

Evidencia Final 2

Fernando Israel Rios Garcia

2023-04-30

Evidencia 2

¿Cuáles son los virus “similares” a SARS-CoV-2 en el mundo?

El SRAS-Cov-2 es un virus de la familia coronavirus, los síntomas que un humano presenta al contagiarse son los de la gripe. Por lo tanto los virus similares al COVID-19 presentan síntomas parecidos en los huéspedes que contagian, sus diferencias varían en la efectividad en la que se contagian y la tasa de mortalidad. En los humanos hay siete virus similares al SARS-Cov-2:

1. SARS-CoV-2
2. SARS-CoV
3. MERS-CoV
4. HCoV-229E
5. HCoV-NL63
6. HCoV-OC43
7. HCoV-HKU1

Especies que pueden ser cercanos al genoma de SARS-CoV-2? Incluye de qué especies son y menciona las referencias de los artículos consultados en PUBMED.

En el artículo Similarities and Dissimilarities of COVID-19 and Other Coronavirus Diseases escrito por Sing Fung y Ding Xiang Liu mencionan los siguientes virus similares al genoma del SARS-Cov-2:

Host	Coronavirus
Humans	SARS-CoV
	MERS-CoV
	HCoV-229E
	HCoV-NL63
	HCoV-OC43
	HCoV-HKU1
Pig	TGEV
	PRCoV
	PHEV
	PEDV
	PDCoV
	SADS-CoV
Mouse	MHV
Rat	SDAV
	RCoV-P
	Horse
Cattle	ECoV
	BCoV

Host	Coronavirus
Cat	FIPV FECV
Dog	CCoV CRCoV
Chicken	IBV

En relación con la situación actual reflexiona, ¿qué propondrías que se deba hacer durante la contingencia del SARS-CoV-2 en comunidades de bajos recursos? Si tu vivieras en una situación de escasos recursos, ¿qué harías? Justifica tu respuesta.

Durante la contingencia del SARS-CoV-2 en comunidades de bajos recursos, es importante tomar medidas para proteger la salud de las personas y prevenir la propagación del virus. Algunas de las medidas que se pueden tomar incluyen:

1. Promover la educación sobre el virus: es importante difundir información precisa y confiable sobre el virus y cómo prevenir su propagación. Se pueden utilizar carteles, folletos o anuncios en los medios locales para educar a la comunidad.
2. Promover el distanciamiento físico: aunque puede ser difícil en comunidades de bajos recursos donde las personas viven en condiciones más cercanas, se debe alentar a la gente a mantener una distancia física de al menos 2 metros entre ellos.
3. Fomentar el uso de mascarillas: las mascarillas son una herramienta importante para prevenir la propagación del virus, y es importante educar a la comunidad sobre cómo usarlas correctamente y proporcionar mascarillas de forma gratuita para quienes no las tengan.
4. Promover el lavado de manos frecuente: el lavado de manos es una de las mejores maneras de prevenir la propagación del virus, y es importante proporcionar agua y jabón para que las personas puedan lavarse las manos con frecuencia.
5. Proporcionar acceso a suministros médicos: se debe asegurar que las personas tengan acceso a suministros médicos, como termómetros, oxígeno y medicamentos, en caso de que se enfermen.
6. Proporcionar ayuda financiera: las comunidades de bajos recursos pueden verse más afectadas por la crisis económica causada por la pandemia, y es importante proporcionar ayuda financiera a las personas que lo necesitan para ayudarles a hacer frente a la crisis.

Analizar las secuencias de SARS-CoV-2 reportadas en los 20 países con más casos reportados. Y puedes tratar de responder a la pregunta: ¿Son muy diferentes las variantes entre cada país? ¿Es diferente el SARS-CoV-2 entre las diferentes poblaciones: Asiática, Hispana, Europea o Africana?

