : Uncertain Significance 2  3
## 17 AD : Uncertain Significance, FTD : Not Classified 1  1
## 18      Atypical Dementia : Not Classified 2  1
## 19               bvFTD : Not Classified 1  1
```

```
## 20          Dementia : Not Classified      1  1
## 21          FTD : Not Classified          1  3
## 22          MCI : Not Classified          1  1
## 23          PD : Not Classified           1  1
## 24          svPPA : Not Classified        1  1
```

```
#make a new variable with these frequencies for plotting
freq_path_ingrps$Freq_grp <- ifelse(freq_path_ingrps$group == "1", freq_path_ingrps$n/53*100, freq_path_
#for the pathogenicity variable into a normal variable
freq_path_ingrps$Pathogenicity <- freq_path_ingrps$barchart$Pathogenicity`

#get a table with these freq
print(table <- freq_path_ingrps %>% arrange(group, desc(Freq_grp)))
```

```
##          barchart$Pathogenicity group  n  Freq_grp
## 1          AD : Pathogenic           0 11 52.380952
## 2          AD : Not Classified        0  7 33.333333
## 3          AD : Likely Pathogenic     0  3 14.285714
## 4          AD : Not Classified        1 18 33.962264
## 5          AD : Benign                1 16 30.188679
## 6          AD : Uncertain Significance 1  4  7.547170
## 7          AD : Likely Pathogenic     1  3  5.660377
## 8          FTD : Not Classified        1  3  5.660377
## 9          AD : Benign, FTD : Not Classified 1  1  1.886792
## 10         AD : Likely Benign          1  1  1.886792
## 11         AD : Pathogenic             1  1  1.886792
## 12 AD : Uncertain Significance, FTD : Not Classified 1  1  1.886792
## 13         bvFTD : Not Classified      1  1  1.886792
## 14         Dementia : Not Classified   1  1  1.886792
## 15         MCI : Not Classified        1  1  1.886792
## 16         PD : Not Classified         1  1  1.886792
## 17         svPPA : Not Classified      1  1  1.886792
## 18         AD : Not Classified         2 12 57.142857
## 19         AD : Uncertain Significance  2  3 14.285714
## 20         AD : Pathogenic             2  2  9.523810
## 21         AD : Benign                 2  1  4.761905
## 22         AD : Likely Pathogenic      2  1  4.761905
## 23         AD : Not Classified, DLB : Not Classified 2  1  4.761905
## 24         Atypical Dementia : Not Classified 2  1  4.761905
##          Pathogenicity
## 1          AD : Pathogenic
## 2          AD : Not Classified
## 3          AD : Likely Pathogenic
## 4          AD : Not Classified
## 5          AD : Benign
## 6          AD : Uncertain Significance
## 7          AD : Likely Pathogenic
## 8          FTD : Not Classified
## 9          AD : Benign, FTD : Not Classified
## 10         AD : Likely Benign
## 11         AD : Pathogenic
## 12 AD : Uncertain Significance, FTD : Not Classified
## 13         bvFTD : Not Classified
## 14         Dementia : Not Classified
```

```
## 15                MCI : Not Classified
## 16                PD : Not Classified
## 17                svPPA : Not Classified
## 18                AD : Not Classified
## 19                AD : Uncertain Significance
## 20                AD : Pathogenic
## 21                AD : Benign
## 22                AD : Likely Pathogenic
## 23                AD : Not Classified, DLB : Not Classified
## 24                Atypical Dementia : Not Classified
```

```
### fig2 A00 vs ab4240 ###
```

```
data_merge_rmWT$homolog <- as.factor(data_merge_rmWT$homolog)
data_merge_rmWT$AD_path2 <- as.factor(data_merge_rmWT$AD_path2)
```

```
# correlation between A00 and ab4240 in all PSEN2 #
```

```
data_merge_rmWT.allpsen2 <- data_merge_rmWT %>%
  filter(PSEN == "2")
```

```
cor.test(data_merge_rmWT.allpsen2$A00, data_merge_rmWT.allpsen2$Abeta_4240)
```

```
##
## Pearson's product-moment correlation
##
## data: data_merge_rmWT.allpsen2$A00 and data_merge_rmWT.allpsen2$Abeta_4240
## t = -2.2786, df = 65, p-value = 0.02598
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.48077375 -0.03398606
## sample estimates:
## cor
## -0.2719744
```

```
#
# data: data_merge_rmWT.allpsen2$A00 and data_merge_rmWT.allpsen2$Abeta_4240
# t = -2.2786, df = 65, p-value = 0.02598
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# -0.48077375 -0.03398606
# sample estimates:
# cor
# -0.2719744
```

```
print(fig2a <- data_merge_rmWT %>%
  filter(PSEN == "2")%>%
  ggplot( aes(x=A00, y=Abeta_4240)) +
  geom_hline (yintercept=0.1451153, linetype="dashed", size=1, color= '#3d2251')+
  geom_point(size=3,show.legend=F, aes(color = homolog, shape =AD_path2, fill = homolog), stroke = 2, alpha=0.5)+
  scale_color_manual(values = c("#283974","#ee9c45"), labels=c("non-homolog", "homolog"))+
  scale_fill_manual(values = c("#283974","#ee9c45"), labels=c("non-homolog", "homolog"))+
  #scale_shape_manual(values = c(15,16))+
  scale_shape_manual(values = c(24,22,21), labels = c("Pathogenic", "Benign/VUS", "Not Classified"))+
  labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 42/40"))), title = "All PSEN2"))
```

```

theme_bw() +
xlim(40,90) +
ylim(0,.5)+
  geom_smooth ( method = "lm", size = 2,se=TRUE, color="black")+
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(axis.title.x = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.15,.80))+
  theme(title = element_text(size = 25, face="bold")) +
  theme(plot.margin = margin(2, 1, 1, 1, "cm"))

```

```

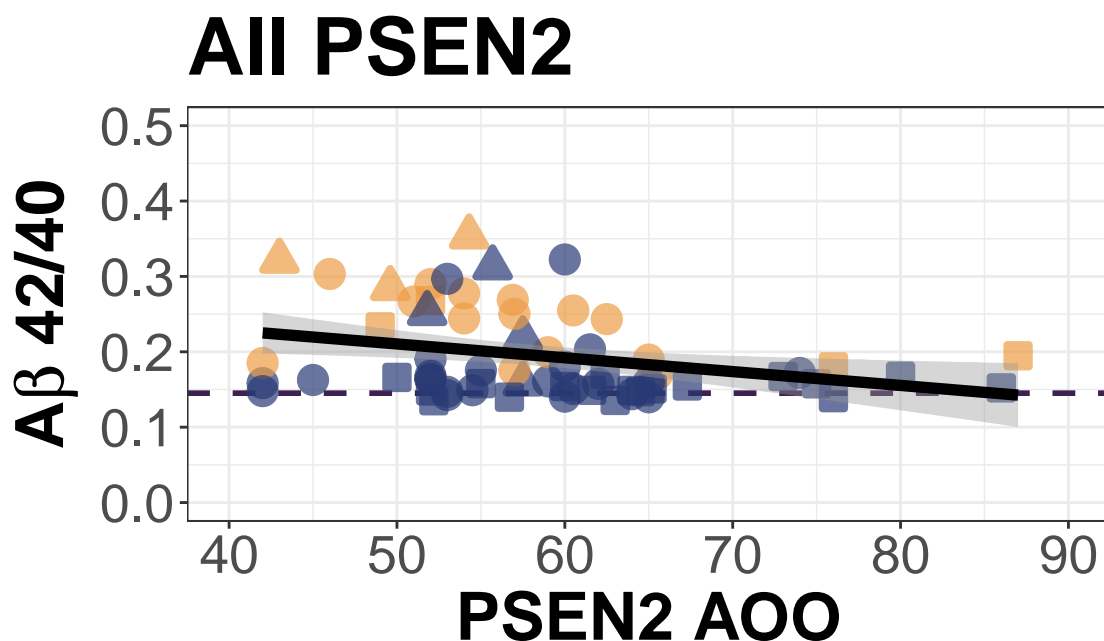
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.

