

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	Pan Corona	8

1 Data

```
home<-" /home/adminmarie/Documents/CIRI_BIBS_projects/"
workdir<-paste0(home, "2020_05_Etienne_covid/2020_dginn_covid19/")
```

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

```
maic<-read.table(paste0(workdir, "data/covid_comp_maic.txt"),
                 h=T)
```

output from covid_comp_dataset.

```
tablo<-read.table(paste0(workdir, "out_tab/primatesVbats.csv"),
                  h=T, sep=",")
```

2 MAIC

2.1 Boxplot

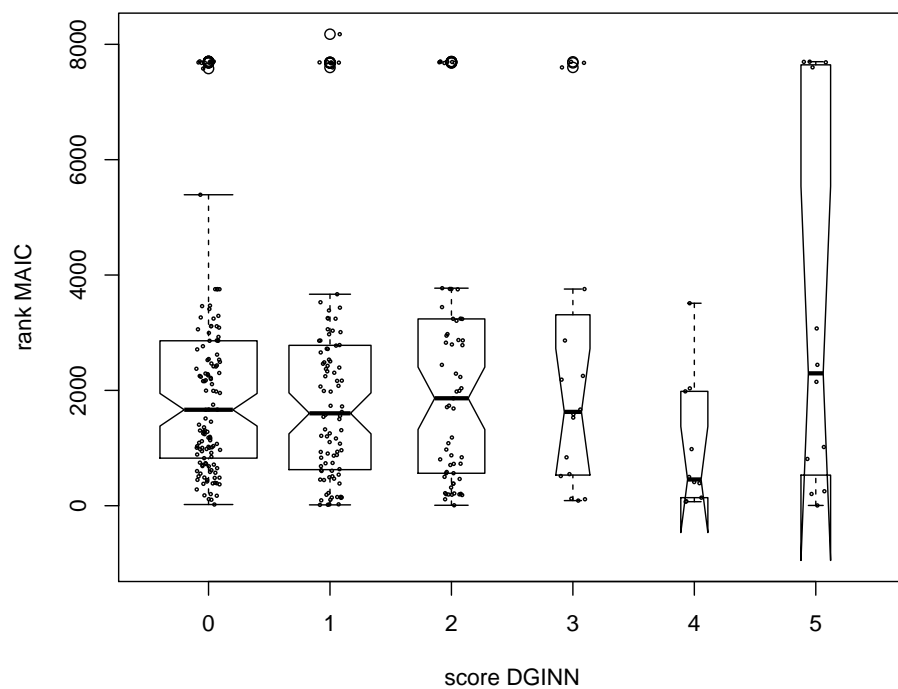
```
par(mfrow=c(2,1))
boxplot(maic$rank~maic$nbats, notch=TRUE, varwidth=TRUE, xlab="score DGINN", ylab="r

## Warning in bxp(structure(list(stats = structure(c(21, 825, 1664,
2860, 5392, : some notches went outside hinges ('box'): maybe
set notch=FALSE

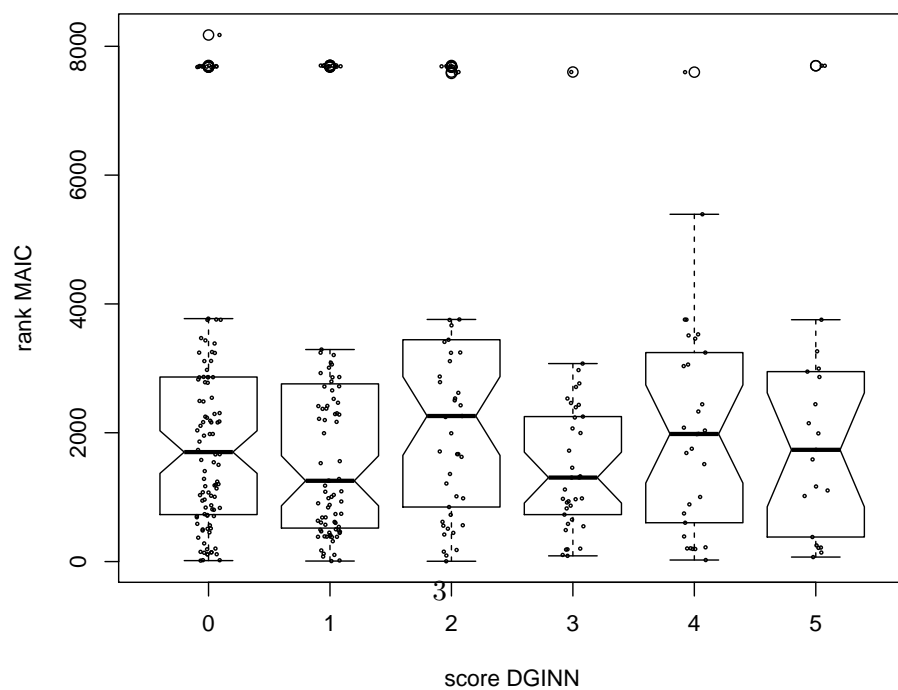
stripchart(maic$rank~maic$nbats, method="jitter", vertical=TRUE, pch=1, cex=0.3, add=

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

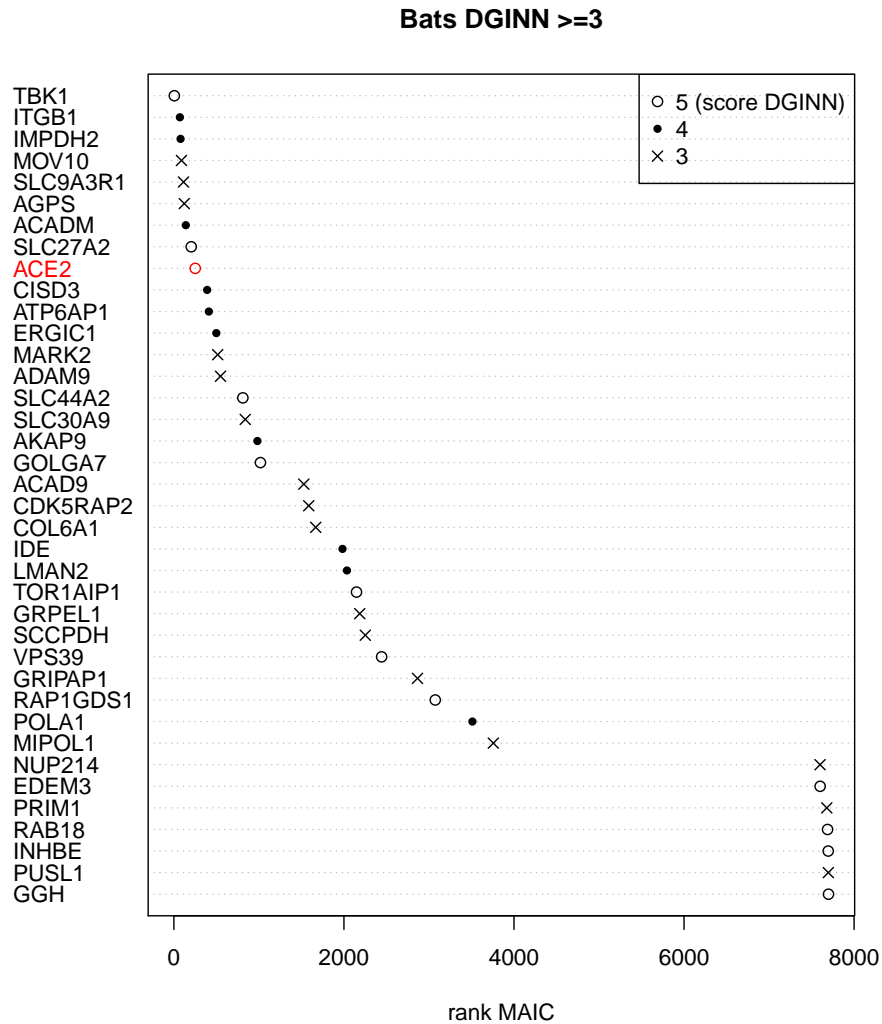


2.2 Dotchart

