

Positive selection on genes interacting with SARS-Cov2, comparison of different analysis, maic

Marie Cariou

March 2021

Contents

| | | |
|----------|--------------------|----------|
| 1 | Data | 2 |
| 2 | MAIC | 2 |
| 2.1 | Boxplot | 2 |
| 2.2 | Dotchart | 4 |
| 3 | Tanglegram | 8 |

1 Data

output from covid_comp_dataset.

```
tablo<-read.table("primatesVbats.csv",  
                  h=T, sep=",")
```

Output MAIC formatted by Léa. This table includes the DGINN "score".

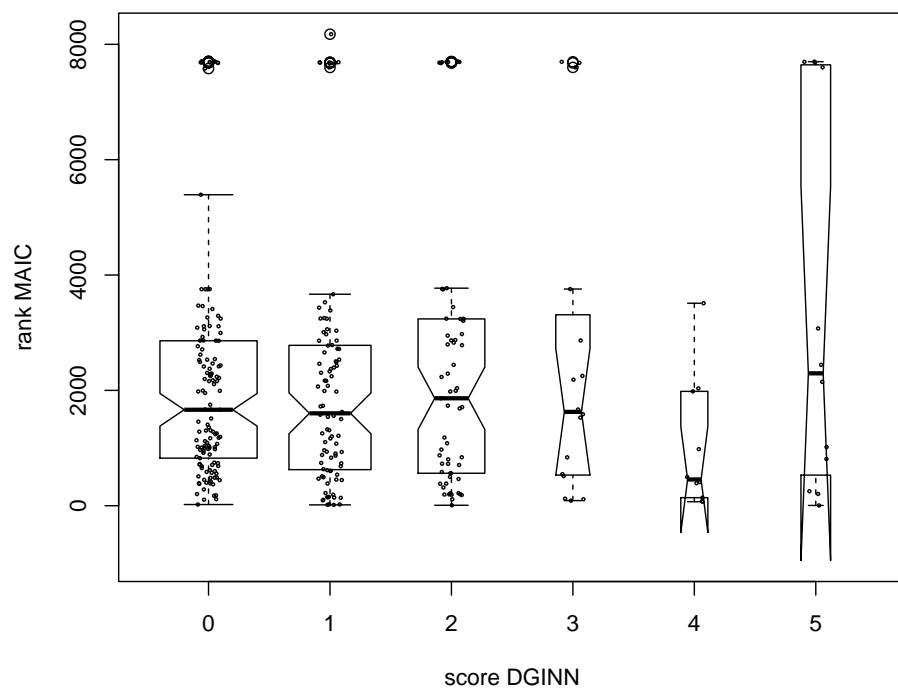
```
home<-"/home/adminmarie/Documents/"  
workdir<-paste0(home, "CIRI_BIBS_projects/2020_05_Etienne_covid/")  
maic<-read.table(paste0(workdir, "data/covid_comp_maic.txt"),  
                 h=T)
```

