



Evidencia 1 | Análisis inicial

Análisis geográfico de las secuencias de SARS-CoV-2: una comparación global de la diversidad genética del virus

Análisis de biología Computacional

Profesor: Heriberto García Coronado

Héctor Alán Gutiérrez Gálvez - A01253031

Camila Guadalupe Rodríguez Martínez - A01253767

Campus Sonora Norte

5 de Mayo, 2023

Título del estudio: Análisis geográfico de las secuencias de SARS-CoV-2: una comparación global de la diversidad genética del virus.

Introducción

Un virus es un agente parasitario microscópico y acelular capaz de reproducirse en el interior de una célula hospedadora, por lo general ocasionando daños. Los virus son capaces de infectar a animales, plantas, bacterias e incluso otros virus llamados virófagos. (Equipo editorial, Etecé, 2021). En este trabajo, nos enfocaremos en el virus SARS-CoV-2 (también conocido como covid-19), el causante de la pandemia mundial iniciada en el 2020. Es una enfermedad que tiene una gran afectación en el pulmón, debido a la forma en que se propaga y se replica en el cuerpo humano, además de que es extremadamente contagiosa. Pertenece a la familia Coronaviridae y se divide en cuatro géneros: alfa, beta, gamma y delta. En la actualidad se han descubierto siete coronavirus en humanos que están encuadrados en los géneros alfa y beta. (Fernández-Pérez et al., 2021).

En el mundo hay 676,609,955 casos, han habido 6,881,955 muertes y existen 13,338,833,198 personas vacunadas. En México se han reportado 7,483,44 casos y 333,188 muertes mientras que 225,063,079 personas ya fueron vacunadas. En el estado de Sonora fueron reportados 201,304 casos y 10,349 muertes. (*COVID-19 Map - Johns Hopkins Coronavirus Resource Center*, n.d.). Sin embargo, hoy en día, el virus está bastante controlado, ya que las personas alrededor del mundo han generado anticuerpos como consecuencia de la vacuna.

La primera variante del virus SARS-CoV-2 que se propagó a nivel mundial fue la variante D614G. Esta variante fue identificada por primera vez en Europa en febrero de 2020 y luego se extendió rápidamente a todo el mundo, convirtiéndose en la cepa dominante del virus a nivel global. (Aguilar-Gamboa et al., 2021).

Las variantes de SARS-CoV-2 más conocidas son la variante del Reino Unido (Alpha), Sudáfrica (Beta), Brasil (Gamma) y la de India (Delta). Otras variantes del virus, que existen en otras regiones del mundo son las siguientes: Epsilon también conocida como variante de California, Eta identificada por primera vez en Diciembre del 2020 en Nigeria y la variante Kappa identificada por primera vez en India en Diciembre del 2020. (*Enfermedad Del Coronavirus 2019 (COVID-19)*, 2020). De igual manera, se han identificado varios coronavirus en otras especies animales que comparten similitudes genéticas con el

SARS-CoV-2. A continuación, se presentan algunas de las especies animales en las que se han identificado coronavirus similares al SARS-CoV-2:

- Pangolines: se ha informado que los coronavirus similares al SARS-CoV-2 se han encontrado en pangolines, específicamente en la especie de pangolín de escamas largas (*Manis pentadactyla*) y la especie de pangolín de escamas chinas (*Manis javanica*). Estos coronavirus tienen una similitud genética del 85% al 92% con el SARS-CoV-2 (Lam, et al., 2020).
- Murciélagos: se ha informado que los coronavirus similares al SARS-CoV-2 se han encontrado en varias especies de murciélagos, incluidos los murciélagos herradura chinos (*Rhinolophus spp.*) y los murciélagos de herradura de alas largas (*Miniopterus spp.*). Estos coronavirus tienen una similitud genética del 96% al 97% con el SARS-CoV-2 (Zhou, et al., 2020).

La biología computacional es una disciplina interdisciplinaria que utiliza herramientas computacionales y técnicas de análisis de datos para abordar problemas biológicos complejos. Se basa en la integración de la biología molecular, la bioquímica, la genética, la estadística y la informática para desarrollar y aplicar métodos computacionales para analizar y comprender grandes conjuntos de datos biológicos. (*Diccionario De Cáncer Del NCI*, n.d.-b)

En resumidas cuentas el análisis de genomas virales mediante herramientas de biología computacional es una herramienta muy importante si se trata del estudio y control de enfermedades virales y este uso puede tener un gran impacto para la prevención y tratamiento de enfermedades infecciosas.

Objetivo

El fin de la realización de este estudio es lograr una comparación global de la diversidad genética del SARS-CoV-2. Nuestra hipótesis es que en los diferentes países alrededor del mundo las características originales de las secuencias del SARS-CoV-2 cambiarán cierto porcentaje. Esto se puede deber a muchos factores, por ejemplo el clima y la zona geográfica. Por lo tanto las preguntas de investigación que se busca responder son las siguientes: ¿Son muy diferentes las variantes entre cada país? ¿Es diferente el SARS-CoV-2 entre los diferentes países?