## 'geom_smooth()' using formula = 'y ~ x'

## Warning: Removed 8 rows containing non-finite outside the scale range
## ('stat_smooth()').

## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').

```



```

print(fig2b <- data_merge_rmWT %>%
      filter(PSEN == "2")%>%
      ggplot( aes(x=A00, y=Abeta_3742)) +
        geom_hline (yintercept=0.4353022, linetype="dashed", size=1, color= '#3d2251')+
      geom_point(size=3,show.legend=F, aes(color = homolog, shape =AD_path2, fill = homolog), stroke = 2, alpha=0.5) +
        scale_color_manual(values = c("#283974","#ee9c45"), labels=c("non-homolog", "homolog"))+
        scale_fill_manual(values = c("#283974","#ee9c45"), labels=c("non-homolog", "homolog"))+
        #scale_shape_manual(values = c(15,16))+
        scale_shape_manual(values = c(24,22,21), labels = c("Pathogenic", "Benign/VUS", "Not Classified"))+
      theme_bw() +
      xlim(40,90) +
      ylim(0,.5)+
      geom_smooth ( method = "lm", size = 2,se=TRUE, color="black")+
        # annotate("text", size = 8, x=55, y=0.05, label= "r = 0.251, p = 0.040",color = "black",
      labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 37/42"))), title = "All PSEN2" )+
      theme(axis.text = element_text(size = 20))+
      theme(legend.title=element_text(size = 25, face="bold")) +
      theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
      theme(axis.title.y = element_text(size = 25, face="bold")) +
      theme(axis.title.x = element_text(size = 25, face="bold")) +
      theme(legend.position = c(0.15,.90))+
      theme(title = element_text(size = 25, face="bold")) +
      theme(plot.margin = margin(2, 1, 1, 1, "cm"))

```

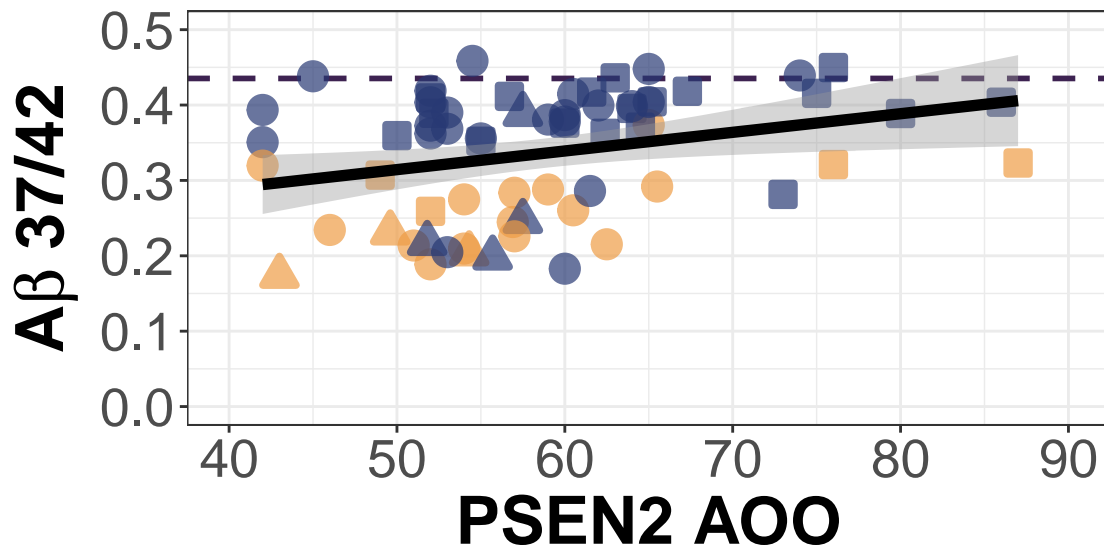
```
## 'geom_smooth()' using formula = 'y ~ x'
```

```

## Warning: Removed 8 rows containing non-finite outside the scale range ('stat_smooth()').
## Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').

```


AII PSEN2



```
data_merge_rmWT.allpsen2 <- data_merge_rmWT %>%
  filter(PSEN == "2")

cor.test(data_merge_rmWT.allpsen2$A00, data_merge_rmWT.allpsen2$Abeta_3742)

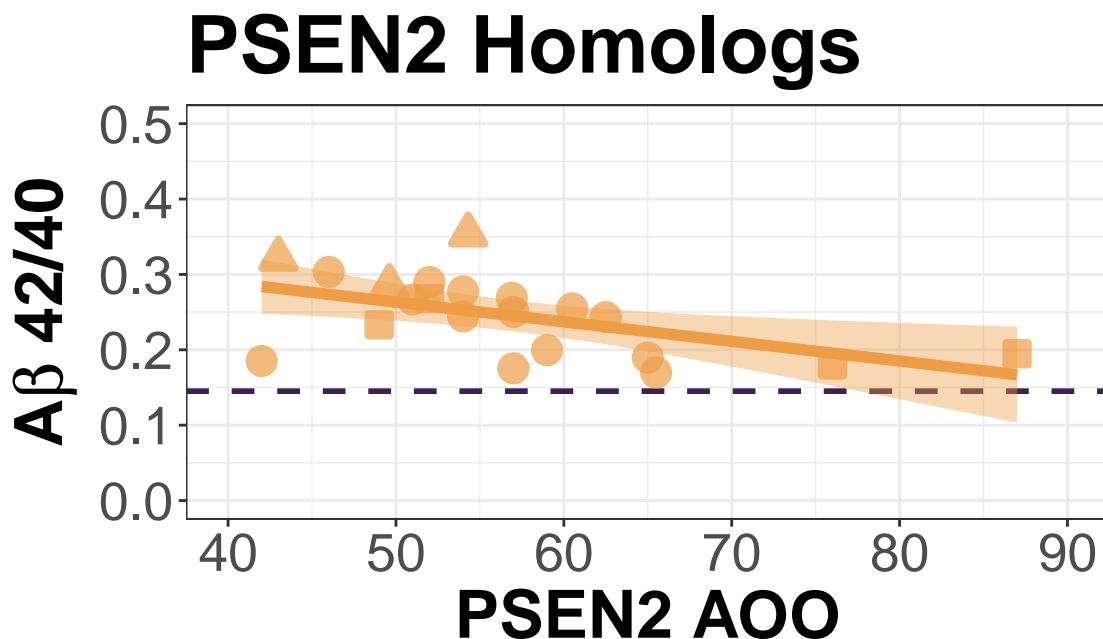
##
## Pearson's product-moment correlation
##
## data: data_merge_rmWT.allpsen2$A00 and data_merge_rmWT.allpsen2$Abeta_3742
## t = 2.0946, df = 65, p-value = 0.0401
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.01197507 0.46366241
## sample estimates:
##      cor
## 0.2514604

#
# data: data_merge_rmWT.allpsen2$A00 and data_merge_rmWT.allpsen2$Abeta_3742
# t = 2.0946, df = 65, p-value = 0.0401
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
#  0.01197507 0.46366241
# sample estimates:
#      cor
```

```
# 0.2514604

print(fig2c <- data_merge_rmWT.allpsen2 %>%
  filter(homolog == "1")%>%
  ggplot( aes(x=A00, y=Abeta_4240)) +
    geom_hline (yintercept=0.1451153, linetype="dashed", size=1, color= '#3d2251')+
    geom_point(size=3,show.legend=F, aes(color = homolog, shape =AD_path2, fill = homolog), stroke = 2, alpha=0.5)+
    scale_color_manual(values = c("#ee9c45"))+
    scale_fill_manual(values = c("#ee9c45"))+
    scale_shape_manual(values = c(24,22,21), labels = c("Pathogenic", "Benign/VUS", "Not Classified")),
  labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 42/40"))), title = "PSEN2 Homologs"),
  theme_bw() +
  xlim(40,90) +
  ylim(0,.5)+
  geom_smooth ( method = "lm", size = 2,se=TRUE, color= '#ee9c45', fill = '#ee9c45')+
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(axis.title.x = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.15,.80))+
  theme(title = element_text(size = 25, face="bold")) +
  theme(plot.margin = margin(2, 1, 1, 1, "cm")))
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
data_merge.psen2.homolog <- data_merge_rmWT.allpsen2 %>%
  filter(homolog == "1")

cor.test(data_merge.psen2.homolog$A00, data_merge.psen2.homolog$Abeta_4240)

##
## Pearson's product-moment correlation
##
## data: data_merge.psen2.homolog$A00 and data_merge.psen2.homolog$Abeta_4240
## t = -2.7218, df = 19, p-value = 0.01354
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.7824334 -0.1269896
## sample estimates:
## cor
## -0.5296416
```

```
#
# data: data_merge.psen2.homolog$A00 and data_merge.psen2.homolog$Abeta_4240
# t = -2.7218, df = 19, p-value = 0.01354
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# -0.7824334 -0.1269896
# sample estimates:
# cor
# -0.5296416
```

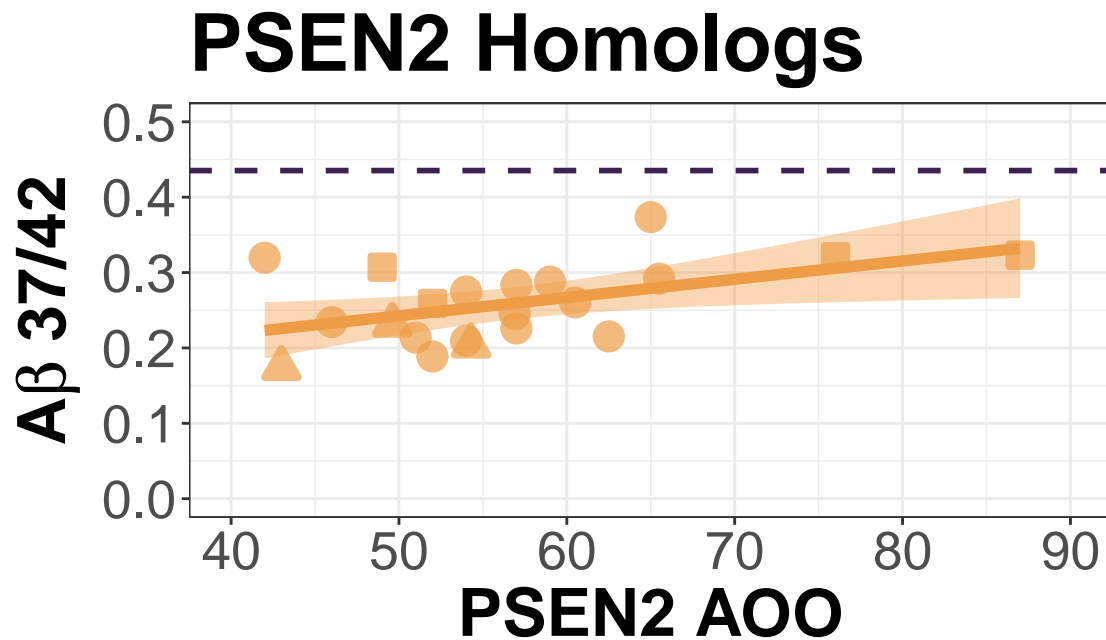
```
print(fig2d <- data_merge_rmWT.allpsen2 %>%
  filter(homolog == "1")%>%
  ggplot( aes(x=A00, y=Abeta_3742)) +
    geom_hline (yintercept=0.4353022, linetype="dashed", size=1, color= '#3d2251')+
    geom_point(size=3,show.legend=F, aes(color = homolog, shape =AD_path2, fill = homolog), stroke
      scale_color_manual(values = c("#ee9c45"))+
      scale_fill_manual(values = c("#ee9c45"))+
      #scale_shape_manual(values = c(15,16))+
      scale_shape_manual(values = c(24,22,21), labels = c("Pathogenic", "Benign/VUS", "Not Classifi
        labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 37/42"))), title = "PSEN2 Hom
          #geom_text_repel(aes(label=ifelse(Abeta_3742>0.3,as.character(Orig_Variant),'')),box.padd

