

# Patristic distances analysis

Roberto Cárcamo Calvo

4/26/2022

```
tree_cluster<- ape::read.tree("./GI_VP1_complete.nwk")
print(tree_cluster)
```

```
##
## Phylogenetic tree with 101 tips and 99 internal nodes.
##
## Tip labels:
##   GI.3/FE8_2015/BBD74614, GI.3/FE2_2015/BBD74612, GI.3/FE29_2015/BBD74616, GI.3/CFIA_FV_448_1b_43479
## Node labels:
##   , 0.5520, 1.0000, 0.9980, 0.9380, 1.0000, ...
##
## Unrooted; includes branch lengths.
```

```
d_cluster <- distTips(tree_cluster, tips = "all", method = ("patristic"), useC = TRUE)
write.csv(as.matrix(d_cluster), "tabla_noemi_patristic_completa_cluster.csv")
tabla_completa_gaps <- read.csv("tabla_noemi_patristic_completa_cluster.csv", sep="," , row.names = 1)
```

```
#colnames(tabla_completa_gaps) <- str_remove(colnames(tabla_completa_gaps), "_.")
#colnames(tabla_completa_gaps)
```

```
#rownames(tabla_completa_gaps) <- str_remove(rownames(tabla_completa_gaps), "_.")
#rownames(tabla_completa_gaps)
```

```
#Intra-clusters comparisons
```

```
comparacion_gi3gi3 <- tabla_completa_gaps[c(1:55), c(1:55),]
comparacion_gi3gi3 <- unlist(comparacion_gi3gi3,use.names=FALSE)
comparacion_gi3gi3 <- fun.zero.omit(comparacion_gi3gi3)
comparacion_gi3gi3 <- unique(comparacion_gi3gi3)
media_gi3gi3 <- mean(comparacion_gi3gi3)
sd_gi3gi3 <- sd(comparacion_gi3gi3)
```

```
comparacion_gina1gina1 <- tabla_completa_gaps[c(59:67), c(59:67)]
comparacion_gina1gina1 <- unlist(comparacion_gina1gina1,use.names=FALSE)
comparacion_gina1gina1 <- fun.zero.omit(comparacion_gina1gina1)
comparacion_gina1gina1 <- unique(comparacion_gina1gina1)
media_gina1gina1 <- mean(comparacion_gina1gina1)
sd_gina1gina1 <- sd(comparacion_gina1gina1)
```

```

comparacion_gina2gina2 <- tabla_completa_gaps[c(56:58), c(56:58)]
comparacion_gina2gina2 <- unlist(comparacion_gina2gina2,use.names=FALSE)
comparacion_gina2gina2 <- fun.zero.omit(comparacion_gina2gina2)
comparacion_gina2gina2 <- unique(comparacion_gina2gina2)
media_gina2gina2 <- mean(comparacion_gina2gina2)
sd_gina2gina2 <- sd(comparacion_gina2gina2)

```

```

comparacion_gi7gi7 <- tabla_completa_gaps[c(80:86), c(80:86)]
comparacion_gi7gi7 <- unlist(comparacion_gi7gi7,use.names=FALSE)
comparacion_gi7gi7 <- fun.zero.omit(comparacion_gi7gi7)
comparacion_gi7gi7 <- unique(comparacion_gi7gi7)
media_gi7gi7 <- mean(comparacion_gi7gi7)
sd_gi7gi7 <- sd(comparacion_gi7gi7)

```

```

comparacion_gi8gi8 <- tabla_completa_gaps[c(68:73), c(68:73)]
comparacion_gi8gi8 <- unlist(comparacion_gi8gi8,use.names=FALSE)
comparacion_gi8gi8 <- fun.zero.omit(comparacion_gi8gi8)
comparacion_gi8gi8 <- unique(comparacion_gi8gi8)
media_gi8gi8 <- mean(comparacion_gi8gi8)
sd_gi8gi8 <- sd(comparacion_gi8gi8)