Métodos y resultados

1. Obtén las secuencias de los genomas de las variantes de SARS-CoV-2 y otros virus que tu decides (8-10 genomas en total) desde el NCBI o el buscador de virus del NCBI.

```
1 #Descargamos los paquetes que contienen las funciones y herramientas necesarias
2 #para trabajar con los datos biológicos
3
4
5 library(viridisLite)
6 library(ape)
7 library(ade4)
8 library(seqrnr)
9 library(adegenet)
10 library(Biostrings)
11 library(DECIIPHER)
12 library(ggtree)
13 library(ggplot2)
14 library(tidyr)
15
16 #Utilizamos el getwd() para usar el directorio de trabajo actual
17 #y el setwd() le da a R una dirección a un directorio diferente
18 #para que ahí mismo se almacene y guarde la lectura y escritura de archivos.
19
20 getwd()
21 setwd("/Users/camilardgzm/AnalisisDeBiologiaComputacional/Virus/")
22
23 #EJERCICIO 1
24 #Obtener los 10 genomas de diferentes países
25
26 corona_virus <- c("NC_045512", "OP435368", "Q918256", "BS007312", "Q913932", "OP848485", "ON291271", "MT994849",
27 "OK096766", "MW466791")
28
29
30 # Leer archivo GenBank
31 virus_sequences <- read.GenBank(corona_virus)
32 write.dna(virus_sequences,file = "virus_coronavirus",format = "fasta")
33
```

2. Calcula la longitud de las secuencias de cada variante. Muestra el resultado impreso en consola y el código que utilizaste para realizar este punto.

Código:

```
35 #EJERCICIO 2
36 #Longitud de las secuencias de cada variante
37
38 enumerar <- c(1:10)
39 longitud <- sapply(virus_sequences, length)
40 gen_long <- data.frame(Numero_Genoma = enumerar, Secuencia_ID = corona_virus, Longitud = lengths)
41
42 print(gen_long, row.names = FALSE)
43
```

Resultado en consola:

```
> print(gen_long, row.names = FALSE)
  Numero_Genoma Secuencia_ID Longitud
    1      NC_045512    29903
    2      OP435368    29799
    3      OQ918256    29010
    4      BS007312    29737
    5      OQ913932    29660
    6      OP848485    29714
    7      ON291271    29689
    8      MT994849    29819
    9      OK096766    29766
   10      MW466791    29902
```

3. Genera una gráfica en la que se visualice la longitud de los 10 genomas virales que analizaste. Muestra el código que utilizaste para realizar este punto.

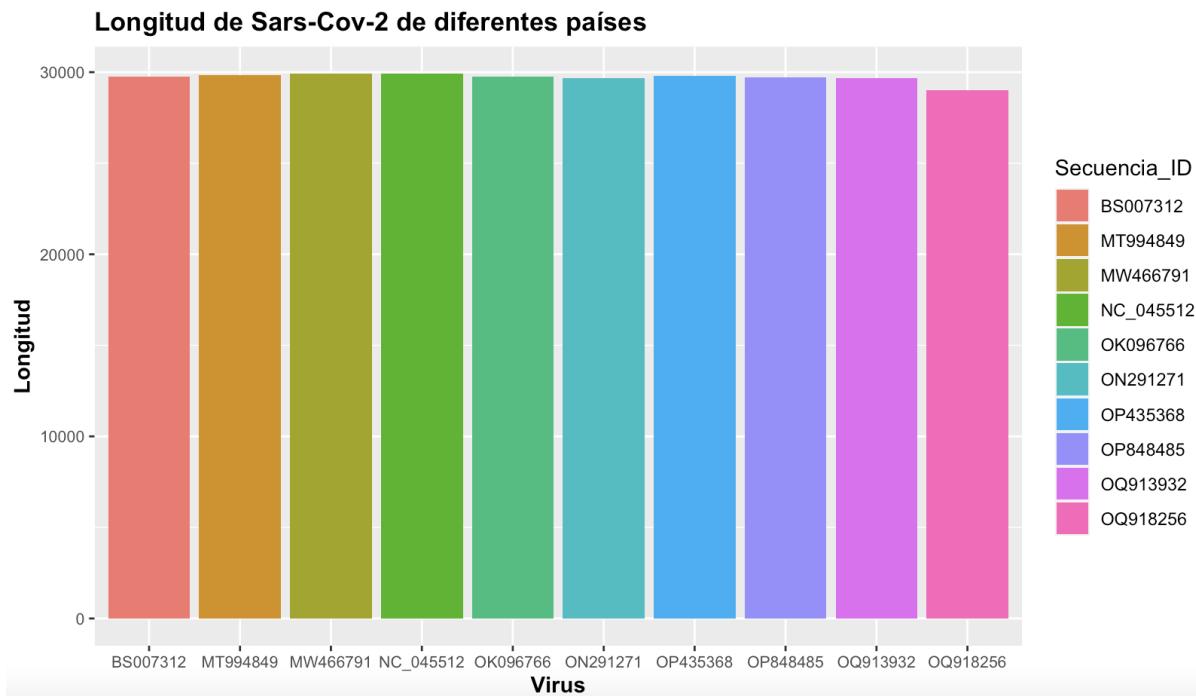
Código:

```

45 #EJERCICIO 3
46 #Gráfico de longitudes
47
48 grafica_longitud <- ggplot(gen_long, aes(x = Secuencia_ID, y = Longitud, fill = Secuencia_ID)) +
  geom_bar(stat = "identity") +
  labs(x = "Virus", y = "Longitud", title = "Longitud de Sars-Cov-2 de diferentes países") +
  theme(axis.text.x = element_text(size = 8),
        axis.text.y = element_text(size = 8),
        axis.title.x = element_text(face = "bold"),
        axis.title.y = element_text(face = "bold"),
        plot.title = element_text(face = "bold", size = 13))
56
57 grafica_longitud