    geom_smooth ( method = "lm", size = 2,se=TRUE, color= '#ee9c45', fill = '#ee9c45')+
      # annotate("text", size = 8, x=55, y=0.05, label= "r = 0.490, p = 0.024",color = "bla
        labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 37/42"))), title = "PSEN2 Hom

  theme_bw() +
  xlim(40,90) +
  ylim(0,.5)+
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
```

```
theme(axis.title.x = element_text(size = 25, face="bold")) +
theme(legend.position = c(0.15,.90))+
theme(plot.margin = margin(2, 1, 1, 1, "cm")))
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
data_merge.psen2.homolog <- data_merge_rmWT.allpsen2 %>%
  filter(homolog == "1")

cor.test(data_merge.psen2.homolog$A00, data_merge.psen2.homolog$Abeta_3742)
```

```
##
## Pearson's product-moment correlation
##
## data: data_merge.psen2.homolog$A00 and data_merge.psen2.homolog$Abeta_3742
## t = 2.4513, df = 19, p-value = 0.02408
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.07417662 0.76085780
## sample estimates:
## cor
## 0.4901677
```

```

#
# data: data_merge.psen2.homolog$A00 and data_merge.psen2.homolog$Abeta_3742
# t = 2.4513, df = 19, p-value = 0.02408
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# 0.07417662 0.76085780
# sample estimates:
# cor
# 0.4901677
#

print(fig2e <- data_merge_rmWT.allpsen2 %>%
  filter(homolog != "1")%>%      ggplot( aes(x=A00, y=Abeta_4240, label = Orig_Variant))
  geom_hline (yintercept=0.1451153, linetype="dashed", size=1, color= '#3d2251')+
  geom_point(size=3,show.legend=F, aes(color = homolog, shape =AD_path2, fill = homolog), stroke = 2, al
  geom_text_repel(aes(label=ifelse(Abeta_4240>0.2,as.character(Orig_Variant),'')),box.padding = unit(
    scale_color_manual(values = c("#283974"))+
    scale_fill_manual(values = c("#283974"))+
    #scale_shape_manual(values = c(15,16))+
    scale_shape_manual(values = c(24,22,21), labels = c("AD Pathogenic", "AD unclear/benign", "non
    labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 42/40"))), title = "PSEN2 Non
  theme_bw() +
  xlim(40,90) +
  ylim(0,.5)+
  geom_smooth ( method = "lm", size = 2,se=TRUE, color= '#283974', fill = '#283974')+
  # annotate("text", size = 8, x=55, y=0.05, label= "r = -0.105, p = 0.487",color = "black", fo
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(axis.title.x = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.15,.80))+
  theme(title = element_text(size = 25, face="bold")) +
  theme(plot.margin = margin(2, 1, 1, 1, "cm")))

## 'geom_smooth()' using formula = 'y ~ x'

## Warning: Removed 8 rows containing non-finite outside the scale range
## ('stat_smooth()').

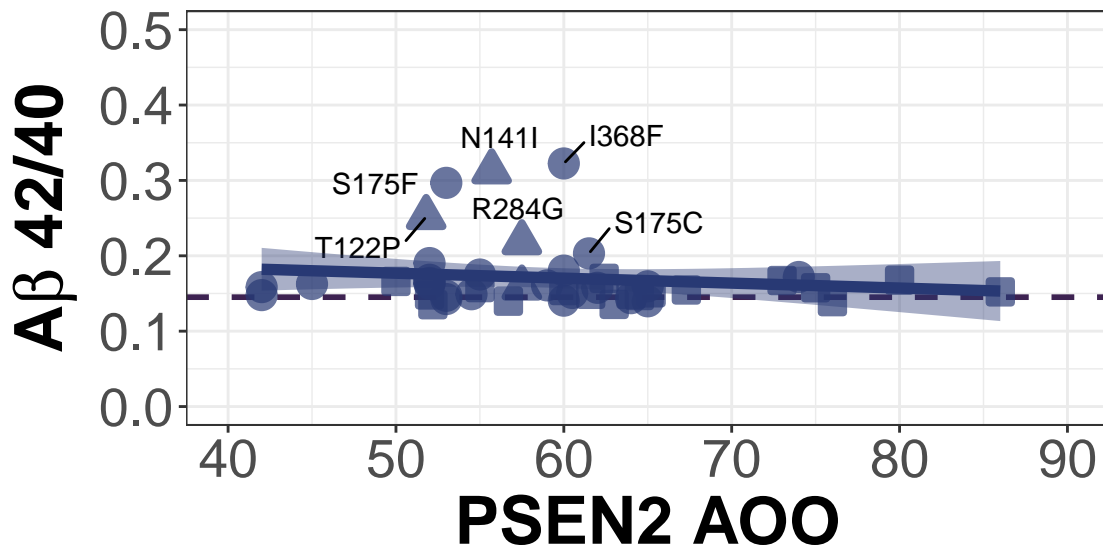
## Warning: The following aesthetics were dropped during statistical transformation: label.
## i This can happen when ggplot fails to infer the correct grouping structure in
## the data.
## i Did you forget to specify a 'group' aesthetic or to convert a numerical
## variable into a factor?

## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').

## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_text_repel()').

```

PSEN2 Non-Homologs



```
data_merge.psen2.nonhomolog <- data_merge_rmWT.allpsen2 %>%
  filter(homolog == "0")
```

```
cor.test(data_merge.psen2.nonhomolog$AOO, data_merge.psen2.nonhomolog$Abeta_4240)
```

```
##
## Pearson's product-moment correlation
##
## data: data_merge.psen2.nonhomolog$AOO and data_merge.psen2.nonhomolog$Abeta_4240
## t = -0.70114, df = 44, p-value = 0.4869
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.3837051 0.1910110
## sample estimates:
## cor
## -0.1051158
```

```
# data: data_merge.psen2.nonhomolog$AOO and data_merge.psen2.nonhomolog$Abeta_4240
# t = -0.70114, df = 44, p-value = 0.4869
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# -0.3837051 0.1910110
# sample estimates:
# cor
# -0.1051158
```

```
summary(lm(Abeta_3742 ~A00 + AD_path2,data_merge_rmWT.allpsen2))
```

```
##
## Call:
## lm(formula = Abeta_3742 ~ A00 + AD_path2, data = data_merge_rmWT.allpsen2)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.151417	-0.052135	0.009577	0.049763	0.147830

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.1961184	0.0589698	3.326	0.001474 **
A00	0.0007746	0.0009878	0.784	0.435838
AD_path22	0.1314390	0.0333167	3.945	0.000203 ***
AD_path23	0.0917089	0.0302592	3.031	0.003537 **

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07296 on 63 degrees of freedom
## (7 observations deleted due to missingness)
## Multiple R-squared:  0.249, Adjusted R-squared:  0.2132
## F-statistic: 6.963 on 3 and 63 DF, p-value: 0.0004038
```

```
# lm(formula = Abeta_3742 ~ A00 + AD_path2, data = data_merge_rmWT.allpsen2)
#
# Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.151417	-0.052135	0.009577	0.049763	0.147830

```
#
# Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.1961184	0.0589698	3.326	0.001474 **
A00	0.0007746	0.0009878	0.784	0.435838
AD_path22	0.1314390	0.0333167	3.945	0.000203 ***
AD_path23	0.0917089	0.0302592	3.031	0.003537 **

```
summary(lm(Abeta_4240 ~A00 + AD_path2,data_merge_rmWT.allpsen2))
```

```
##
## Call:
## lm(formula = Abeta_4240 ~ A00 + AD_path2, data = data_merge_rmWT.allpsen2)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.11046	-0.03416	-0.01462	0.03309	0.12706

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.3021536  0.0399735   7.559 2.11e-10 ***
## A00          -0.0005823  0.0006696  -0.870 0.387777
## AD_path22    -0.1015028  0.0225842  -4.494 3.05e-05 ***
## AD_path23    -0.0717701  0.0205116  -3.499 0.000863 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04945 on 63 degrees of freedom
## (7 observations deleted due to missingness)
## Multiple R-squared:  0.2989, Adjusted R-squared:  0.2655
## F-statistic: 8.951 on 3 and 63 DF,  p-value: 5.037e-05