```

```

comparacion_gi9gi9 <- tabla_completa_gaps[c(74:79), c(74:79)]
comparacion_gi9gi9 <- unlist(comparacion_gi9gi9,use.names=FALSE)
comparacion_gi9gi9 <- fun.zero.omit(comparacion_gi9gi9)
comparacion_gi9gi9 <- unique(comparacion_gi9gi9)
media_gi9gi9 <- mean(comparacion_gi9gi9)
sd_gi9gi9 <- sd(comparacion_gi9gi9)

```

### *#Inter GI.3 comparisons*

```

comparacion_gi3gina1 <- tabla_completa_gaps[c(1:55), c(59:67)]
comparacion_gi3gina1 <- unlist(comparacion_gi3gina1,use.names=FALSE)
media_gi3gina1 <- mean(comparacion_gi3gina1)
sd_gi3gina1 <- sd(comparacion_gi3gina1)

```

```

comparacion_gi3gina2 <- tabla_completa_gaps[c(1:55), c(56:58)]
comparacion_gi3gina2 <- unlist(comparacion_gi3gina2,use.names=FALSE)
media_gi3gina2 <- mean(comparacion_gi3gina2)
sd_gi3gina2 <- sd(comparacion_gi3gina2)

```

```

comparacion_gi3gi7 <- tabla_completa_gaps[c(1:55), c(80:86)]
comparacion_gi3gi7 <- unlist(comparacion_gi3gi7,use.names=FALSE)
media_gi3gi7 <- mean(comparacion_gi3gi7)
sd_gi3gi7 <- sd(comparacion_gi3gi7)

```

```

comparacion_gi3gi8 <- tabla_completa_gaps[c(1:55), c(68:73)]
comparacion_gi3gi8 <- unlist(comparacion_gi3gi8,use.names=FALSE)
media_gi3gi8 <- mean(comparacion_gi3gi8)
sd_gi3gi8 <- sd(comparacion_gi3gi8)

```

```

comparacion_gi3gi9 <- tabla_completa_gaps[c(1:55), c(74:79)]
comparacion_gi3gi9 <- unlist(comparacion_gi3gi9,use.names=FALSE)
media_gi3gi9 <- mean(comparacion_gi3gi9)
sd_gi3gi9 <- sd(comparacion_gi3gi9)

#Inter GI.NA1 comparisons

comparacion_gina1gina2 <- tabla_completa_gaps[c(59:67), c(56:58)]
comparacion_gina1gina2 <- unlist(comparacion_gina1gina2,use.names=FALSE)
media_gina1gina2 <- mean(comparacion_gina1gina2)
sd_gina1gina2 <- sd(comparacion_gina1gina2)

comparacion_gina1gi7 <- tabla_completa_gaps[c(59:67), c(80:86)]
comparacion_gina1gi7 <- unlist(comparacion_gina1gi7,use.names=FALSE)
media_gina1gi7 <- mean(comparacion_gina1gi7)
sd_gina1gi7 <- sd(comparacion_gina1gi7)

comparacion_gina1gi8 <- tabla_completa_gaps[c(59:67), c(68:73)]
comparacion_gina1gi8 <- unlist(comparacion_gina1gi8,use.names=FALSE)
media_gina1gi8 <- mean(comparacion_gina1gi8)
sd_gina1gi8 <- sd(comparacion_gina1gi8)

comparacion_gina1gi9 <- tabla_completa_gaps[c(59:67), c(74:79)]
comparacion_gina1gi9 <- unlist(comparacion_gina1gi9,use.names=FALSE)
media_gina1gi9 <- mean(comparacion_gina1gi9)
sd_gina1gi9 <- sd(comparacion_gina1gi9)

#Inter GI.NA2 comparisons

comparacion_gina2gi7 <- tabla_completa_gaps[c(56:58), c(80:86)]
comparacion_gina2gi7 <- unlist(comparacion_gina2gi7,use.names=FALSE)
media_gina2gi7 <- mean(comparacion_gina2gi7)
sd_gina2gi7 <- sd(comparacion_gina2gi7)

comparacion_gina2gi8 <- tabla_completa_gaps[c(56:58), c(68:73)]
comparacion_gina2gi8 <- unlist(comparacion_gina2gi8,use.names=FALSE)
media_gina2gi8 <- mean(comparacion_gina2gi8)
sd_gina2gi8 <- sd(comparacion_gina2gi8)

comparacion_gina2gi9 <- tabla_completa_gaps[c(56:58), c(74:79)]
comparacion_gina2gi9 <- unlist(comparacion_gina2gi9,use.names=FALSE)
media_gina2gi9 <- mean(comparacion_gina2gi9)
sd_gina2gi9 <- sd(comparacion_gina2gi9)

#Plots

comparaciones <- c("GI.3-to-GI.3", "GI.NA1-to-GI.NA1", "GI.NA2-to-GI.NA2", "GI.7-to-GI.7", "GI.8-to-GI.8")