“SARS-CoV-2USA”, “SARS-CoV-2India”, “SARS-CoV-2Francia”, “SARS-CoV-2 Alemania”, “SARS-CoV-2 Brasil”, “SARS-CoV-2 Japon”, “SARS-CoV-2 Corea del Sur”, “SARS-CoV-2 Italia”, “SARS-CoV-2 Rusia”, “SARS-CoV-2 Turquía”, “SARS-CoV-2 España”, “SARS-CoV-2 Vietnam”, “SARS-CoV-2 Australia”, “SARS-CoV-2 Argentina”, “SARS-CoV-2 Taiwán”, “SARS-CoV-2 Holanda”, “SARS-CoV-2 Iran”, “SARS-CoV-2 México”, “SARS-CoV-2 Indonesia”, “SARS-CoV-2 Polonia”

Referencias

Torun Ş, Kesim Ç, Süner A, Botan Yıldırım B, Özen Ö, Akçay Ş. Influenza viruses and SARS-CoV-2 in adult: ‘Similarities and differences’. Tuberk Toraks. 2021 Dec;69(4):458-468. English. doi: 10.5578/tt.20219603. PMID: 34957739.

Chen Z, Boon SS, Wang MH, Chan RWY, Chan PKS. Genomic and evolutionary comparison between SARS-CoV-2 and other human coronaviruses. J Virol Methods. 2021 Mar;289:114032. doi: 10.1016/j.jviromet.2020.114032. Epub 2020 Dec 5. PMID: 33290786; PMCID: PMC7718587.

Keshta AS, Mallah SI, Al Zubaidi K, Ghorab OK, Keshta MS, Alarabi D, Abousaleh MA, Salman MT, Taha OE, Zeidan AA, Elsaid MF, Tang P. COVID-19 versus SARS: A comparative review. J Infect Public Health. 2021 Jul;14(7):967-977. doi: 10.1016/j.jiph.2021.04.007. Epub 2021 Apr 24. PMID: 34130121; PMCID: PMC8064890.

Sing, T. & Liu, D. (2021). Similarities and Dissimilarities of COVID-19 and Other Coronavirus Diseases. Annual Review of Microbiology, Vol. 75, 19-47. <https://doi.org/10.1146/annurev-micro-110520-023212>

<https://rdr.io/bioc/QSutils/man/DNA.dist.html>

Code

```
# Call packages
library(ade4)
library(ape)
library(adegenet) # Packages to read DNA seq
library(gridExtra)
library(ggplot2)
library(stringr)
library(seqinr)

# Bring files, functions files
source("../lengthSeq.r")
source("../percentageBases.r")
source("../phylogeneticTreeGraph.r")

seq1 <- readLines("../secuenciasEvidencia2/Mexico.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqMexico <- strsplit(var1, " ")[[1]]

seq1 <- readLines("../secuenciasEvidencia2/Argentina.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqArgentina <- strsplit(var1, " ")[[1]]

seq1 <- readLines("../secuenciasEvidencia2/Polonia.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqPoland <- strsplit(var1, " ")[[1]]

seq1 <- readLines("../secuenciasEvidencia2/Vietnam.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqVietnam <- strsplit(var1, " ")[[1]]

seq1 <- readLines("../secuenciasEvidencia2/CoreaDelSur.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqSouthKorea <- strsplit(var1, " ")[[1]]

seq1 <- readLines("../secuenciasEvidencia2/Alemania.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqGermany <- strsplit(var1, " ")[[1]]

seq1 <- readLines("../secuenciasEvidencia2/Australia.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
```

```

seqAustralia <- strsplit(var1, "")[[1]]

seq1 <- readLines("./secuenciasEvidencia2/PaisesBajos.fasta", n = -1)[-1]
var1 <- paste(seq1, collapse = " ")
seqNetherlands <- strsplit(var1, "")[[1]]

dnaSeq_array <- array(c(seqMexico, seqArgentina, seqPoland, seqVietnam, seqSouthKorea, seqGermany, seqA

# Call and store returns values of functions
# Call length function
lengthSEQMexico <-lengthSeq(seqMexico)
lengthSEQArgentina <-lengthSeq(seqArgentina)
lengthSEQPoland <-lengthSeq(seqPoland)
lengthSEQVietnam <-lengthSeq(seqVietnam)
lengthSEQSouthKorea <-lengthSeq(seqSouthKorea)
lengthSEQGermany <-lengthSeq(seqGermany)
lengthSEQAustralia <-lengthSeq(seqAustralia)
lengthSEQNetherlands <-lengthSeq(seqNetherlands)
# Call percentageBases function
basesSEQMexico <- percentageBases(seqMexico)
basesSEQArgentina <- percentageBases(seqArgentina)
basesSEQPoland <- percentageBases(seqPoland)
basesSEQVietnam <- percentageBases(seqVietnam)
basesSEQSouthKorea <- percentageBases(seqSouthKorea)
basesSEQGermany <- percentageBases(seqGermany)
basesSEQAustralia <- percentageBases(seqAustralia)
basesSEQNetherlands <- percentageBases(seqNetherlands)

print("Listado de la longitud de las secuencias de COVID-19 utilizadas en este reporte")

## [1] "Listado de la longitud de las secuencias de COVID-19 utilizadas en este reporte"
print(paste("Longitud de la secuencia de COVID-19 en Mexico",lengthSEQMexico))

## [1] "Longitud de la secuencia de COVID-19 en Mexico 29968"
print(paste("Longitud de la secuencia de COVID-19 en Argentina",lengthSEQArgentina))

## [1] "Longitud de la secuencia de COVID-19 en Argentina 30110"
print(paste("Longitud de la secuencia de COVID-19 en Poland",lengthSEQPoland))

## [1] "Longitud de la secuencia de COVID-19 en Poland 30068"
print(paste("Longitud de la secuencia de COVID-19 en Vietnam",lengthSEQVietnam))

## [1] "Longitud de la secuencia de COVID-19 en Vietnam 30173"
print(paste("Longitud de la secuencia de COVID-19 en South Korea",lengthSEQSouthKorea))

## [1] "Longitud de la secuencia de COVID-19 en South Korea 30117"
print(paste("Longitud de la secuencia de COVID-19 en Germany",lengthSEQGermany))

## [1] "Longitud de la secuencia de COVID-19 en Germany 30242"
print(paste("Longitud de la secuencia de COVID-19 en Australia",lengthSEQAustralia))

## [1] "Longitud de la secuencia de COVID-19 en Australia 29745"