#
# lm(formula = Abeta_4240 ~ A00 + AD_path2, data = data_merge_rmWT.allpsen2)
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -0.11046 -0.03416 -0.01462  0.03309  0.12706
#
# Coefficients:
#           Estimate Std. Error t value Pr(>|t|)
# (Intercept)  0.3021536  0.0399735   7.559 2.11e-10 ***
# A00          -0.0005823  0.0006696  -0.870 0.387777
# AD_path22    -0.1015028  0.0225842  -4.494 3.05e-05 ***
# AD_path23    -0.0717701  0.0205116  -3.499 0.000863 ***

print(fig2f <- data_merge_rmWT.allpsen2 %>%
  filter(homolog != "1")%>%
  ggplot( aes(x=A00, y=Abeta_3742, label = Orig_Variant)) +
    geom_hline (yintercept=0.4353022, linetype="dashed", size=1, color= '#3d2251')+
  geom_point(size=3,show.legend=F, aes(color = homolog, shape =AD_path2, fill = homolog), stroke = 2, alpha=0.5)+
  scale_color_manual(values = c("#283974"))+
  scale_fill_manual(values = c("#283974"))+
  geom_text_repel(aes(label=ifelse(Abeta_3742<0.3,as.character(Orig_Variant),'')),box.padding = unit(0.5pt, 'cm'))

  #scale_shape_manual(values = c(15,16))+
  scale_shape_manual(values = c(24,22,21), labels = c("Pathogenic", "Benign/VUS", "Not Classified"))+
  geom_smooth ( method = "lm", size = 2,se=TRUE, color= '#283974', fill = '#283974')+
  # annotate("text", size = 8, x=55, y=0.05, label= "r = 0.147, p = 0.328",color = "black")+
  labs( x="PSEN2 A00", y=expression(bold(paste("A", beta, " 37/42"))), title = "PSEN2 Non-Pathogenic")
  theme_bw() +
  xlim(40,90) +
  ylim(0,.5)+
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
  theme(axis.title.x = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.15,.90))+
```



```
theme(plot.margin = margin(2, 1, 1, 1, "cm"))
```

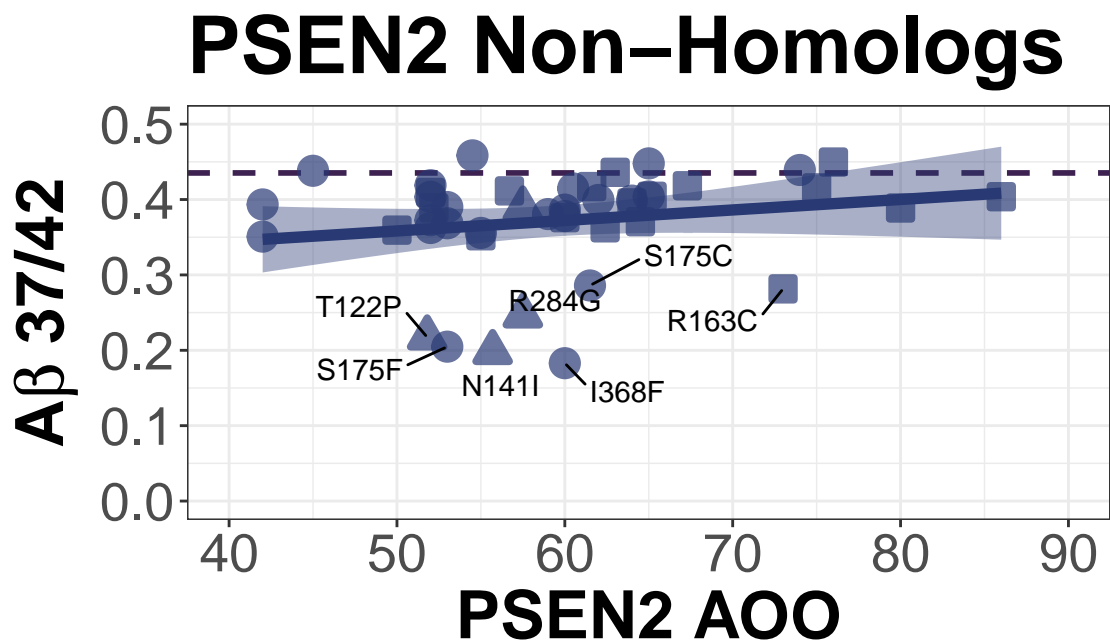
```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 8 rows containing non-finite outside the scale range
## ('stat_smooth()').
```

```
## Warning: The following aesthetics were dropped during statistical transformation: label.
## i This can happen when ggplot fails to infer the correct grouping structure in
## the data.
## i Did you forget to specify a 'group' aesthetic or to convert a numerical
## variable into a factor?
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_text_repel()').
```



```
data_merge.psen2.nonhomolog <- data_merge_rmWT.allpsen2 %>%
  filter(homolog == "0")
```

```
cor.test(data_merge.psen2.nonhomolog$A00, data_merge.psen2.nonhomolog$Abeta_3742)
```

```
##
## Pearson's product-moment correlation
##
## data: data_merge.psen2.nonhomolog$A00 and data_merge.psen2.nonhomolog$Abeta_3742
## t = 0.98887, df = 44, p-value = 0.3281
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.1492383 0.4197777
## sample estimates:
## cor
## 0.1474477
```

```
#
# data: data_merge.psen2.nonhomolog$A00 and data_merge.psen2.nonhomolog$Abeta_3742
# t = 0.98887, df = 44, p-value = 0.3281
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# -0.1492383 0.4197777
# sample estimates:
# cor
# 0.1474477
```

```
#### merge together ####
```

```
combined_fig2 <- plot_grid(fig2a, fig2b, fig2c, fig2d,fig2e,fig2f, nrow = 3, labels = "AUTO", label_size = 12)
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 8 rows containing non-finite outside the scale range
## ('stat_smooth()').
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 8 rows containing non-finite outside the scale range ('stat_smooth()').
## Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').
```

```
## 'geom_smooth()' using formula = 'y ~ x'
## 'geom_smooth()' using formula = 'y ~ x'
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 8 rows containing non-finite outside the scale range
## ('stat_smooth()').
```

```
## Warning: The following aesthetics were dropped during statistical transformation: label.
## i This can happen when ggplot fails to infer the correct grouping structure in
## the data.
## i Did you forget to specify a 'group' aesthetic or to convert a numerical
## variable into a factor?
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_text_repel()').
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 8 rows containing non-finite outside the scale range
## ('stat_smooth()').
```

```
## Warning: The following aesthetics were dropped during statistical transformation: label.
## i This can happen when ggplot fails to infer the correct grouping structure in
## the data.
## i Did you forget to specify a 'group' aesthetic or to convert a numerical
## variable into a factor?
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_point()').
```

```
## Warning: Removed 8 rows containing missing values or values outside the scale range
## ('geom_text_repel()').
```

```
dev.new()
pdf("~/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/PSEN1_2_homolog/02_figures/raw/Figure
combined_fig2
dev.off()
```

```
## pdf
## 2
```

```
# Figure 3A; boxplot of AOO by PSEN1 and PSEN2 groupings #
print(boxplot_AOO_PSEN <- data_merge_rmWT %>%
  filter(AOO > 0) %>%
  ggplot( aes(PSEN, AOO, fill = PSEN_homolog_cat, color = PSEN_homolog_cat), show.legend=F) +
    geom_line(aes(group=PSEN2_homolog.x), size=.5,color= "#992428", alpha = .5, show.legend=F) +
    geom_rain(alpha = .6,rain.side = 'f', show.legend=F,
      boxplot.args = list(color = "black", outlier.shape = NA),
      violin.args = list(color = "black", outlier.shape = NA, alpha = .5),
      point.args = list(aes(shape=AD_path2), size = 3, alpha = .7),
      violin.args.pos = list(width = 1, position = position_nudge(x = c(rep(-.5, 256*2), rep(.5, 256*2), rep(1.5, 256*2))),
      boxplot.args.pos = list( width = .3, position = position_nudge(x = c(-.3, .3,.3))))+
    ylim(0, 90) +
    scale_y_continuous(breaks=seq(20,90,20))+
    theme(axis.text.y.right = element_blank()) +
    labs( fill = "PSEN", color = "PSEN", shape = "ACMG",x = "")+
    scale_color_manual(values = c("#992428", "#283974","#ee9c45"), labels= c("PSEN1", "PSEN2 non-pathogenic"),
    scale_fill_manual(values = c("#992428", "#283974","#ee9c45"), labels= c("PSEN1", "PSEN2 non-pathogenic"),
    scale_shape_manual(values = c(17,15,16,18), labels = c("Pathogenic", "Benign/VUS", "Not Classified", "Unknown"))+
    theme_classic() +
    scale_x_discrete(labels=c('PSEN1', 'PSEN2'))+
```

```

theme(axis.title.y = element_text(margin = margin(t = 0, r = 10, b = 0, l = 0), size = 20, face = "bold"),
      legend.position = "right")+
theme(legend.title = element_text(size = 15, family = "Helvetica", face = "bold"))+
theme(legend.background = element_rect(fill = "transparent"), legend.key = element_rect(fill = "transparent", stroke = "black", stroke.width = 1))+
theme(legend.text = element_text(size = 15, family = "Helvetica", face = "bold"))+
theme(axis.text.y = element_text(size = 15, face = "bold", color = "black"))+
  theme(axis.text.x = element_text(hjust = c(.75, .25), size = 20, face = "bold", color = "black"))+
  theme(plot.margin = unit(c(0.5, 0.5, 0.5, 0.5), "inches"))

```

```

## Warning in (function (mapping = NULL, data = NULL, stat = "density", position
## = "dodge", : Ignoring unknown parameters: 'outlier.shape'

```