```

```

media <- c(media_gi3gi3, media_gina1gina1, media_gina2gina2, media_gi7gi7, media_gi8gi8, media_gi9gi9, media_gi3gina1, media_gi3gina2, media_gi3gi7, media_gi3gina7)
sd <- c(sd_gi3gi3, sd_gina1gina1, sd_gina2gina2, sd_gi7gi7, sd_gi8gi8, sd_gi9gi9, sd_gi3gina1, sd_gi3gina2, sd_gi3gi7, sd_gi3gina7)
sd <- as.numeric(sd)

df <- data.frame(comparaciones, media, sd)

boxplot <- ggplot(df, aes(x=comparaciones, fill = comparaciones), show.legend = TRUE) + geom_boxplot(aes(fill = comparaciones))
boxplot_crossbar <- ggplot(df, aes(x=comparaciones, fill = comparaciones), show.legend = TRUE) + geom_crossbar(aes(fill = comparaciones))

r <- boxplot + scale_fill_brewer(palette="Blues") + theme(legend.position = "none") + ylab("Distance")

r <- boxplot + scale_fill_brewer(palette="Blues") + theme(legend.position = "none") + ylab("Distance")
r <- r + theme(axis.text = element_text(size = 10, face="bold"))
r <- r + theme(axis.title = element_text(size = 12.5))
r <- r + geom_hline(yintercept = 0.15, linetype="dashed",
                    color = "red", size=0.3)
r <- r + coord_flip()

comparisons <- fct_inorder(comparaciones)
r <- r + aes(x = comparisons)

r <- r + geom_segment(aes(x = 0.75, y = media_gi3gi3-2*sd_gi3gi3, xend = 1.25, yend = media_gi3gi3-2*sd_gi3gi3))
r <- r + geom_segment(aes(x = 0.75, y = media_gi3gi3+2*sd_gi3gi3, xend = 1.25, yend = media_gi3gi3+2*sd_gi3gi3))

r <- r + geom_segment(aes(x = 1.75, y = media_gina1gina1-2*sd_gina1gina1, xend = 2.25, yend = media_gina1gina1-2*sd_gina1gina1))
r <- r + geom_segment(aes(x = 1.75, y = media_gina1gina1+2*sd_gina1gina1, xend = 2.25, yend = media_gina1gina1+2*sd_gina1gina1))

r <- r + geom_segment(aes(x = 2.75, y = media_gina2gina2-2*sd_gina2gina2, xend = 3.25, yend = media_gina2gina2-2*sd_gina2gina2))
r <- r + geom_segment(aes(x = 2.75, y = media_gina2gina2+2*sd_gina2gina2, xend = 3.25, yend = media_gina2gina2+2*sd_gina2gina2))

r <- r + geom_segment(aes(x = 3.75, y = media_gi7gi7-2*sd_gi7gi7, xend = 4.25, yend = media_gi7gi7-2*sd_gi7gi7))
r <- r + geom_segment(aes(x = 3.75, y = media_gi7gi7+2*sd_gi7gi7, xend = 4.25, yend = media_gi7gi7+2*sd_gi7gi7))

r <- r + geom_segment(aes(x = 4.75, y = media_gi8gi8-2*sd_gi8gi8, xend = 5.25, yend = media_gi8gi8-2*sd_gi8gi8))
r <- r + geom_segment(aes(x = 4.75, y = media_gi8gi8+2*sd_gi8gi8, xend = 5.25, yend = media_gi8gi8+2*sd_gi8gi8))

r <- r + geom_segment(aes(x = 5.75, y = media_gi9gi9-2*sd_gi9gi9, xend = 6.25, yend = media_gi9gi9-2*sd_gi9gi9))
r <- r + geom_segment(aes(x = 5.75, y = media_gi9gi9+2*sd_gi9gi9, xend = 6.25, yend = media_gi9gi9+2*sd_gi9gi9))

r <- r + geom_segment(aes(x = 6.75, y = media_gi3gina1-2*sd_gi3gina1, xend = 7.25, yend = media_gi3gina1-2*sd_gi3gina1))
r <- r + geom_segment(aes(x = 6.75, y = media_gi3gina1+2*sd_gi3gina1, xend = 7.25, yend = media_gi3gina1+2*sd_gi3gina1))

r <- r + geom_segment(aes(x = 7.75, y = media_gi3gina2-2*sd_gi3gina2, xend = 8.25, yend = media_gi3gina2-2*sd_gi3gina2))
r <- r + geom_segment(aes(x = 7.75, y = media_gi3gina2+2*sd_gi3gina2, xend = 8.25, yend = media_gi3gina2+2*sd_gi3gina2))

r <- r + geom_segment(aes(x = 8.75, y = media_gi3gi7-2*sd_gi3gi7, xend = 9.25, yend = media_gi3gi7-2*sd_gi3gi7))
r <- r + geom_segment(aes(x = 8.75, y = media_gi3gi7+2*sd_gi3gi7, xend = 9.25, yend = media_gi3gi7+2*sd_gi3gi7))