```

```
print(paste("Longitud de la secuencia de COVID-19 en Netherlands",lengthSEQNetherlands))
```

```
## [1] "Longitud de la secuencia de COVID-19 en Netherlands 30036"
```

```
  gwuhan = data.frame(  
    bases = c("A", "C", "G", "T"),  
    g1 = percentageBases(seqMexico)  
  )
```

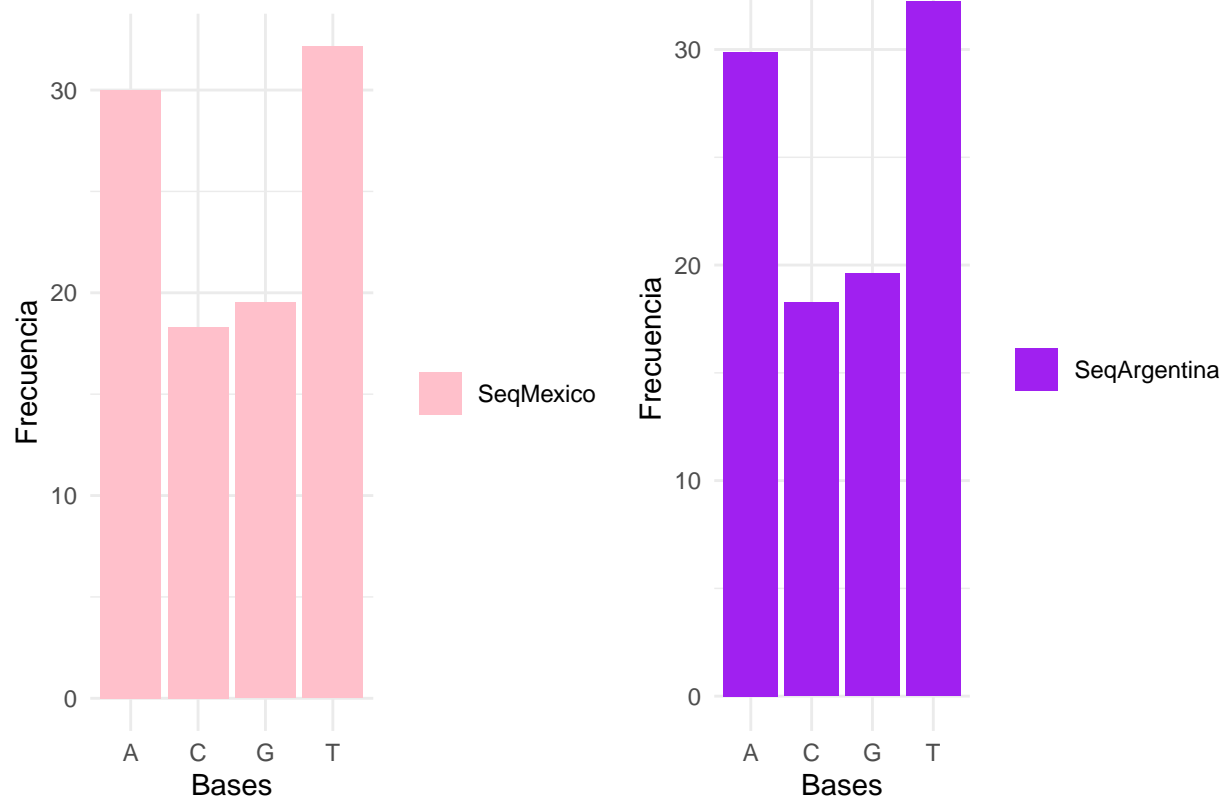
```
q <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +  
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqMexico"), stat = "identity", position = "stack") +  
  labs(title = "Comparación de las bases de ADN de las variantes del virus", x = "Bases", y = "Frecuencia") +  
  scale_fill_manual(values = c("SeqMexico" = "pink")) +  
  theme_minimal()
```

```
  gwuhan = data.frame(  
    bases = c("A", "C", "G", "T"),  
    g1 = percentageBases(seqArgentina)  
  )
```

```
w <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +  
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqArgentina"), stat = "identity", position = "stack") +  
  labs(x = "Bases", y = "Frecuencia", fill = "") +  
  scale_fill_manual(values = c("SeqArgentina" = "purple")) +  
  theme_minimal()
```

```
grid.arrange(q,w, ncol=2)
```