```

## Scale for y is already present.
## Adding another scale for y, which will replace the existing scale.

```

```

## Warning: Vectorized input to 'element_text()' is not officially supported.
## i Results may be unexpected or may change in future versions of ggplot2.

```

```

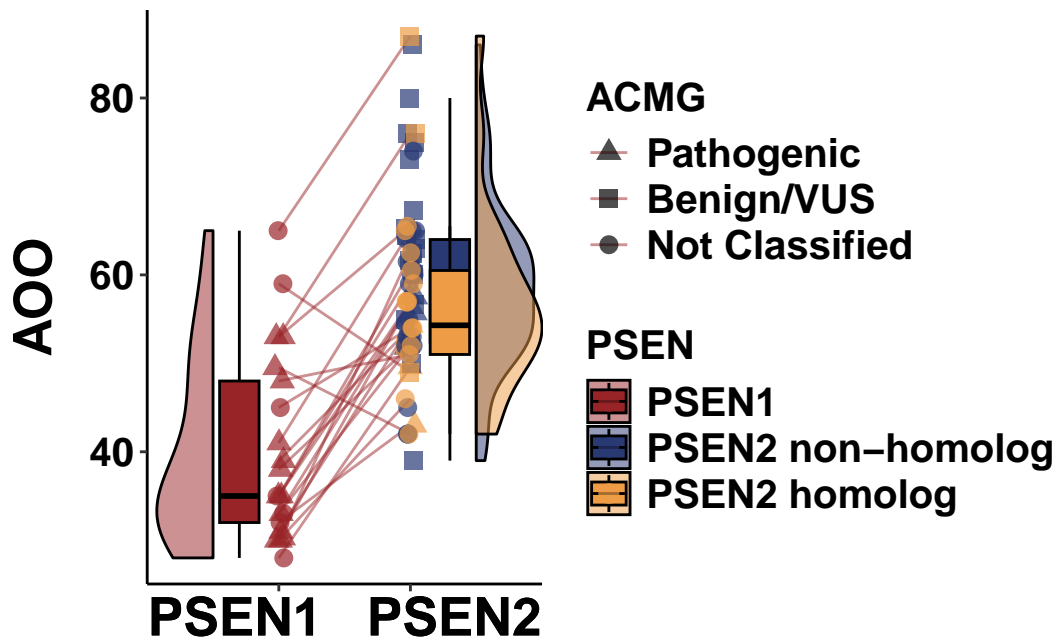
## Warning in setup_data(...): geom_paired_raincloud is only useful for visualizing groupings of length
##                               Check out packages {vioplot} and {see} for alternative ways of plotting split

```

```

## Warning: Using the 'size' aesthetic with geom_polygon was deprecated in ggplot2 3.4.0.
## i Please use the 'linewidth' aesthetic instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.

```



```
leg <- get_legend(boxplot_A00_PSEN)
```




```
## Warning in setup_data(...): geom_paired_raincloud is only useful for visualizing groupings of length
##                               Check out packages {vioplot} and {see} for alternative ways of plotting split v
```

```
print(leg <-as_ggplot(leg))
```

ACMG

- ▲ Pathogenic
- Benign/VUS
- Not Classified

PSEN

-  PSEN1
-  PSEN2 non-homolog
-  PSEN2 homolog

```
# comparing A00 between PSEN1 and PSEN2 groups #
t.test(data_merge_rmWT$A00~data_merge_rmWT$PSEN)
```

```
##
## Welch Two Sample t-test
##
## data: data_merge_rmWT$A00 by data_merge_rmWT$PSEN
## t = -6.9835, df = 31.715, p-value = 6.841e-08
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
## -23.64585 -12.96388
## sample estimates:
## mean in group 1 mean in group 2
## 40.13095 58.43582
```

```
#
# data: data_merge_rmWT$A00 by data_merge_rmWT$PSEN
# t = -6.9835, df = 31.715, p-value = 6.841e-08
# alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
# 95 percent confidence interval:
# -23.64585 -12.96388
# sample estimates:
# mean in group 1 mean in group 2
# 40.13095 58.43582

# comparing A00 between PSEN1 homologs and PSEN2 homologs #
data_merge_rmWT_homolog <- data_merge_rmWT %>%
  filter(homolog == "1")
t.test(data_merge_rmWT_homolog$A00~data_merge_rmWT_homolog$PSEN)
```

```
##
## Welch Two Sample t-test
##
## data: data_merge_rmWT_homolog$A00 by data_merge_rmWT_homolog$PSEN
## t = -5.1105, df = 39.995, p-value = 8.323e-06
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
## -23.29451 -10.09121
## sample estimates:
## mean in group 1 mean in group 2
## 40.13095 56.82381
```

```
#
# data: data_merge_rmWT_homolog$A00 by data_merge_rmWT_homolog$PSEN
# t = -5.1105, df = 39.995, p-value = 8.323e-06
# alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
# 95 percent confidence interval:
# -23.29451 -10.09121
# sample estimates:
# mean in group 1 mean in group 2
# 40.13095 56.82381

# Figure 3B, A00 in PSEN1 homologs vs A00 in PSEN2 homologs ##
data_merge_wide_rmWT$AD_path2.PSEN2 <- as.factor(data_merge_wide_rmWT$AD_path2.PSEN2)
data_merge_wide_rmWT_fig1d <- data_merge_wide_rmWT %>%
  filter(homolog.PSEN2 == "1")

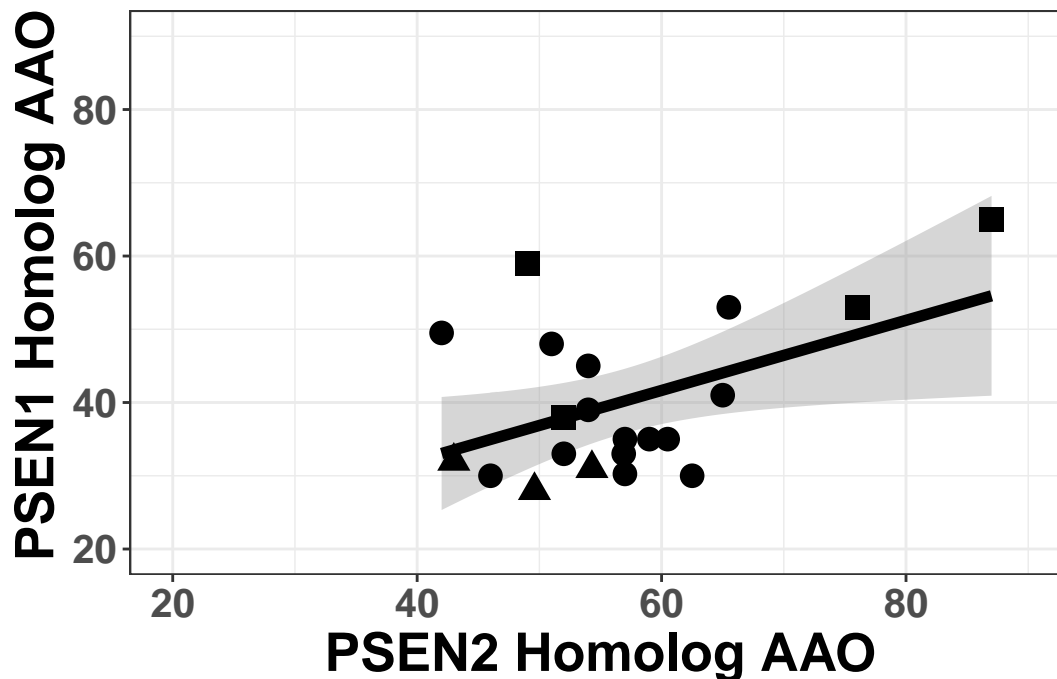
print(fig3b <- data_merge_wide_rmWT_fig1d %>%
  ggplot( aes(x=A00.PSEN2, y=A00.PSEN1)) +
  geom_smooth ( method = "lm", size = 2,se=TRUE, color="black")+
  scale_shape_manual(values = c(17,15,16), labels = c("Pathogenic", "Benign/VUS", "Not Classified")) +
  geom_point(size=4,show.legend=F, aes(shape =AD_path2.PSEN2 ), color ="black") +
  labs( x="PSEN2 Homolog AAO", y=" PSEN1 Homolog AAO",shape = "ACMG Classification")+
  theme_bw() +
  ylim(20,90) +
  xlim(20,90) +
```

```

theme(axis.text = element_text(size = 15, face="bold"))+
theme(legend.title=element_text(size = 15, face="bold")) +
theme(legend.text=element_text(size= 15, face = "bold",family="Helvetica")) +
theme(axis.title.y = element_text(size = 20, face="bold")) +
theme(axis.title.x = element_text(size = 20, face="bold")) +
theme(legend.position = c(0.2,.80))+
theme(plot.margin = unit(c(0.5, 0.5, 0.5, 0.5), "inches"))

```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```

# correlations between PSEN1 homolog AAO and PSEN2 homolog AAO #
cor.test(data_merge_wide_rmWT_fig1d$A00.PSEN2,data_merge_wide_rmWT_fig1d$A00.PSEN1)

```

```

##
## Pearson's product-moment correlation
##
## data: data_merge_wide_rmWT_fig1d$A00.PSEN2 and data_merge_wide_rmWT_fig1d$A00.PSEN1
## t = 2.3403, df = 19, p-value = 0.03033
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.05195341 0.75130129
## sample estimates:
## cor
## 0.4730312