```

```

r <- r + geom_segment(aes(x = 9.75, y = media_gi3gi8-2*sd_gi3gi8, xend = 10.25, yend = media_gi3gi8-2*
r <- r + geom_segment(aes(x = 9.75, y = media_gi3gi8+2*sd_gi3gi8, xend = 10.25, yend = media_gi3gi8+2*

r <- r + geom_segment(aes(x = 10.75, y = media_gi3gi9-2*sd_gi3gi9, xend = 11.25, yend = media_gi3gi9-2*
r <- r + geom_segment(aes(x = 10.75, y = media_gi3gi9+2*sd_gi3gi9, xend = 11.25, yend = media_gi3gi9+2*

r <- r + geom_segment(aes(x = 11.75, y = media_gina1gina2-2*sd_gina1gina2, xend = 12.25, yend = media_g
r <- r + geom_segment(aes(x = 11.75, y = media_gina1gina2+2*sd_gina1gina2, xend = 12.25, yend = media_g

r <- r + geom_segment(aes(x = 12.75, y = media_gina1gi7-2*sd_gina1gi7, xend = 13.25, yend = media_gina
r <- r + geom_segment(aes(x = 12.75, y = media_gina1gi7+2*sd_gina1gi7, xend = 13.25, yend = media_gina

r <- r + geom_segment(aes(x = 13.75, y = media_gina1gi8-2*sd_gina1gi8, xend = 14.25, yend = media_gina
r <- r + geom_segment(aes(x = 13.75, y = media_gina1gi8+2*sd_gina1gi8, xend = 14.25, yend = media_gina