```

Resultado en consola:



4. Agrega una interpretación escrita, desde el punto de vista biológico, para la gráfica anterior.

Se puede observar que los virus mostrados en la gráfica tienen un cambio muy leve en el tamaño de las secuencias. Esto significa que evidentemente dependiendo de la zona geográfica que se encuentre el virus puede variar su longitud en un porcentaje bastante pequeño. Al observar que el virus casi se mantiene igual en las diferentes zonas geográficas y tiempo transcurrido, indica que se ha mantenido muy parecido a como era desde un principio, teniendo una evolución lenta.

Además los virus que se están analizando tienen entre 1 a 2 años de diferencia, por lo que es otra variable que tiene un impacto en este “leve” cambio. Estos hallazgos podrían tener implicaciones en el desarrollo de tratamientos y vacunas contra el virus, así como en la comprensión de la evolución y la propagación del virus en todo el mundo.

5. Calcula la composición de nucleótidos a cada una de las variantes del virus. Muestra el resultado impreso en consola y el código que utilizaste para realizar este punto.

Código:

```
60 #EJERCICIO 5
61 #Composición de nucleótidos de cada genoma
62
63 virus_sequences_character <- c(as.character(virus_sequences))
64 nucleotidos_gen <- sapply(virus_sequences_character, count, 1)
65
66 nucleotidos_gen
67
```

Resultado en consola:

```
> nucleotidos_gen
NC_045512 0P435368 0Q918256 BS007312 0Q913932 0P848485 ON291271 MT994849 0K096766 MW466791
a      8954    8907    5827    8890    8867    8872    8867    8905    8889    8953
c      5492    5454    3618    5437    5427    5458    5439    5484    5458    5493
g      5863    5838    3840    5827    5814    5833    5827    5852    5842    5862
t      9594    9599    6402    9583    9540    9550    9555    9578    9577    9594
```

6. Genera una gráfica en la que se visualice la composición de nucleótidos de los 10 genomas virales que analizaste. Muestra el código que utilizaste para realizar este punto.

```

72 nucleotidos <- as.data.frame(nucleotidos_gen)
73
74 nucleotidos
75
76 #-----NC_045512-----
77
78 # Definir colores
79 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
80
81 # Crear gráfico de barras
82 grafica_nucleotidos <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = NC_045512, fill = rownames(nucleotidos))) +
83   geom_bar(stat = "identity") +
84   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma NC_045512") +
85   scale_x_discrete(labels = c("A", "C", "G", "T")) +
86   scale_fill_manual(values = colores) +
87   theme(axis.text.x = element_text(face = "bold"),
88         axis.text.y = element_text(face = "bold"),
89         axis.title.x = element_text(face = "bold"),
90         axis.title.y = element_text(face = "bold"),
91         plot.title = element_text(face = "bold", size = 13))
92
93
94 # Imprimir gráfico
95 print(grafica_nucleotidos)
96
97 #-----OP435368-----
98
99 # Definir colores
100 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
101
102 # Crear gráfico de barras
103 grafica_nucleotidos2 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = OP435368, fill = rownames(nucleotidos))) +
104   geom_bar(stat = "identity") +
105   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma OP435368") +
106   scale_x_discrete(labels = c("A", "C", "G", "T")) +
107   scale_fill_manual(values = colores) +
108   theme(axis.text.x = element_text(face = "bold"),
109         axis.text.y = element_text(face = "bold"),
110         axis.title.x = element_text(face = "bold"),
111         axis.title.y = element_text(face = "bold"),
112         plot.title = element_text(face = "bold", size = 13))
113
114 # Imprimir gráfico
115 print(grafica_nucleotidos2)
116
117 #-----0Q918256-----
118
119 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
120
121 # Crear gráfico de barras
122 grafica_nucleotidos3 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = 0Q918256, fill = rownames(nucleotidos))) +
123   geom_bar(stat = "identity") +
124   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma 0Q918256") +
125   scale_x_discrete(labels = c("A", "C", "G", "T")) +
126   scale_fill_manual(values = colores) +
127   theme(axis.text.x = element_text(face = "bold"),
128         axis.text.y = element_text(face = "bold"),
129         axis.title.x = element_text(face = "bold"),
130         axis.title.y = element_text(face = "bold"),
131         plot.title = element_text(face = "bold", size = 13))
132
133 # Imprimir gráfico
134 print(grafica_nucleotidos3)
135
136 #-----BS007312-----
137
138 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
139
140 # Crear gráfico de barras
141 grafica_nucleotidos4 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = BS007312, fill = rownames(nucleotidos))) +
142   geom_bar(stat = "identity") +
143   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma BS007312") +
144   scale_x_discrete(labels = c("A", "C", "G", "T")) +
145   scale_fill_manual(values = colores) +
146   theme(axis.text.x = element_text(face = "bold"),
147         axis.text.y = element_text(face = "bold"),
148         axis.title.x = element_text(face = "bold"),
149         axis.title.y = element_text(face = "bold"),
150         plot.title = element_text(face = "bold", size = 13))
151
152
153 # Imprimir gráfico
154 print(grafica_nucleotidos4)
```