```
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))
```

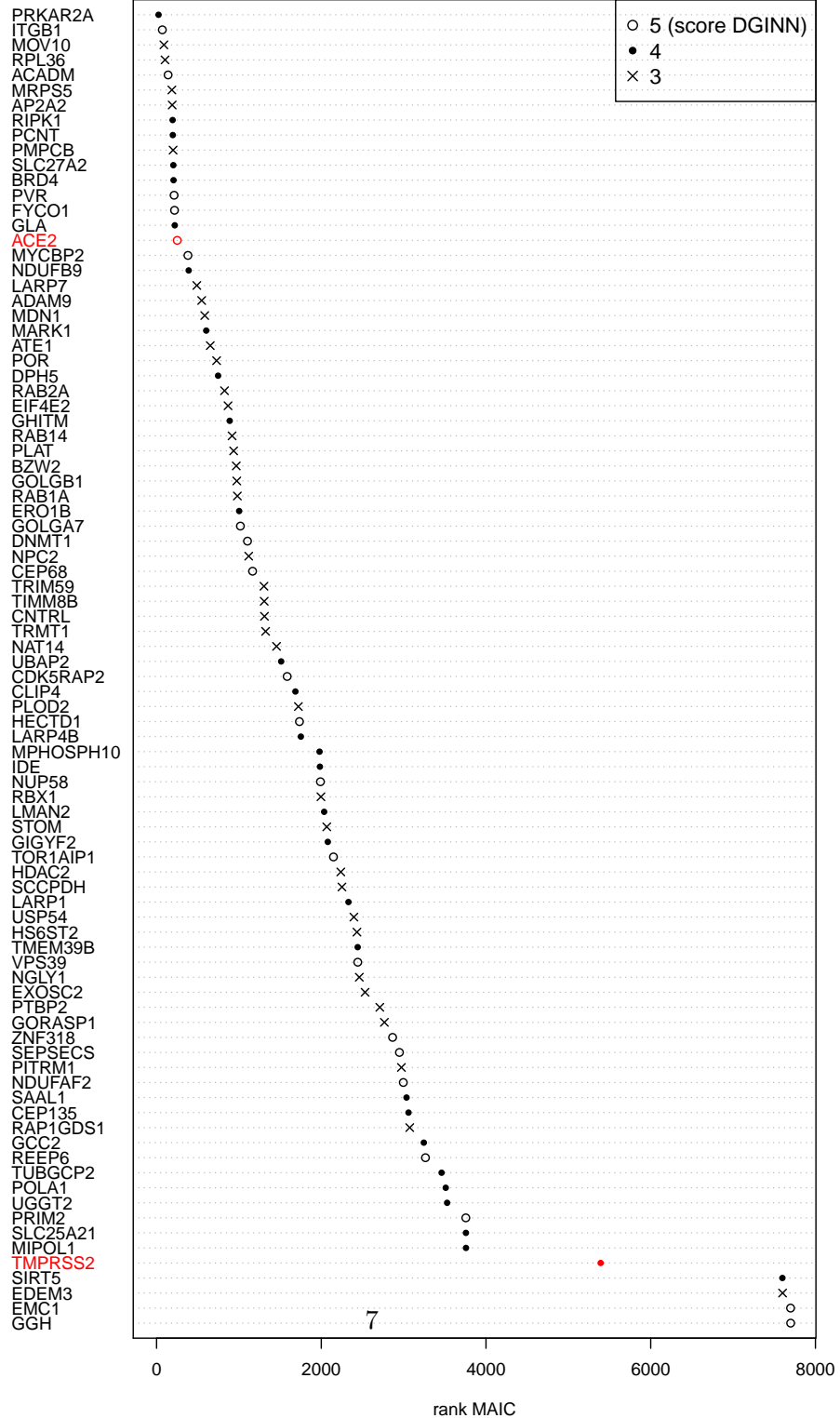


