

analisis_inicial

cyrce salinas

2024-04-28

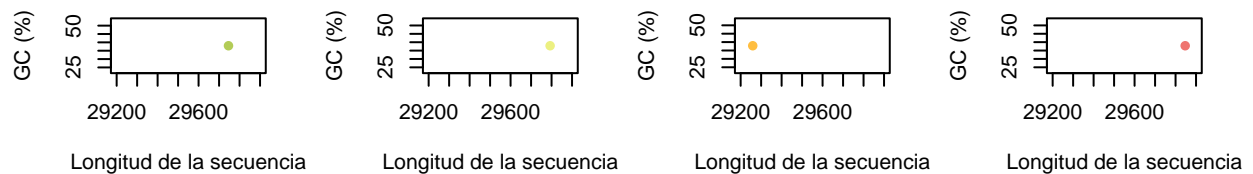
reading files

##	viruses	size	a	c	g	t	gc_percent
## 1	B16401	29746	29.84065	18.36452	19.61295	32.18189	37.97747
## 2	BA2752	29793	29.87678	18.30180	19.59464	32.22679	37.89644
## 3	BA286	29259	29.88360	18.22321	19.60288	32.29031	37.82609
## 4	EG51	29847	29.95611	18.28995	19.56981	32.18414	37.85975
## 5	JN1	29478	29.87991	18.26447	19.60106	32.25456	37.86553
## 6	XBB191	29685	29.88715	18.31565	19.61597	32.18124	37.93162
## 7	XBB116	29502	29.89628	18.29029	19.59867	32.21477	37.88896
## 8	XBB23	29706	29.89160	18.29720	19.60342	32.20778	37.90062
## 9	ZETA	29782	29.86703	18.35337	19.62931	32.15029	37.98267
## 10	ALFA	29734	29.84944	18.35360	19.62343	32.17353	37.97703

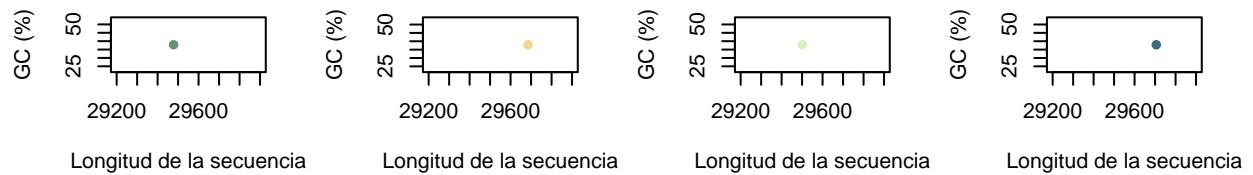
```
## Secuencia: B16401
## Complementaria:
## CAACCAACTT ... AATTTTAGTA
##
## Secuencia: BA2752
## Complementaria:
## TAAAGGTTTA ... CATGTGATT
##
## Secuencia: BA286
## Complementaria:
## ATAATAATT ... CCTTACCGCA
##
## Secuencia: EG51
## Complementaria:
## ACTAAAGGTT ... AAAAAAAAAA
##
## Secuencia: JN1
## Complementaria:
## CTGCATGCTT ... GTCTACTCTT
##
## Secuencia: XBB191
## Complementaria:
## CGGCTGCATG ... CTATCCCAT
##
## Secuencia: XBB116
## Complementaria:
## TCGTCCGTGT ... CATTTTCACC
```

```
##
## Secuencia: XBB23
## Complementaria:
## AAAATCTGTG ... CTATCCCCAT
##
## Secuencia: ZETA
## Complementaria:
## AGATCTGTTC ... AGTGCTATCC
##
## Secuencia: ALFA
## Complementaria:
## ACTTTCGATC ... TGCTAGGGAG
```

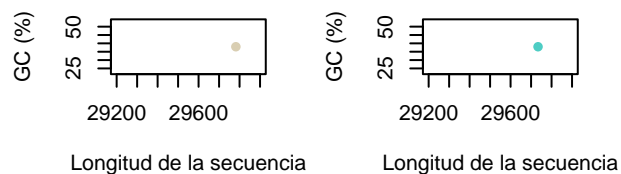
GC% vs. Longitud – B16 GC% vs. Longitud – BA2 GC% vs. Longitud – BA1 GC% vs. Longitud – EG