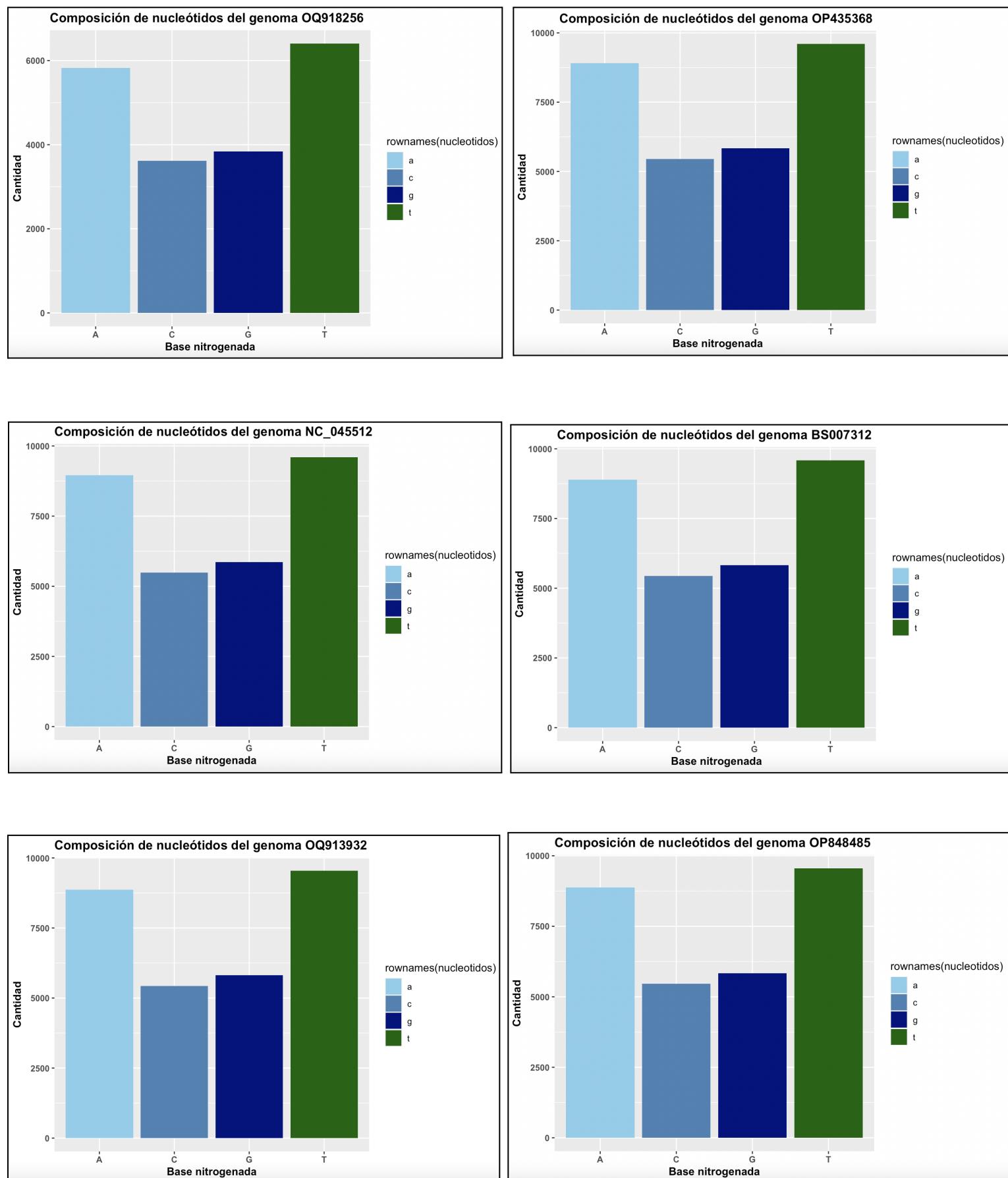
```

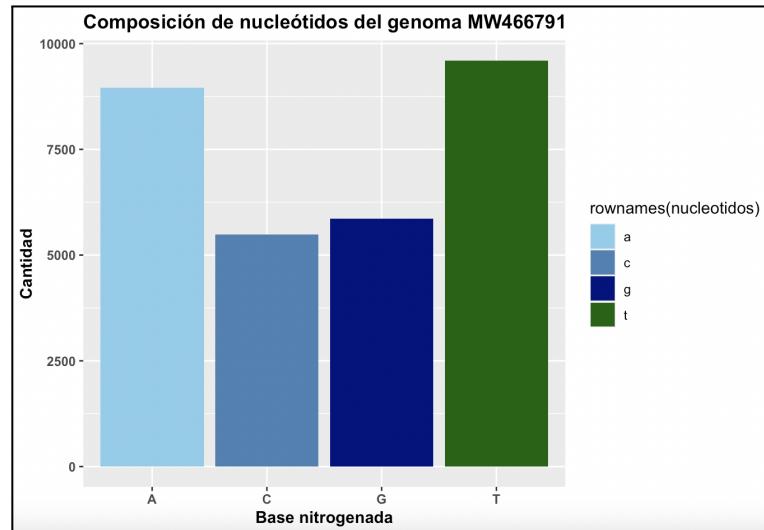
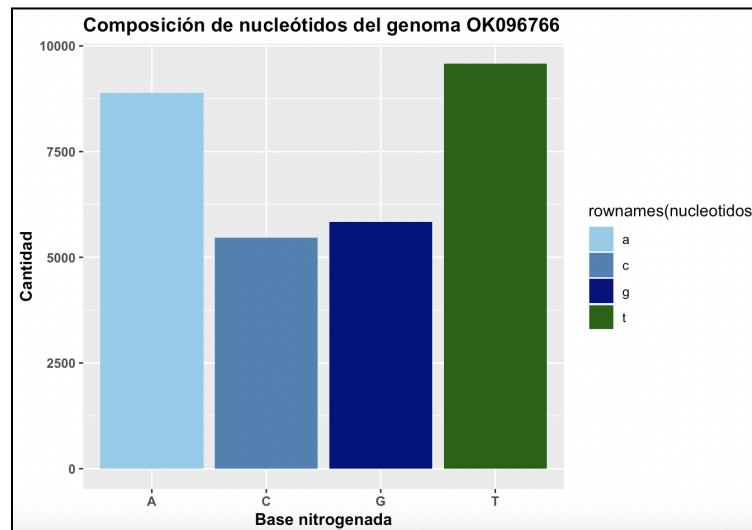
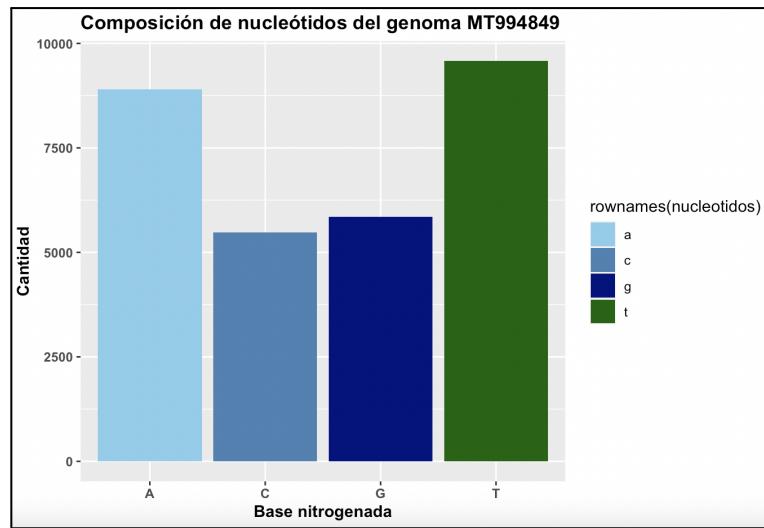
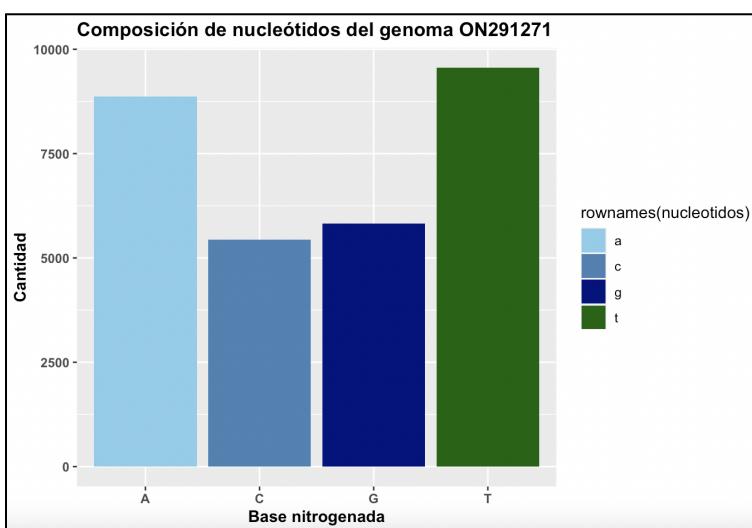
176 #-----OP848485-----
177 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
178
179 # Crear gráfico de barras
180 grafica_nucleotidos6 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = OP848485, fill = rownames(nucleotidos))) +
181   geom_bar(stat = "identity") +
182   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma OP848485") +
183   scale_x_discrete(labels = c("A", "C", "G", "T")) +
184   scale_fill_manual(values = colores) +
185   theme(axis.text.x = element_text(face = "bold"),
186         axis.text.y = element_text(face = "bold"),
187         axis.title.x = element_text(face = "bold"),
188         axis.title.y = element_text(face = "bold"),
189         plot.title = element_text(face = "bold", size = 13))
190
191 # Imprimir gráfico
192 print(grafica_nucleotidos6)
193
194 #-----ON291271-----
195 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
196
197 # Crear gráfico de barras
198 grafica_nucleotidos7 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = ON291271, fill = rownames(nucleotidos))) +
199   geom_bar(stat = "identity") +
200   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma ON291271") +
201   scale_x_discrete(labels = c("A", "C", "G", "T")) +
202   scale_fill_manual(values = colores) +
203   theme(axis.text.x = element_text(face = "bold"),
204         axis.text.y = element_text(face = "bold"),
205         axis.title.x = element_text(face = "bold"),
206         axis.title.y = element_text(face = "bold"),