```

```

#
# data: data_merge_wide_rmWT_fig1d$A00.PSEN2 and data_merge_wide_rmWT_fig1d$A00.PSEN1
# t = 2.3403, df = 19, p-value = 0.03033
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# 0.05195341 0.75130129
# sample estimates:
# cor
# 0.4730312

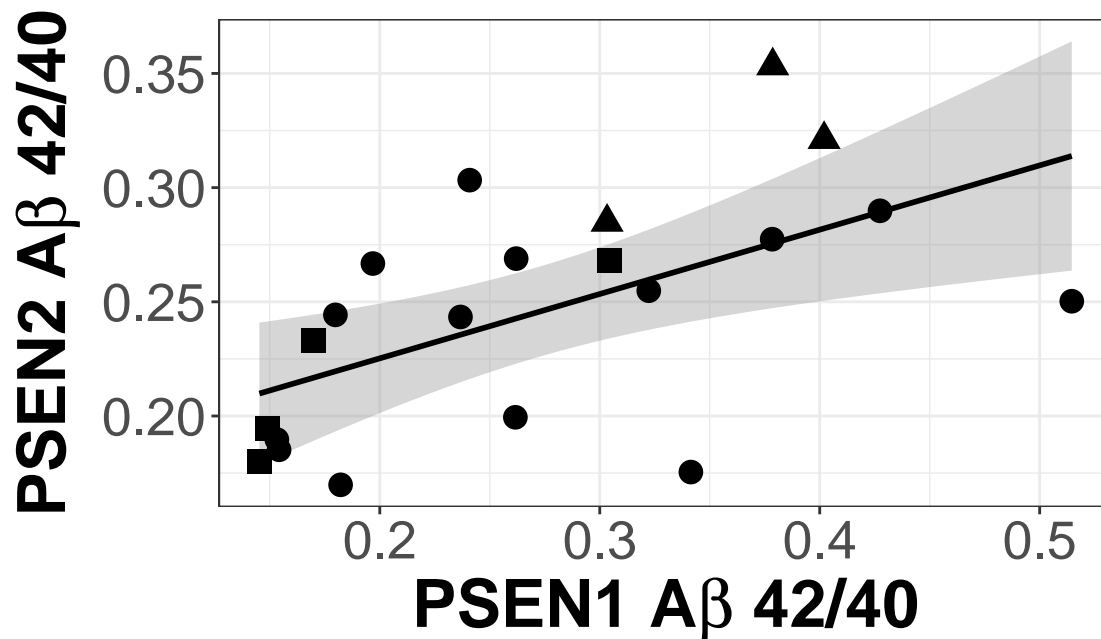
#### comparing ab4240 in PSEN1 homolog and PSEN2 homolog ####

homolog_only_data_merge_wide_rmWT <- data_merge_wide_rmWT %>% filter(data_merge_wide_rmWT$homolog.PSEN2

print(fig3e <- homolog_only_data_merge_wide_rmWT %>%
  ggplot( aes(x=Abeta_4240.PSEN1, y=Abeta_4240.PSEN2)) +
  geom_smooth ( method = "lm", size = 1,se=TRUE, color="black")+
  geom_point(size=4,show.legend=F, aes(shape = AD_path2.PSEN2)) +
  scale_shape_manual(values = c(17,15,16, 18), labels = c("AD Pathogenic", "AD unclear/benign", "non-AD
    #annotate("text", size = 8, x=.5, y=0.05, label= "r = 0.147, p = 0.328",color = "black")
  labs( x=expression(bold(paste("PSEN1 A", beta, " 42/40"))), y=expression(bold(paste("PSEN2 A", beta, " 42/40"))))
  theme_bw() +
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
  theme(axis.title.x = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.15,.90))+
  theme(plot.margin = margin(2, 1, 1, 1, "cm")))

## 'geom_smooth()' using formula = 'y ~ x'

```

```
cor.test(homolog_only_data_merge_wide_rmWT$Abeta_4240.PSEN1, homolog_only_data_merge_wide_rmWT$Abeta_4240.PSEN2)
```

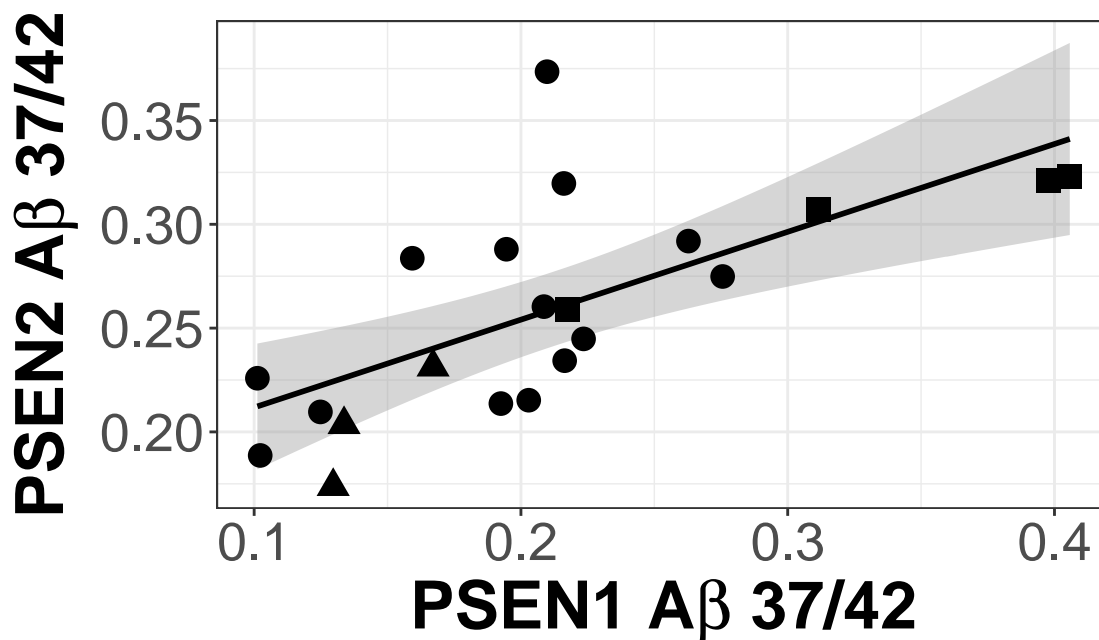
```
##
## Pearson's product-moment correlation
##
## data: homolog_only_data_merge_wide_rmWT$Abeta_4240.PSEN1 and homolog_only_data_merge_wide_rmWT$Abeta_4240.PSEN2
## t = 3.1029, df = 19, p-value = 0.005857
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1977430 0.8090658
## sample estimates:
##      cor
## 0.5799256
```

```
#
# data: homolog_only_data_merge_wide_rmWT$Abeta_4240.PSEN1 and homolog_only_data_merge_wide_rmWT$Abeta_4240.PSEN2
# t = 3.1029, df = 19, p-value = 0.005857
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
#  0.1977430 0.8090658
# sample estimates:
#      cor
# 0.5799256
```

```
#### comparing ab3742 in PSEN1 homolog and PSEN2 homolog ####
```

```
print(fig3f <- homolog_only_data_merge_wide_rmWT %>%
  ggplot( aes(x=Abeta_3742.PSEN1, y=Abeta_3742.PSEN2)) +
  geom_smooth ( method = "lm", size = 1,se=TRUE, color="black")+
  geom_point(size=4,show.legend=F, aes(shape = AD_path2.PSEN2)) +
  scale_shape_manual(values = c(17,15,16, 18), labels = c("AD Pathogenic", "AD unclear/benign", "non-AD", "AD Pathogenic"))
  #annotate("text", size = 8, x=.5, y=0.05, label= "r = 0.147, p = 0.328",color = "black")
labs( x=expression(bold(paste("PSEN1 A", beta, " 37/42"))), y=expression(bold(paste("PSEN2 A", beta, " 37/42"))))
  theme_bw() +
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
  theme(axis.title.x = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.15,.90))+
  theme(plot.margin = margin(2, 1, 1, 1, "cm")))
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```

cor.test(homolog_only_data_merge_wide_rmWT$Abeta_3742.PSEN1, homolog_only_data_merge_wide_rmWT$Abeta_3742.PSEN2)

##
## Pearson's product-moment correlation
##
## data: homolog_only_data_merge_wide_rmWT$Abeta_3742.PSEN1 and homolog_only_data_merge_wide_rmWT$Abeta_3742.PSEN2
## t = 4.0207, df = 19, p-value = 0.0007308
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.3482621 0.8584442
## sample estimates:
## cor
## 0.6780159

#
# data: homolog_only_data_merge_wide_rmWT$Abeta_3742.PSEN1 and homolog_only_data_merge_wide_rmWT$Abeta_3742.PSEN2
# t = 4.0207, df = 19, p-value = 0.0007308
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval:
# 0.3482621 0.8584442
# sample estimates:
# cor
# 0.6780159

#### comparing A00 vs ab4240 in PSEN1 homologs and PSEN2 homologs ####

data_merge_rmWT$AD_path2 <- as.character(data_merge_rmWT$AD_path2)

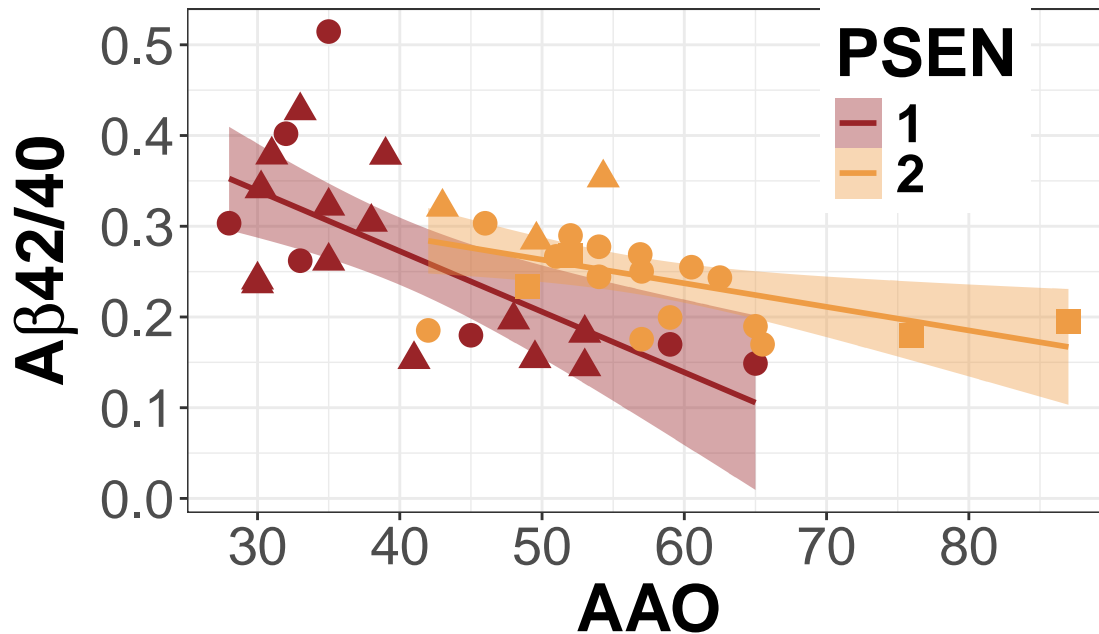
data_merge_rmWT$AD_path2 <- data_merge_rmWT$AD_path2 %>% replace_na("4")

print(fig3c <- data_merge_rmWT %>% filter (homolog == "1") %>%
  ggplot( aes(x=A00, y=Abeta_4240)) +
  geom_smooth ( method = "lm", size = 1,se=TRUE, aes(color= PSEN , fill = PSEN))+
  geom_point(size=4,show.legend=F, aes(color =PSEN, shape = AD_path2)) +
  #annotate("text", size = 8, x=.5, y=0.05, label= "r = 0.147, p = 0.328",color = "black",
  scale_fill_manual(values = c("#992428","#ee9c45"), labels=labels)+
  scale_color_manual(values = c("#992428","#ee9c45"), labels=labels)+
  #scale_shape_manual(values = c(17,16), labels=labels)+ #Titles have to be done for shapes and colors
  scale_shape_manual(values = c(17,15,16, 18), labels = c("AD Pathogenic", "AD unclear/uncertain", "AD non-pathogenic", "AD unknown"))+

  labs( x="A00", y = expression(bold(paste("A", beta, "42/40"))))+
  theme_bw() +
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold",family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
  theme(legend.position = c(0.8,.80))+
  theme(plot.margin = margin(2, 1, 1, 1, "cm"))+
  theme(axis.title.x = element_text(size = 25, face="bold")))