Comparación de las bases de ADN de las variantes del virus



```

    gwuhan = data.frame(
      bases = c("A", "C", "G", "T"),
      g1 = percentageBases(seqPoland)
    )

q <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqPoland"), stat = "identity", position = "dodge") +
  labs(title = "Comparación de las bases de ADN de las variantes del virus", x = "Bases", y = "Frecuencia") +
  scale_fill_manual(values = c("SeqPoland" = "pink")) +
  theme_minimal()

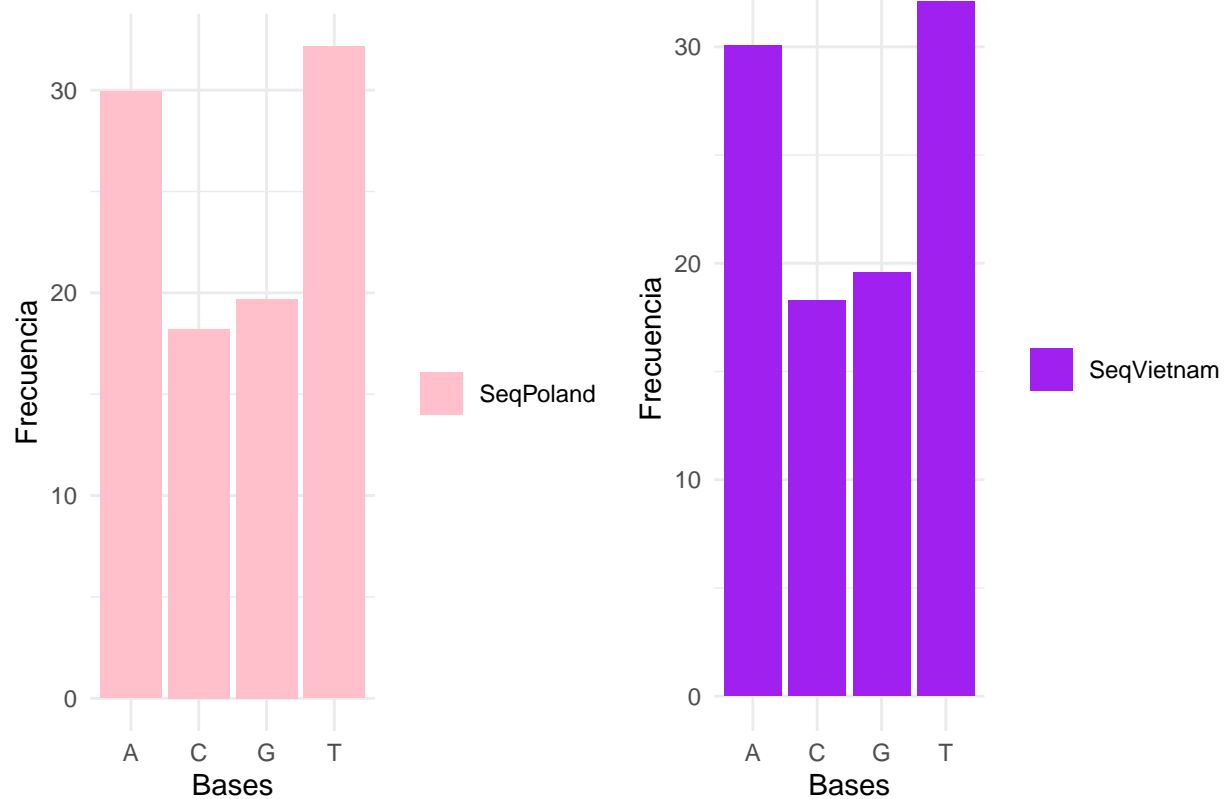
    gwuhan = data.frame(
      bases = c("A", "C", "G", "T"),
      g1 = percentageBases(seqVietnam)
    )

w <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqVietnam"), stat = "identity", position = "dodge") +
  labs(x = "Bases", y = "Frecuencia", fill = "") +
  scale_fill_manual(values = c("SeqVietnam" = "purple")) +
  theme_minimal()

grid.arrange(q,w, ncol=2)