207         plot.title = element_text(face = "bold", size = 13))
208
209 # Imprimir gráfico
210 print(grafica_nucleotidos7)
211
212
213 #-----MT994849-----
214 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
215
216 # Crear gráfico de barras
217 grafica_nucleotidos8 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = MT994849, fill = rownames(nucleotidos))) +
218   geom_bar(stat = "identity") +
219   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma MT994849") +
220   scale_x_discrete(labels = c("A", "C", "G", "T")) +
221   scale_fill_manual(values = colores) +
222   theme(axis.text.x = element_text(face = "bold"),
223         axis.text.y = element_text(face = "bold"),
224         axis.title.x = element_text(face = "bold"),
225         axis.title.y = element_text(face = "bold"),
226         plot.title = element_text(face = "bold", size = 13))
227
228 # Imprimir gráfico
229 print(grafica_nucleotidos8)
230
231
232 #-----OK096766-----
233 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
234
235 # Crear gráfico de barras
236 grafica_nucleotidos9 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = OK096766, fill = rownames(nucleotidos))) +
237   geom_bar(stat = "identity") +
238   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma OK096766") +
239   scale_x_discrete(labels = c("A", "C", "G", "T")) +
240   scale_fill_manual(values = colores) +
241   theme(axis.text.x = element_text(face = "bold"),
242         axis.text.y = element_text(face = "bold"),
243         axis.title.x = element_text(face = "bold"),
244         axis.title.y = element_text(face = "bold"),
245         plot.title = element_text(face = "bold", size = 13))
246
247 # Imprimir gráfico
248 print(grafica_nucleotidos9)
249
250