```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
data_merge_rmWT_homolog <- data_merge_rmWT %>% filter (homolog == "1")

summary(lm(Abeta_4240~A00*PSEN,data_merge_rmWT_homolog))
```

```
##
## Call:
## lm(formula = Abeta_4240 ~ A00 * PSEN, data = data_merge_rmWT_homolog)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.112612 -0.043278  0.005749  0.026983  0.208725
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.539527   0.057160   9.439 1.66e-11 ***
## A00          -0.006676   0.001379  -4.842 2.18e-05 ***
## PSEN2        -0.146373   0.098754  -1.482  0.1465
## A00:PSEN2     0.004076   0.001961   2.078  0.0445 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06564 on 38 degrees of freedom
## Multiple R-squared:  0.4293, Adjusted R-squared:  0.3843
```

```
## F-statistic: 9.53 on 3 and 38 DF, p-value: 7.971e-05
```

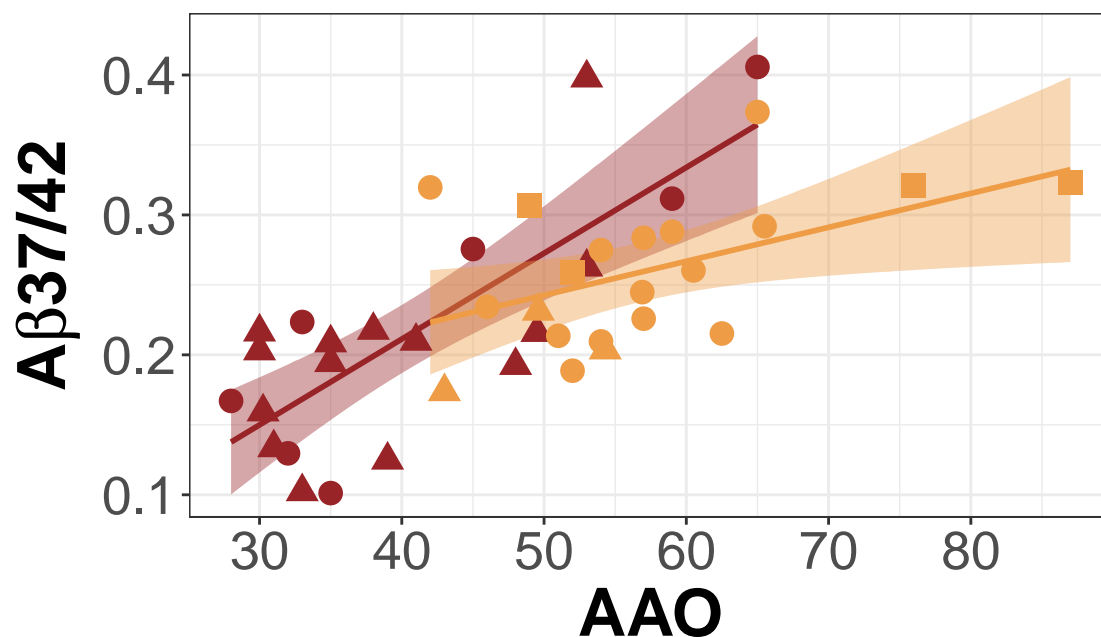
```
# Call:
# lm(formula = Abeta_4240 ~ A00 * PSEN, data = data_merge_rmWT_homolog)
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -0.112612 -0.043278  0.005749  0.026983  0.208725
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept)  0.539527   0.057160   9.439 1.66e-11 ***
# A00          -0.006676   0.001379  -4.842 2.18e-05 ***
# PSEN2        -0.146373   0.098754  -1.482  0.1465
# A00:PSEN2     0.004076   0.001961   2.078  0.0445 *
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# Residual standard error: 0.06564 on 38 degrees of freedom
# Multiple R-squared:  0.4293, Adjusted R-squared:  0.3843
# F-statistic: 9.53 on 3 and 38 DF, p-value: 7.971e-05
```

```
#### comparing A00 vs ab3742 in PSEN1 homologs and PSEN2 homologs ####
```

```
print(fig3d <- data_merge_rmWT %>% filter (homolog == "1") %>%
  ggplot( aes(x=A00, y=Abeta_3742)) +
  geom_smooth ( method = "lm", size = 1, se=TRUE, aes(color= PSEN, fill = PSEN ), show.legend=F)+
  geom_point(size=4, show.legend=F, aes(color =PSEN, shape = AD_path2)) +
  #annotate("text", size = 8, x=.5, y=0.05, label= "r = 0.147, p = 0.328", color = "black",
  #         scale_fill_manual(values = c("#992428", "#ee9c45"), labels=labels)+
  scale_color_manual(values = c("#992428", "#ee9c45"), labels=labels)+
  scale_shape_manual(values = c(17,15,16, 18), labels = c("AD Pathogenic", "AD unclear/
  labs( x="A00", y = expression(bold(paste("A", beta, "37/42"))))+
  theme_bw() +
  theme(axis.text = element_text(size = 20))+
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold", family="Helvetica")) +
  theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
  theme(plot.margin = margin(2, 1, 1, 1, "cm"))+

  theme(axis.title.x = element_text(size = 25, face="bold")))
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
summary(lm(Abeta_3742~A00*PSEN,data_merge_rmWT_homolog))
```

```
##
## Call:
## lm(formula = Abeta_3742 ~ A00 * PSEN, data = data_merge_rmWT_homolog)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.080307 -0.033466 -0.003125  0.027016  0.106920
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.034066   0.043592  -0.781   0.4394
## A00           0.006133   0.001052   5.833 9.66e-07 ***
## PSEN2        0.155457   0.075313   2.064  0.0459 *
## A00:PSEN2    -0.003709   0.001496  -2.480  0.0177 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05006 on 38 degrees of freedom
## Multiple R-squared:  0.5608, Adjusted R-squared:  0.5261
## F-statistic: 16.17 on 3 and 38 DF,  p-value: 6.223e-07
```

```
# Call:
# lm(formula = Abeta_3742 ~ A00 * PSEN, data = data_merge_rmWT_homolog)
```