```

Comparación de las bases de ADN de las variantes del virus



```

    gwuhan = data.frame(
      bases = c("A", "C", "G", "T"),
      g1 = percentageBases(seqSouthKorea)
    )

q <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqSouthKorea"), stat = "identity", position = "dodge") +
  labs(title = "Comparación de las bases de ADN de las variantes del virus", x = "Bases", y = "Frecuencia") +
  scale_fill_manual(values = c("SeqSouthKorea" = "pink")) +
  theme_minimal()

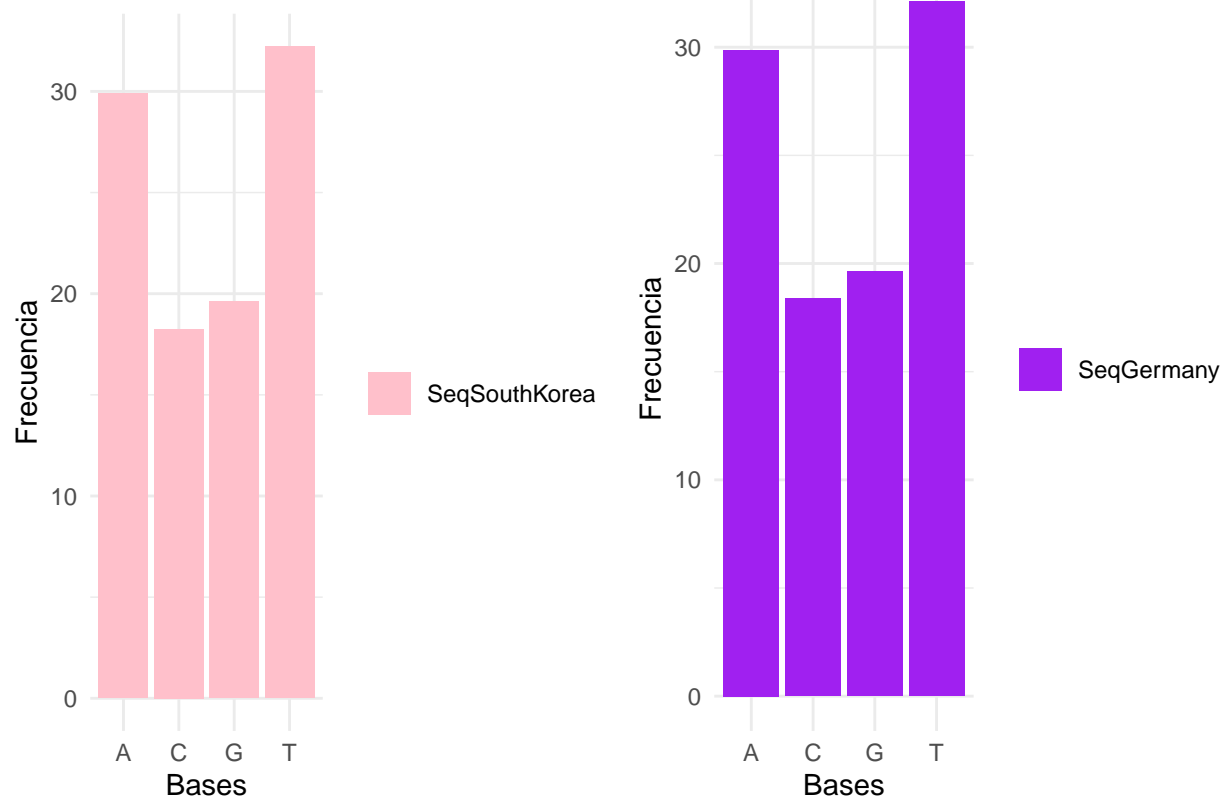
    gwuhan = data.frame(
      bases = c("A", "C", "G", "T"),
      g1 = percentageBases(seqGermany)
    )

w <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqGermany"), stat = "identity", position = "dodge") +
  labs(x = "Bases", y = "Frecuencia", fill = "") +
  scale_fill_manual(values = c("SeqGermany" = "purple")) +
  theme_minimal()

grid.arrange(q,w, ncol=2)