```

```
257 * #-----MW466791-----
258
259 colores <- c("skyblue", "steelblue", "navy", "darkgreen")
260
261 # Crear gráfico de barras
262 grafica_nucleotidos10 <- ggplot(nucleotidos, aes(x = rownames(nucleotidos), y = MW466791, fill = rownames(nucleotidos))) +
263   geom_bar(stat = "identity") +
264   labs(x = "Base nitrogenada", y = "Cantidad", title = "Composición de nucleótidos del genoma MW466791") +
265   scale_x_discrete(labels = c("A", "C", "G", "T")) +
266   scale_fill_manual(values = colores) +
267   theme(axis.text.x = element_text(face = "bold"),
268         axis.text.y = element_text(face = "bold"),
269         axis.title.x = element_text(face = "bold"),
270         axis.title.y = element_text(face = "bold"),
271         plot.title = element_text(face = "bold", size = 13))
272
273
274 # Imprimir gráfico
275 print(grafica_nucleotidos10)
```

Resultado en consola:





7. Agrega una interpretación escrita, desde el punto de vista biológico, para la gráfica anterior.

Las gráficas muestran la composición de nucleótidos en 10 genomas secuenciados (todos provenientes del virus SARS-CoV-2). En general se observa que las cantidades de las bases nitrogenadas cambian ligeramente. Aunque hay pequeñas variaciones en la cantidad de cada tipo de nucleótido, la composición general sigue siendo similar en todos los genomas.

En resumen, estos resultados sugieren que la estructura genética del Sars-Cov-2 no ha sufrido cambios significativos en su composición de nucleótidos recientemente. La alta proporción de Timina y Adenina tiene relación con el cómo se adapta el virus al entrar a las células de las personas infectadas, y también se relaciona con la capacidad del virus a su capacidad para evadir la respuesta inmunológica de quien es infectado. Sin embargo, es importante tener en cuenta que la información proporcionada se refiere solo a la composición de nucleótidos y no a la función de los genes individuales, que pueden haber experimentado cambios o mutaciones en el tiempo.

8. Calcula el %GC de cada variante. Muestra el resultado impreso en consola y el código que utilizaste para realizar este punto.

Código:

```

273 #EJERCICIO 8
274 #calcula el %GC de cada variante
275
276 NC_045512_GC <- GC(virus_sequences_character[[1]])*100
277 OP435368_GC <- GC(virus_sequences_character[[2]])*100
278 OQ918256_GC <- GC(virus_sequences_character[[3]])*100
279 BS007312_GC <- GC(virus_sequences_character[[4]])*100
280 OQ913932_GC <- GC(virus_sequences_character[[5]])*100
281 OP848485_GC <- GC(virus_sequences_character[[6]])*100
282 ON291271_GC <- GC(virus_sequences_character[[7]])*100
283 MT994849_GC <- GC(virus_sequences_character[[8]])*100
284 OK096766_GC <- GC(virus_sequences_character[[9]])*100
285 MW466791_GC <- GC(virus_sequences_character[[10]])*100
286
287 gc_porcentaje <- c(NC_045512_GC, OP435368_GC, OQ918256_GC, BS007312_GC, OQ913932_GC, OP848485_GC, ON291271_GC, MT994849_GC, OK096766_GC, MW466791_GC)
288 genomas <- c("NC_045512", "OP435368", "OQ918256", "BS007312", "OQ913932", "OP848485", "ON291271", "MT994849", "OK096766", "MW466791")
289 gc_final <- data.frame(Genoma = genomas, GC_Porcentaje = gc_porcentaje)
290
291 gc_final
292