```

#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -0.080307 -0.033466 -0.003125  0.027016  0.106920
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept) -0.034066   0.043592  -0.781   0.4394
# APO         0.006133   0.001052   5.833 9.66e-07 ***
# PSEN2       0.155457   0.075313   2.064  0.0459 *
# APO:PSEN2   -0.003709   0.001496  -2.480  0.0177 *
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# Residual standard error: 0.05006 on 38 degrees of freedom
# Multiple R-squared:  0.5608, Adjusted R-squared:  0.5261
# F-statistic: 16.17 on 3 and 38 DF,  p-value: 6.223e-07

data_merge_rmWT_homolog.psen1 <- data_merge_rmWT_homolog %>% filter(PSEN== "1")
data_merge_rmWT_homolog.psen2 <- data_merge_rmWT_homolog %>% filter(PSEN== "2")

mean(data_merge_rmWT_homolog.psen1$Abeta_3742) #0.212076

## [1] 0.212076

mean(data_merge_rmWT_homolog.psen2$Abeta_3742) #0.2591818

## [1] 0.2591818

mean(data_merge_rmWT_homolog.psen1$Abeta_4240) # 0.2716151

## [1] 0.2716151

mean(data_merge_rmWT_homolog.psen2$Abeta_4240) # 0.2454094

## [1] 0.2454094

t.test(data_merge_rmWT_homolog$Abeta_4240 ~ data_merge_rmWT_homolog$PSEN)

##
## Welch Two Sample t-test
##
## data:  data_merge_rmWT_homolog$Abeta_4240 by data_merge_rmWT_homolog$PSEN
## t = 1.0155, df = 28.939, p-value = 0.3183
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
##  -0.02657659  0.07898794
## sample estimates:
## mean in group 1 mean in group 2
##      0.2716151      0.2454094

```

```
#
# data: data_merge_rmWT_homolog$Abeta_4240 by data_merge_rmWT_homolog$PSEN
# t = 1.0155, df = 28.939, p-value = 0.3183
# alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
# 95 percent confidence interval:
# -0.02657659 0.07898794
# sample estimates:
# mean in group 1 mean in group 2
# 0.2716151 0.2454094

t.test(data_merge_rmWT_homolog$Abeta_3742 ~ data_merge_rmWT_homolog$PSEN)
```

```
##
## Welch Two Sample t-test
##
## data: data_merge_rmWT_homolog$Abeta_3742 by data_merge_rmWT_homolog$PSEN
## t = -2.1947, df = 33.521, p-value = 0.03522
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
## -0.090747903 -0.003463845
## sample estimates:
## mean in group 1 mean in group 2
## 0.2120760 0.2591818
```

```
#
# data: data_merge_rmWT_homolog$Abeta_3742 by data_merge_rmWT_homolog$PSEN
# t = -2.1947, df = 33.521, p-value = 0.03522
# alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
# 95 percent confidence interval:
# -0.090747903 -0.003463845
# sample estimates:
# mean in group 1 mean in group 2
# 0.2120760 0.2591818
```

```
data_merge_rmWT_homolog$Abeta_GSC <- (data_merge_rmWT_homolog$Abeta_37 + data_merge_rmWT_homolog$Abeta_42)
```

```
#### Fig 2x ####
```

```
print(fig2x <- data_merge_rmWT_homolog %>%
  ggplot( aes(x=A00, y=Abeta_GSC)) +
  geom_smooth ( method = "lm", size = 1, se=TRUE, aes(color= PSEN, fill = PSEN ), show.legend=F) +
  geom_point(size=4, show.legend=F, aes(color =PSEN, shape = AD_path2)) +
  #annotate("text", size = 8, x=.5, y=0.05, label= "r = 0.147, p = 0.328", color = "black") +
  scale_fill_manual(values = c("#992428", "#ee9c45"), labels=labels) +
  scale_color_manual(values = c("#992428", "#ee9c45"), labels=labels) +
  scale_shape_manual(values = c(17,15,16, 18), labels = c("AD Pathogenic", "AD unclear/")) +
  labs( x="A00", y = expression(bold(paste("A", beta, "37+40/42")))) +
  theme_bw() +
  theme(axis.text = element_text(size = 20)) +
  theme(legend.title=element_text(size = 25, face="bold")) +
  theme(legend.text=element_text(size= 20, face = "bold", family="Helvetica")) +
```



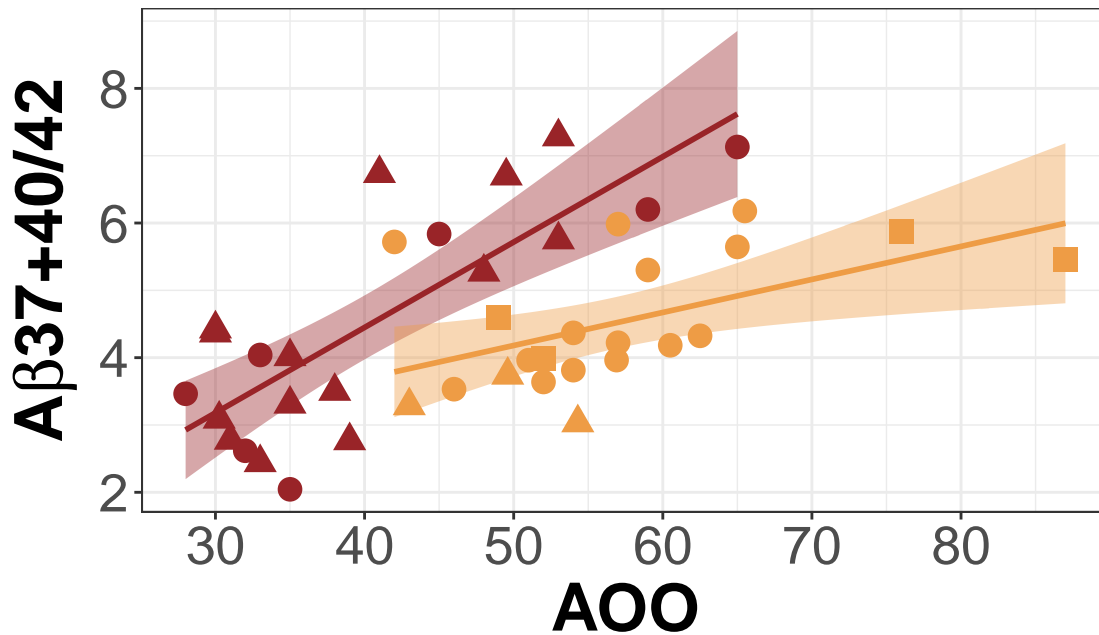
```

theme(axis.title.y = element_text(size = 25, face="bold")) +
  theme(title = element_text(size = 25, face="bold")) +
  theme(plot.margin = margin(2, 1, 1, 1, "cm"))+

  theme(axis.title.x = element_text(size = 25, face="bold"))

```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
summary(lm(Abeta_GSC~A00*PSEN,data_merge_rmWT_homolog))
```

```
##
## Call:
## lm(formula = Abeta_GSC ~ A00 * PSEN, data = data_merge_rmWT_homolog)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7701  -0.5465  -0.2952   0.6442   2.1616
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.62468    0.82316  -0.759   0.453
## A00           0.12684    0.01986   6.387 1.68e-07 ***
## PSEN2         2.35189    1.42217   1.654   0.106
## A00:PSEN2     -0.07777    0.02824  -2.753   0.009 **

```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9453 on 38 degrees of freedom
## Multiple R-squared:  0.5519, Adjusted R-squared:  0.5165
## F-statistic: 15.6 on 3 and 38 DF,  p-value: 9.043e-07

#
# Call:
# lm(formula = Abeta_GSC ~ APO * PSEN, data = data_merge_rmWT_homolog)
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -1.7701 -0.5465 -0.2952  0.6442  2.1616
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept) -0.62468     0.82316  -0.759   0.453
# APO          0.12684     0.01986   6.387 1.68e-07 ***
# PSEN2        2.35189     1.42217   1.654   0.106
# APO:PSEN2    -0.07777     0.02824  -2.753   0.009 **
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# Residual standard error: 0.9453 on 38 degrees of freedom
# Multiple R-squared:  0.5519, Adjusted R-squared:  0.5165
# F-statistic: 15.6 on 3 and 38 DF,  p-value: 9.043e-07

#combined_fig2ad <- plot_grid(fig2a, fig2b, fig2c,fig2d, nrow =2)

#dev.new()
#pdf("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/combined_fig2ad.pdf", wi
#combined_fig2ad
#dev.off()

#
#
#              combined_fig2ad <- plot_grid( fig2c,fig2d,fig2a, fig2b, nrow =2, labels = "AUTO", lab
#
# dev.new()
# pdf("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/combined_fig2ad051924.p
# combined_fig2ad
# dev.off()

#### Fig 2e ####

#
#
# combined_figAB <- plot_grid( fig_1c,fig1d, nrow = 1 ,labels = c("A", "B"), rel_widths = c(1.2,.9), la
#
#
# dev.new()
# pdf("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/combined_fig3ab_051624.p
# combined_figAB

```

```

# dev.off()
#
#

combined_fig3 <- plot_grid( "", "", fig3c, fig3d, fig3e, fig3f, nrow = 3, labels = c("", "", "C", "D")

## Warning in as_grob.default(plot): Cannot convert object of class character into
## a grob.

## Warning in as_grob.default(plot): Cannot convert object of class character into
## a grob.

## 'geom_smooth()' using formula = 'y ~ x'
## 'geom_smooth()' using formula = 'y ~ x'
## 'geom_smooth()' using formula = 'y ~ x'
## 'geom_smooth()' using formula = 'y ~ x'

#

dev.new()
pdf("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/fig3cf_061924.pdf", width=
combined_fig3
dev.off()

## pdf
## 2

# dev.new()
# pdf("/Users/sws59/Dropbox (Partners HealthCare)/Stephanie (1)/PSEN1_2_homolog/combined_fig3_051624.pdf", width=
# combined_fig3
# dev.off()

```