```

Comparación de las bases de ADN de las variantes del virus



```

    gwuhan = data.frame(
      bases = c("A", "C", "G", "T"),
      g1 = percentageBases(seqAustralia)
    )

q <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqAustralia"), stat = "identity", position = "dodge") +
  labs(title = "Comparación de las bases de ADN de las variantes del virus", x = "Bases", y = "Frecuencia") +
  scale_fill_manual(values = c("SeqAustralia" = "pink")) +
  theme_minimal()

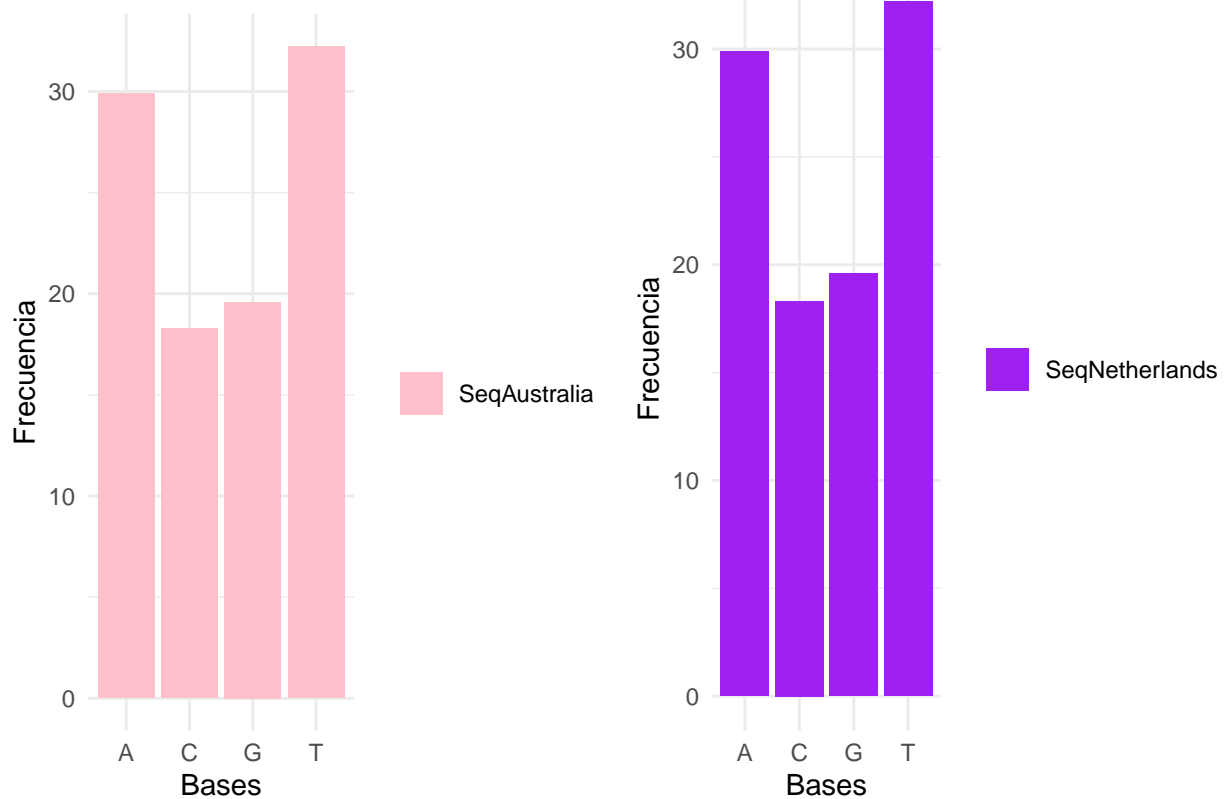
    gwuhan = data.frame(
      bases = c("A", "C", "G", "T"),
      g1 = percentageBases(seqNetherlands)
    )

w <- ggplot(gwuhan, aes(x = bases, y = g1, fill = bases)) +
  geom_bar(data = gwuhan, aes(x = bases, y = g1, fill = "SeqNetherlands"), stat = "identity", position = "dodge") +
  labs(x = "Bases", y = "Frecuencia", fill = "") +
  scale_fill_manual(values = c("SeqNetherlands" = "purple")) +
  theme_minimal()

grid.arrange(q,w, ncol=2)