```

Resultado en consola:

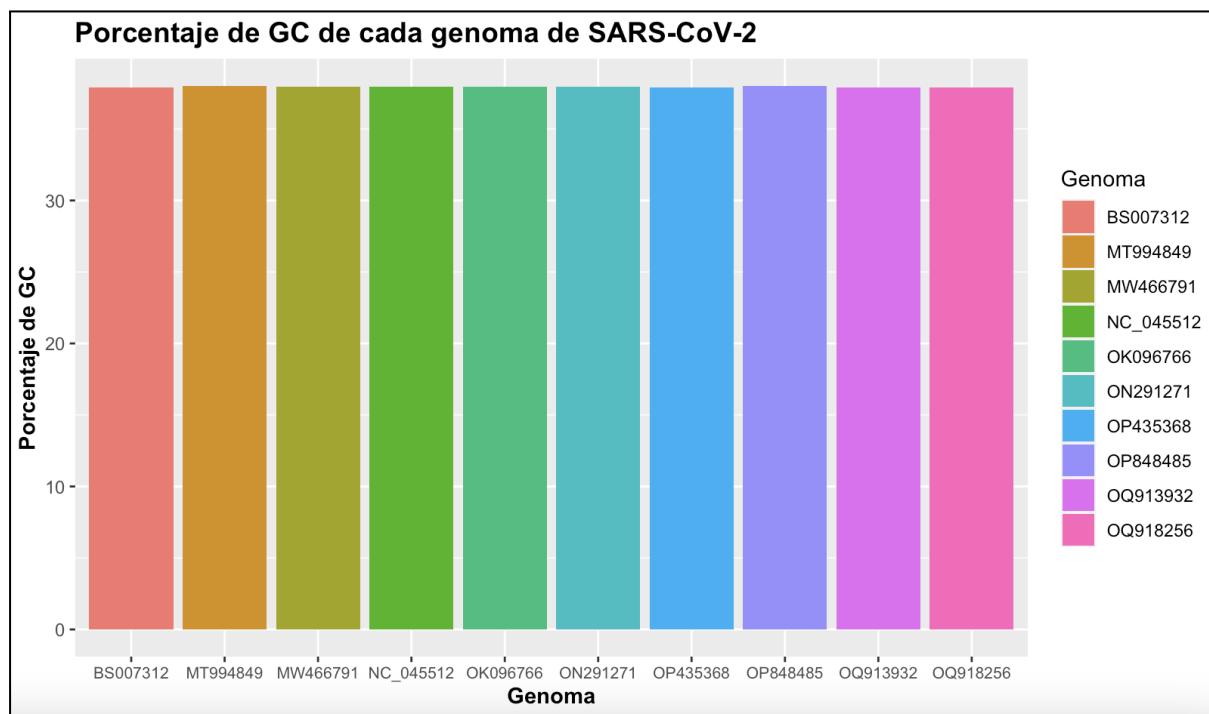
> gc_final		
	Genoma	GC_Porcentaje
1	NC_045512	37.97278
2	OP435368	37.89516
3	OQ918256	37.88287
4	BS007312	37.87874
5	OQ913932	37.91487
6	OP848485	38.00020
7	ON291271	37.94799
8	MT994849	38.01603
9	OK096766	37.96278
10	MW466791	37.97405

9. Genera una gráfica en la que se visualice el %GC de los 10 genomas virales que analizaste. Muestra el código que utilizaste para realizar este punto.

Código:

```
294 #EJERCICIO 9
295
296 ggplot(gc_final, aes(x = Genoma, y = GC_Porcentaje, fill = Genoma)) +
297   geom_bar(stat = "identity") +
298   labs(x = "Genoma", y = "Porcentaje de GC", title = "Porcentaje de GC de cada genoma de SARS-CoV-2") +
299   theme(axis.text.x = element_text(size = 8),
300         axis.text.y = element_text(size = 9),
301         axis.title.x = element_text(face = "bold"),
302         axis.title.y = element_text(face = "bold"),
303         plot.title = element_text(face = "bold", size = 14))
304
```

Resultado en consola:



10. Agrega una interpretación escrita, desde el punto de vista biológico, para la gráfica anterior.

La gráfica muestra una composición de nucleótidos GC (guanina y citosina) en el virus SARS-COV-2. Al observar la gráfica nos damos cuenta que los valores de Citosina y Guanina casi no cambian. Por ejemplo el valor del porcentaje de GC del primer adn del covid-19 encontrado es de 37.97278%, y el más lejano es de 38.00020%, el cual su país proveniente es Irán. En base a esta información se puede inferir que la zona geográfica puede llegar a tener un impacto, sin embargo el cambio sería menos del 1%.

