# Objective:

Compare the median expression levels of PD-L1 and PD-1 and identify the relevant oncological disease cohorts for their solid tumor immunotherapy.

# Code:

```
library(FirebrowseR)
library(ggplot2)

#TCGA_cohort = Metadata.Cohorts(format="csv"","cohort="")$cohort
#Problem in retrieving data for FPPP cohort using FirebrowseR, so ignored

Median=""

TCGA_cohort=c("ACC","BLCA","BRCA","CESC","CHOL","COAD","COADREAD","DLBC","ESCA","GBM","GBMLGG","HNSC","KICH","KIPAN","KIRC",
"KIRP","LAML","LGG","LIHC","LUAD","LUSC","MESO","OV","PAAD","PCPG","PRAD","READ","SARC","SKCM","STAD","STES","TGCT","THCA",
"THYM","UCEC","UCS","UVM")
for(x in (TCGA_cohort)){
    PD1 = Analyses.mRNASeq.Quartiles(format = "csv", gene = "PDCD1", cohort = x, protocol = "RSEM", sample_type = "tumors",
 Exclude = "")
    PDL1 = Analyses.mRNASeq.Quartiles(format = "csv", gene = "CD274", cohort = x, protocol = "RSEM", sample_type = "tumors",
Exclude = "")
    Median[x]= paste (PD1$median, '\t', PDL1$median)
                            }
write.table(Median,"Temp1.tsv", sep="\t",row.names=TRUE, col.names = FALSE, quote = FALSE)
mydata=read.table("Temp1.tsv",sep="\t")
unlink("Temp1.tsv")

lm_eqn <- function(m){
    eq <- substitute(italic(y) == a + b %.% italic(x)*","~~italic(pval)~"="~p*","~~italic(r)^2~"="~r2,
                    list(a = format(coef(m)[1], digits = 2),
                         b = format(coef(m)[2], digits = 2),
                         p = format(summary(fit)$coefficients[2,4], digits = 2),
                         r2 = format(summary(m)$r.squared, digits = 2)))
    as.character(as.expression(eq));
}

fit <- lm(X.1 ~ X, data = mydata)
MyPlot=ggplot(mydata, aes(x=X, y=X.1)) +
    labs(x = "PD-1 (median log2 expression)", y ="PD-L1 (median log2 expression)")+
    geom_point(size=2.5, shape=19,color = "yellowgreen")+
    geom_text(label=rownames(mydata),fontface="bold",size = 3.0,color="orange",hjust=0.25, vjust=1.5)+
    geom_smooth(method="lm", se=TRUE, level=0.5)+
    geom_text(x=7.5,y=2.0,size = 3.0, label = lm_eqn(fit), parse = TRUE)

#MyPlot
```
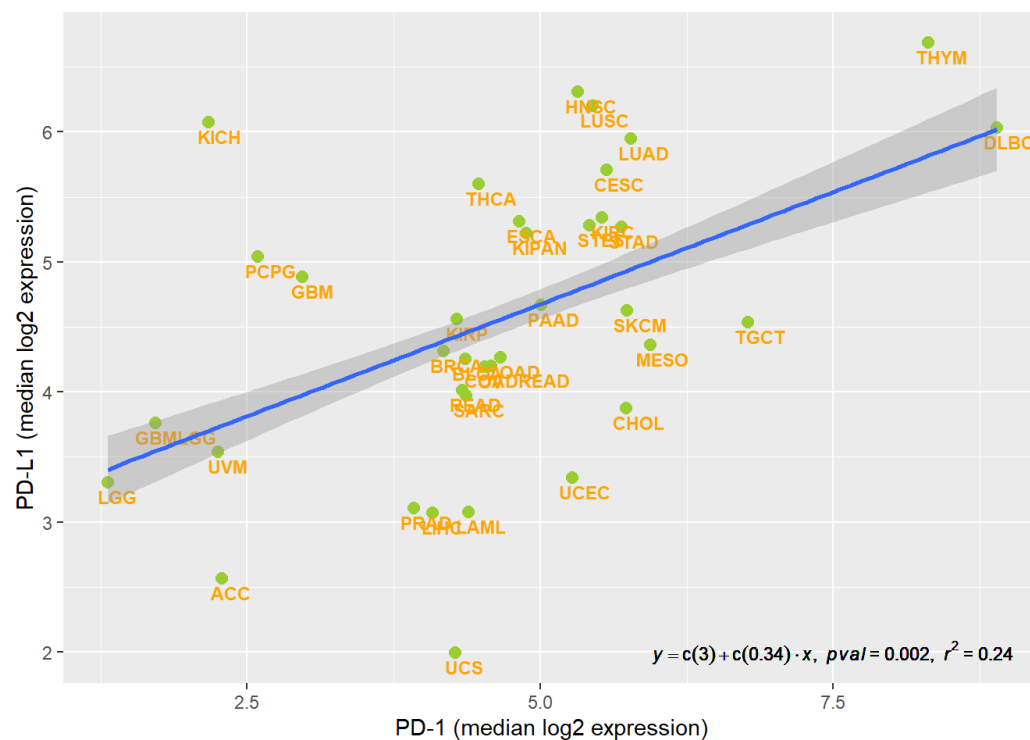
# Scatterplot comparing median RSEM normalized expression per cohort of PD-1 vs PD-L1



## Notes:

Alias=> PD-1:PDCD1|CD279 & PD-L1:CD274

The median RSEM normalized expression of PD-1 and PD-L1 per cohort, sorted by descending combined expression were stored in file **Median_Expression.tsv**

## Summary:

Programmed death-ligand 1 (PD-L1) is a receptor protein that generally over expressed on tumor cells or on non-transformed cells in tumor microenvironment. PD-L1 binds to PD-1 receptors of T cells and significantly supresses T-cells' anti-tumor activity.(Nature Reviews Cancer, 12, 2012).

This work is summarising the median log2 expression for genes; PD-1 and PD-L1 per cancer cohort (from TCGA) as illustrated in the scatterplot. These two genes exhibit a significant association (p=0.002, r^2=0.24) across all reported cancers, further suggesting their consequential role in promoting cancer.The most relevant oncological disease cohorts for their solid tumor immunotherapy based on current data analysis would be DLBC, PAAD and BRCA as they perfectly aligns with the predicted relationship between PD-1 and PD-L1.

## END