2 MAIC

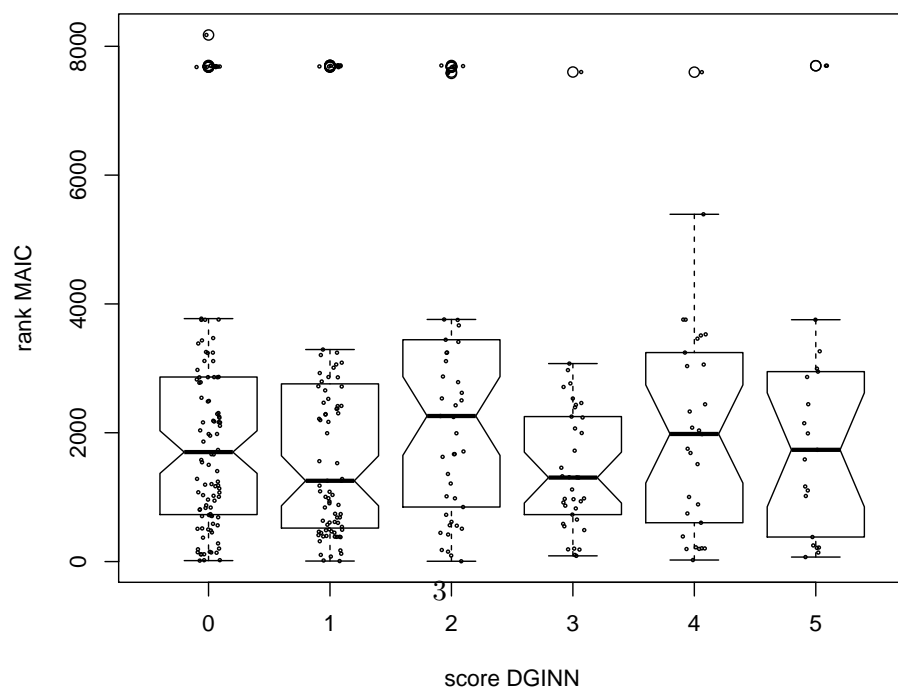
2.1 Boxplot

```
par(mfrow=c(2,1))  
boxplot(maic$rank~maic$nbats, notch=TRUE, varwidth=TRUE, xlab="score DGINN", ylab="r  
  
## Warning in bxp(list(stats = structure(c(21, 825, 1664, 2860,  
5392, 15, 625.5, : some notches went outside hinges ('box')): maybe  
set notch=FALSE  
  
stripchart(maic$rank~maic$nbats, method="jitter", vertical=TRUE, pch=1, cex=0.3, add=  
  
boxplot(maic$rank~maic$nprimates, notch=TRUE, xlab="score DGINN", ylab="rank MAIC",  
stripchart(maic$rank~maic$nprimates, method="jitter", vertical=TRUE, pch=1, cex=0.3,
```