```
tmp<-maic[maic$nprimates>=3, c("gene", "rank", "nprimates")]
tmp<-tmp[order(tmp$rank, decreasing = TRUE),]

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 Pan Corona

```
pancorona<-read.table(paste0(workdir, "data/pancorona_S5.csv"),
                      h=T, fill = TRUE, sep="\t")

## Warning in file(file, "rt"): impossible d'ouvrir le fichier
## Warning in file(file, "rt"): /home/adminmarie/Documents/CIRI_BIBS_projects/2020_05_Etienne_covid/2020_dginn_covid19
: Aucun fichier ou dossier de ce type
## Error in file(file, "rt"): impossible d'ouvrir la connexion

names(pancorona)<-c("tmp.Gene.name", names(pancorona)[-1])

## Error in eval(expr, envir, enclos): objet 'pancorona' introuvable

# Genes en commun
pancorona$tmp.Gene.name[pancorona$tmp.Gene.name %in% tablo$tmp.Gene.name]

## Error in eval(expr, envir, enclos): objet 'pancorona' introuvable

# Uniquement dans le tableau pancorona
sort(pancorona$tmp.Gene.name[(pancorona$tmp.Gene.name %in% tablo$tmp.Gene.name)==FALSE])

## Error in sort(pancorona$tmp.Gene.name[(pancorona$tmp.Gene.name
## Error in sort(pancorona$tmp.Gene.name[(pancorona$tmp.Gene.name
%in% tablo$tmp.Gene.name) == : objet 'pancorona' introuvable

## Uniquement dans tableau
sort(tablo$tmp.Gene.name[(tablo$tmp.Gene.name %in% pancorona$tmp.Gene.name)==FALSE])

## Error in tablo$tmp.Gene.name %in% pancorona$tmp.Gene.name: objet
'pancorona' introuvable

pancorona<-pancorona[,c("tmp.Gene.name", "TOTAL")]

## Error in eval(expr, envir, enclos): objet 'pancorona' introuvable

pandginn<-na.omit(merge(pancorona, tablo, by="tmp.Gene.name", all.x=TRUE))

## Error in merge(pancorona, tablo, by = "tmp.Gene.name", all.x
= TRUE): objet 'pancorona' introuvable
```



```

pandginn<-pandginn[order(pandginn$nprimates),]
## Error in eval(expr, envir, enclos): objet 'pandginn' introuvable
pandginn<-pandginn[order(pandginn$TOTAL),]
## Error in eval(expr, envir, enclos): objet 'pandginn' introuvable
dotchart(as.matrix(pandginn[,2]), labels = pandginn$tmp.Gene.name, xlim=c(0,5))
## Error in as.matrix(pandginn[, 2]): objet 'pandginn' introuvable
points(pandginn[,4], 1:nrow(pandginn), col="blue", pch=20, cex=0.7)
## Error in points(pandginn[, 4], 1:nrow(pandginn), col = "blue",
pch = 20, : objet 'pandginn' introuvable
points(pandginn[,3], 1:nrow(pandginn), col="blue", pch=4)
## Error in points(pandginn[, 3], 1:nrow(pandginn), col = "blue",
pch = 4): objet 'pandginn' introuvable
legend("bottomright", c("pancorona score", "dginn primate score", "dginn bats score"))
## Error in strwidth(legend, units = "user", cex = cex, font = text.font):
plot.new has not been called yet

```

A-t-on un enrichissement en Pan-corona dans nos gènes sous PS?

```

pandginnall<-merge(pancorona, tablo, by="tmp.Gene.name", all.x=FALSE,all.y=TRUE)
## Error in merge(pancorona, tablo, by = "tmp.Gene.name", all.x
= FALSE, : objet 'pancorona' introuvable
dim(pandginnall)
## Error in eval(expr, envir, enclos): objet 'pandginnall' introuvable
# test indépendance: under PS / in the pancorona list
table(is.na(pandginnall$TOTAL)==FALSE)
## Error in table(is.na(pandginnall$TOTAL) == FALSE): objet 'pandginnall'
introuvable
table(pandginnall$nbats>=3)

```

```

## Error in table(pandginnall$nbats >= 3):  objet 'pandginnall'
introuvable

chi<-table(is.na(pandginnall$TOTAL)==FALSE,pandginnall$nbats>=3)

## Error in table(is.na(pandginnall$TOTAL) == FALSE, pandginnall$nbats
>= :  objet 'pandginnall' introuvable

chi

## Error in eval(expr, envir, enclos):  objet 'chi' introuvable

chisq.test(chi)

## Error in is.data.frame(x):  objet 'chi' introuvable

table(is.na(pandginnall$TOTAL)==FALSE)

## Error in table(is.na(pandginnall$TOTAL) == FALSE): objet 'pandginnall'
introuvable

table(pandginnall$nprimates>=3)

## Error in table(pandginnall$nprimates >= 3):  objet 'pandginnall'
introuvable

chi<-table(is.na(pandginnall$TOTAL)==FALSE,pandginnall$nprimates>=3)

## Error in table(is.na(pandginnall$TOTAL) == FALSE, pandginnall$nprimates
>= :  objet 'pandginnall' introuvable

chi

## Error in eval(expr, envir, enclos):  objet 'chi' introuvable

chisq.test(chi)

## Error in is.data.frame(x):  objet 'chi' introuvable

```

No enrichment in PanCORONA in our genes under PS.