En caso de que el virus tuviera alteraciones en los porcentajes de GC, esto indicaría una evolución más drástica y podría generar nuevas capacidades de transmisión y alojamiento, lo que dificultará el desarrollo de tratamientos efectivos. Por lo tanto, la estabilidad en la composición de nucleótidos GC es una ventaja para la humanidad, en cuanto al control y tratamiento del SARS-COV-2.

11. Crea una tabla (marco de datos) en la que se incluya el nombre común de cada virus, el ID (identificador en base de datos) de cada genoma, la longitud y el contenido GC de cada genoma. Muestra el resultado impreso en consola y el código que utilizaste para realizar este punto.

Código:

```

306 #EJERCICIO 11
307
308 virus_nombre <- rep("SARS-coV-2",10)
309 pais <- c("China","Finlandia","India","Japón","Estdos Unidos","Australia","Francia","Iran","Alemania","Corea del Sur")
310 gc_genomas <- sapply(virus_sequences_character, function(x) GC(x) * 100)
311
312 dataFrameFinal <- data.frame(Virus = virus_nombre, ID = corona_virus,País_de_Origen = pais, Longitud = longitud, Porcentaje_GC = gc_genomas)
313
314 print(dataFrameFinal, row.names = FALSE)
315

```

Resultado en consola:

> print(dataFrameFinal, row.names = FALSE)	Virus	ID	País_de_Origen	Longitud	Porcentaje_GC
	SARS-coV-2	NC_045512	China	29903	37.97278
	SARS-coV-2	OP435368	Finlandia	29799	37.89516
	SARS-coV-2	0Q918256	India	29010	37.88287
	SARS-coV-2	BS007312	Japón	29737	37.87874
	SARS-coV-2	0Q913932	Estdos Unidos	29660	37.91487
	SARS-coV-2	OP848485	Australia	29714	38.00020
	SARS-coV-2	ON291271	Francia	29689	37.94799
	SARS-coV-2	MT994849	Iran	29819	38.01603
	SARS-coV-2	OK096766	Alemania	29766	37.96278
	SARS-coV-2	MW466791	Corea del Sur	29902	37.97405

Conclusión

Las herramientas de biología computacional para el análisis de virus son esenciales para examinar sus características y componentes principales, para de esta forma llegar a conclusiones certeras acerca de cómo evolucionan y/o se adaptan. Al terminar este trabajo, se observó que a pesar de las variaciones que existen en los diferentes países del mundo, la composición general del genoma se mantiene bastante similar. Con esta información podemos concluir que el virus del COVID-19, aunque ha habido algunas mutaciones en la secuencia de su ADN, en general se ha mantenido relativamente constante.

Estas nuevas tecnologías que facilitan el estudio de virus son extremadamente importantes para la vida en la actualidad, ya que permite encontrar nuevos tratamientos y diagnósticos para enfermedades como lo es el SARS-CoV-2, el cual ha ocasionado tantas muertes a nivel mundial.

Referencias:

- "Virus en Biología". Autor: Equipo editorial, Etecé. De: Argentina. Para: *Concepto.de*. Disponible en: <https://concepto.de/virus-en-biologia/>. Última edición: 5 de agosto de 2021. Consultado: 02 de mayo de 2023 Fuente: <https://concepto.de/virus-en-biologia/#ixzz80b68fAYF>
- Fernández-Pérez, G. C., Miranda, M. O., Fernández-Rodríguez, P., Casares, M. V., De La Calle, M. C., López, Á. F., Blanco, M., & Cuchat, J. M. O. (2021). SARS-CoV-2: cómo es, cómo actúa y cómo se expresa en la imagen. *Radiología*, 63(2), 115–126. <https://doi.org/10.1016/j.rx.2020.10.006>
- *COVID-19 Map - Johns Hopkins Coronavirus Resource Center*. (n.d.). Johns Hopkins Coronavirus Resource Center. <https://coronavirus.jhu.edu/map.html>
- *Situación de COVID-19 : Secretaría de Salud – COVID 19*. (n.d.). http://covid19.saludsonora.gob.mx/?page_id=2339
- Aguilar-Gamboa, F. R., Suclupe-Campos, D. O., Vega-Fernández, J. A., & Silva-Díaz, H. (2021). Diversidad genómica en SARS-CoV-2: Mutaciones y variantes. *REVISTA DEL CUERPO MÉDICO HOSPITAL NACIONAL ALMANZOR AGUINAGA ASENJO*, 14, 572–582. <https://doi.org/10.35434/rcmhnaaa>
- *Enfermedad del coronavirus 2019 (COVID-19)*. (2020, February 11). Centers for Disease Control and Prevention. <https://espanol.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html>
- Lam, T. T., Jia, N., Zhang, Y. W., Shum, M. H., Jiang, J. F., Zhu, H. C., ... & Holmes, E. C. (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, 583(7815), 282-285.

- Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., ... & Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270-273.
- *Diccionario de cáncer del NCI*. (n.d.-b). Instituto Nacional Del Cáncer. <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/biología-computacional>
- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature Medicine*, 26(4), 450–452. <https://doi.org/10.1038/s41591-020-0820-9>