```


Comparación de las bases de ADN de las variantes del virus



```
# Print phylogeneticTreeGraph
phylogeneticGraph <- phylogeneticTreeGraph() # Call Function that create graph

##
## Converting FASTA alignment into a DNAbin object...
##
##
## Finding the size of a single genome...
##
##
## genome size is: 29,869 nucleotides
##
## ( 375 lines per genome )
##
## Importing sequences...
## .....

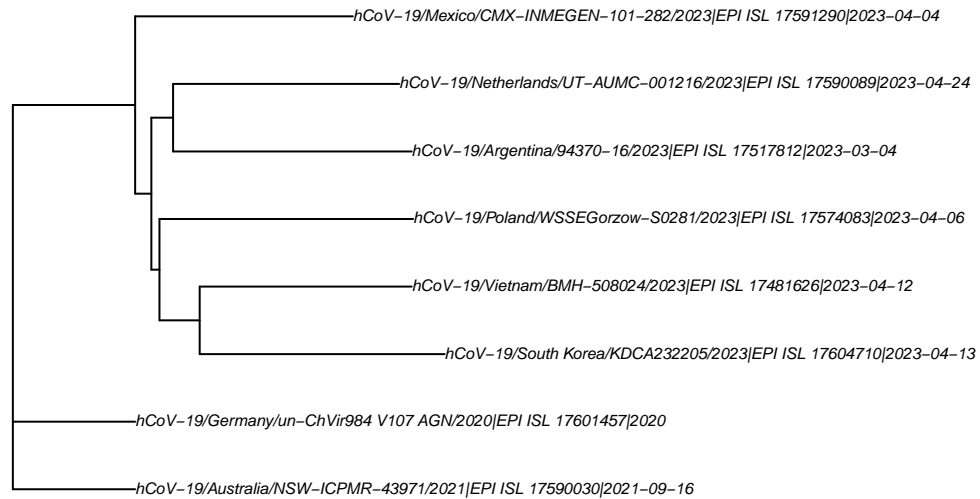
## Warning in split.default(txt, rep(1:nb.ind, each = LINES.PER.IND)): data length
## is not a multiple of split variable

##
## Forming final object...

## Warning in matrix(res, nrow = length(IND.LAB), byrow = TRUE): data length
## [237407] is not a sub-multiple or multiple of the number of rows [8]

##
## ...done.
```

Árbol de tipo NJ



phylogeneticGraph

NULL

Al ver nuestras graficas podemos notar como las diferencias aunque son percibibles, no es un cambio importante pues siguen teniendo parentesco entre si las variantes de cada pais. Revisando las graficas de historgramas en bases a las bases del ADN vemos como todas tienen un porcentaje parecido en su porcentaje correspondiente de cada bases. Revisando el arbol filo genetico de las ocho secuencias analizadas podemos comprender mejor su parentesco. Vemos como hay tres branch, las cuales representan una familia, dos secuencias son distintas y no comparten parentesco entre si mientras que la tercera parte de un nodo que lleva a otras dos difurcaciones, dejando dos clades. Estos dos clades correlacionan a sus secuencias como ancestros entre si. Analizando las ubicaciones geograficas no podemos sacar conclusiones ya que estan en distintos continentes y es aleatorio.