r <- r + geom_segment(aes(x = 14.75, y = media_gina1gi9-2*sd_gina1gi9, xend = 15.25, yend = media_gina
r <- r + geom_segment(aes(x = 14.75, y = media_gina1gi9+2*sd_gina1gi9, xend = 15.25, yend = media_gina

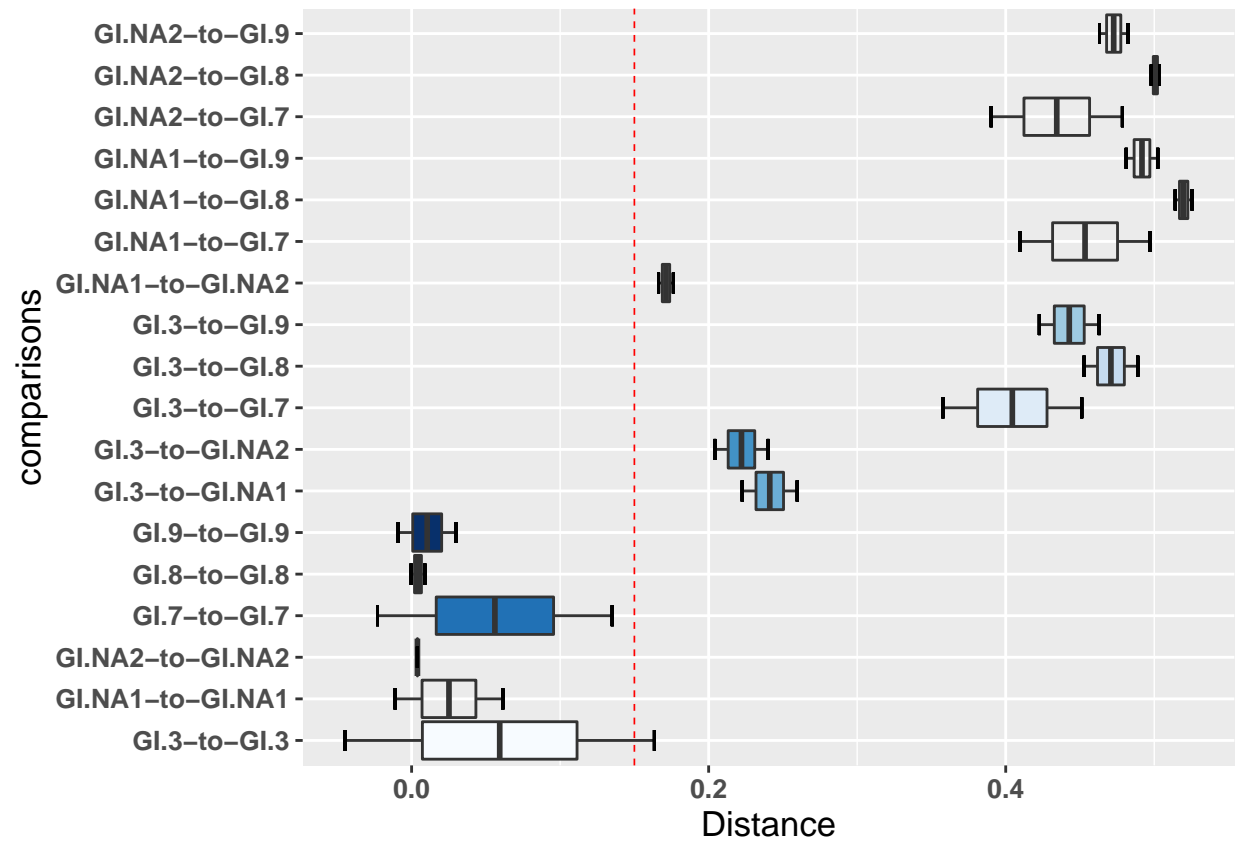
r <- r + geom_segment(aes(x = 15.75, y = media_gina2gi7-2*sd_gina2gi7, xend = 16.25, yend = media_gina
r <- r + geom_segment(aes(x = 15.75, y = media_gina2gi7+2*sd_gina2gi7, xend = 16.25, yend = media_gina

r <- r + geom_segment(aes(x = 16.75, y = media_gina2gi8-2*sd_gina2gi8, xend = 17.25, yend = media_gina
r <- r + geom_segment(aes(x = 16.75, y = media_gina2gi8+2*sd_gina2gi8, xend = 17.25, yend = media_gina

r <- r + geom_segment(aes(x = 17.75, y = media_gina2gi9-2*sd_gina2gi9, xend = 18.25, yend = media_gina
r <- r + geom_segment(aes(x = 17.75, y = media_gina2gi9+2*sd_gina2gi9, xend = 18.25, yend = media_gina

r

```



```
ggsave("patristic_distance_masGI7.tiff", height = 10.97252, width = 17.75354, units = "cm", device='tif
```