Bats



Primates

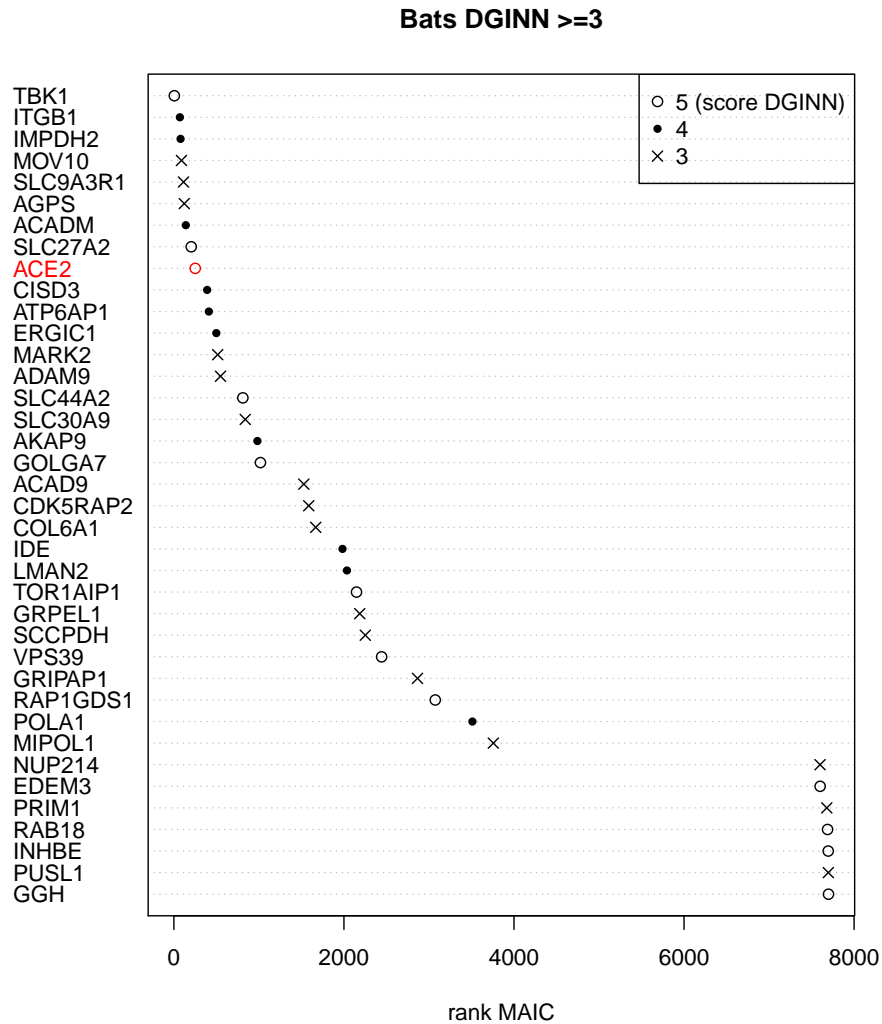


2.2 Dotchart

```
tmp<-maic[maic$nbats>=3, c("gene", "rank", "nbats")]
tmp<-tmp[order(tmp$rank, decreasing = TRUE),]
tmp$col<-"black"
tmp$col[tmp$gene=="ACE2"]<-"red"
tmp$col[tmp$gene=="TPRSS2"]<-"red"

tmp$pch[tmp$nbats==5]<-1
tmp$pch[tmp$nbats==4]<-20
tmp$pch[tmp$nbats==3]<-4

dotchart(tmp$rank, main="Bats DGINN >=3", xlab="rank MAIC", labels=tmp$gene, pch=tmp$pch)
legend("topright", c("5 (score DGINN)", "4", "3"), pch=c(1,20,4))
```



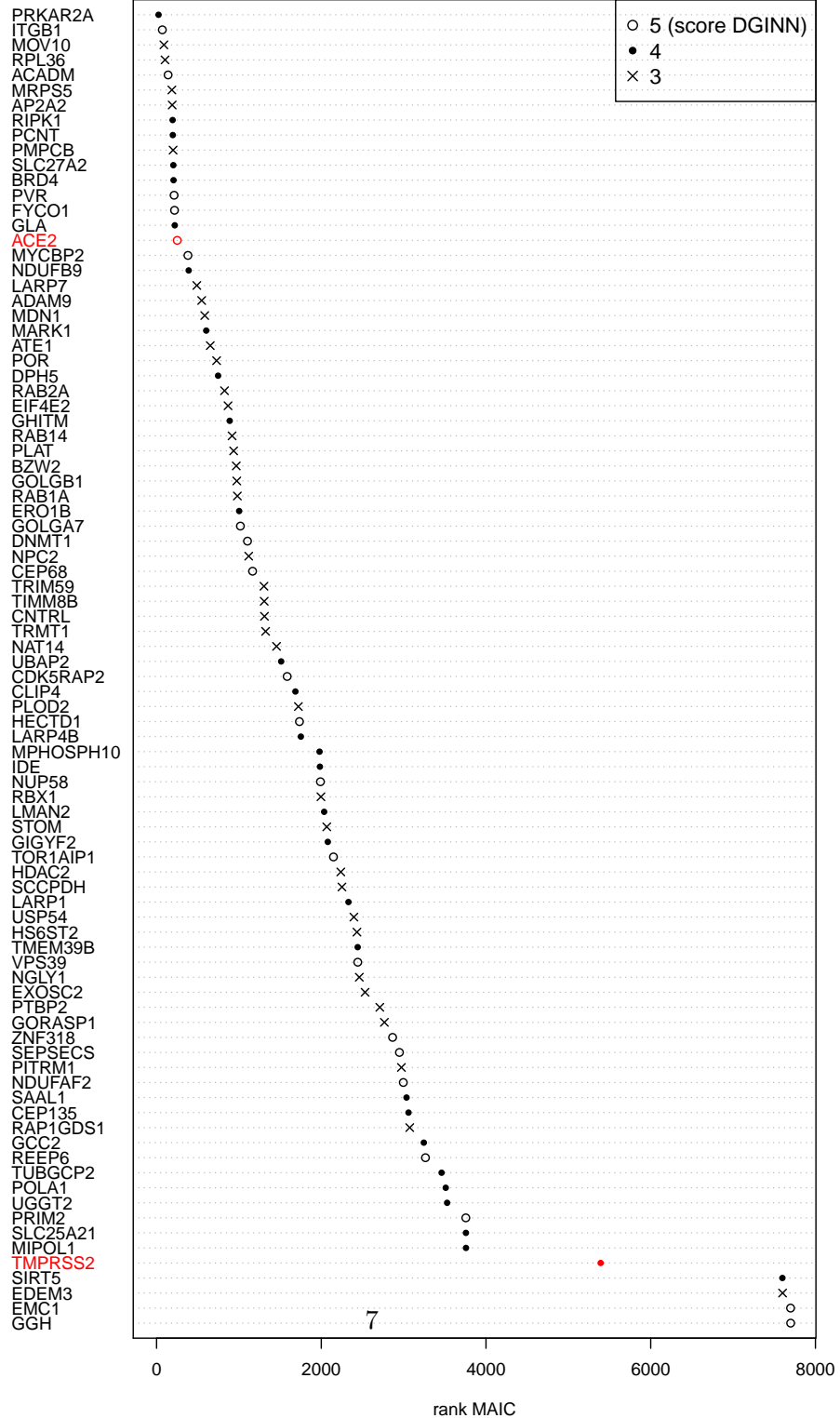
```
tmp<-maic[maic$nprimates>=3, c("gene", "rank", "nprimates")]
tmp<-tmp[order(tmp$rank, decreasing = TRUE),]

tmp$pch[tmp$nprimates==5]<-1
tmp$pch[tmp$nprimates==4]<-20
tmp$pch[tmp$nprimates==3]<-4
tmp$col<-"black"
tmp$col[tmp$gene=="ACE2"]<-"red"
```

```
tmp$col[tmp$gene=="TPRSS2"]<-"red"

dotchart(tmp$rank, main="Primates DGINN >=3", xlab="rank MAIC", labels=tmp$gene, pch=
legend("topright", c("5 (score DGINN)", "4", "3"), pch=c(1,20,4))
```

Primates DGINN >=3



3 Tanglegram

```
#install.packages('dendextend') # stable CRAN version
library(dendextend) # load the package
#install.packages("phytools") # stable CRAN version
library(phytools) # load the package
library(ggraph)
library(igraph)
library(tidyverse)

##
tmp<-tablo[(tablo$nbats!=0 | tablo$nprimates!=0),]
tmp<-head(tablo, 20)
#tmp<-rbind(as.matrix(tmp), c("outgroup", 50, 50))
tmp<-as.data.frame(tmp)
matbats<-hclust(dist(tmp$nbats))
matpri<-hclust(dist(tmp$nprimates))

tmp[order(tmp$nbats),]

dendpri<-as.dendrogram(matpri)
dendbats<-as.dendrogram(matbats)

labels(dendpri)<-as.character(tmp$tmp$Gene.name[labels(dendpri)])
labels(dendbats)<-as.character(tmp$tmp$Gene.name[labels(dendbats)])

tmp[order(tmp$nprimates, decreasing=FALSE),]$'tmp$Gene.name'> order
dendpri<-dendextend::rotate(dendpri, order=order)

tmp[order(tmp$nbats, decreasing=FALSE),]$'tmp$Gene.name'> order
dendbats<-dendextend::rotate(dendbats, order=order)

#### Il faut swapper certains neuds de l'arbres

class(labels(dendpri))
```



```

dend12 <- dendlist(dendbats, dendpri)

?png
png("tanglegram.png", width = 1800, height = 3000)
tanglegram(dend12, columns_width=c(3, 3,3), axes=FALSE,
            edge.lwd=0, margin_inner=6,
            margin_top=2,
            main_left="bats",
            main_right = "primates",
            lwd=0.5,
            cex_main=1,
            lab.cex=1,
            k_labels=6)

dev.off()

tmp
?tanglegram

```