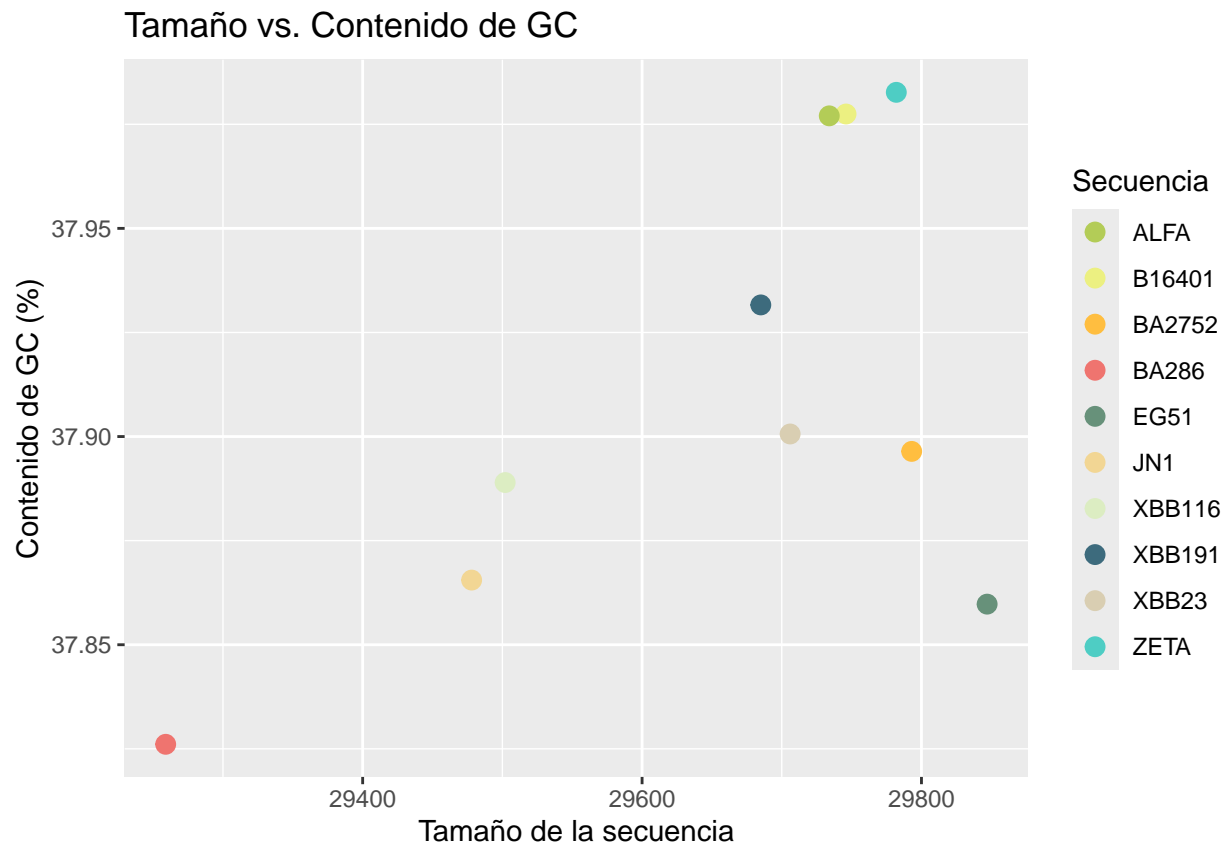


GC% vs. Longitud – JN GC% vs. Longitud – XBE GC% vs. Longitud – XBI



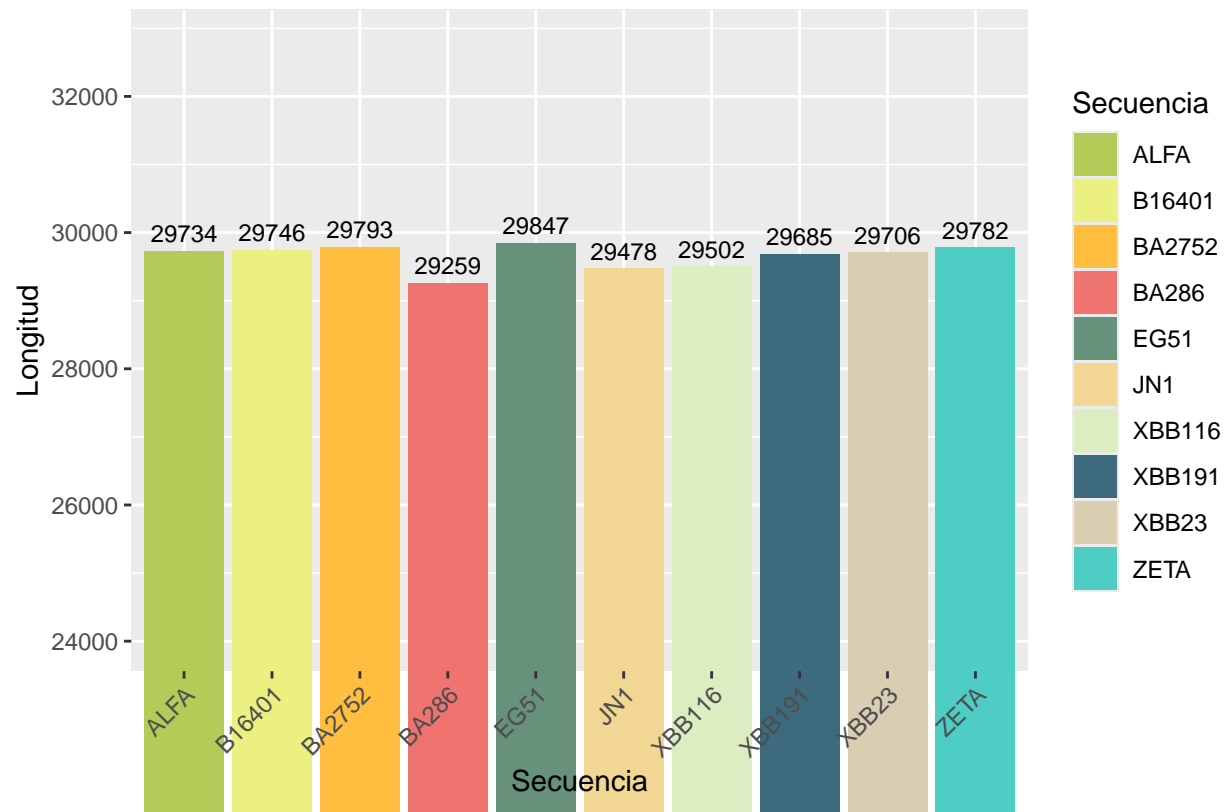
GC% vs. Longitud – ZE GC% vs. Longitud – AL



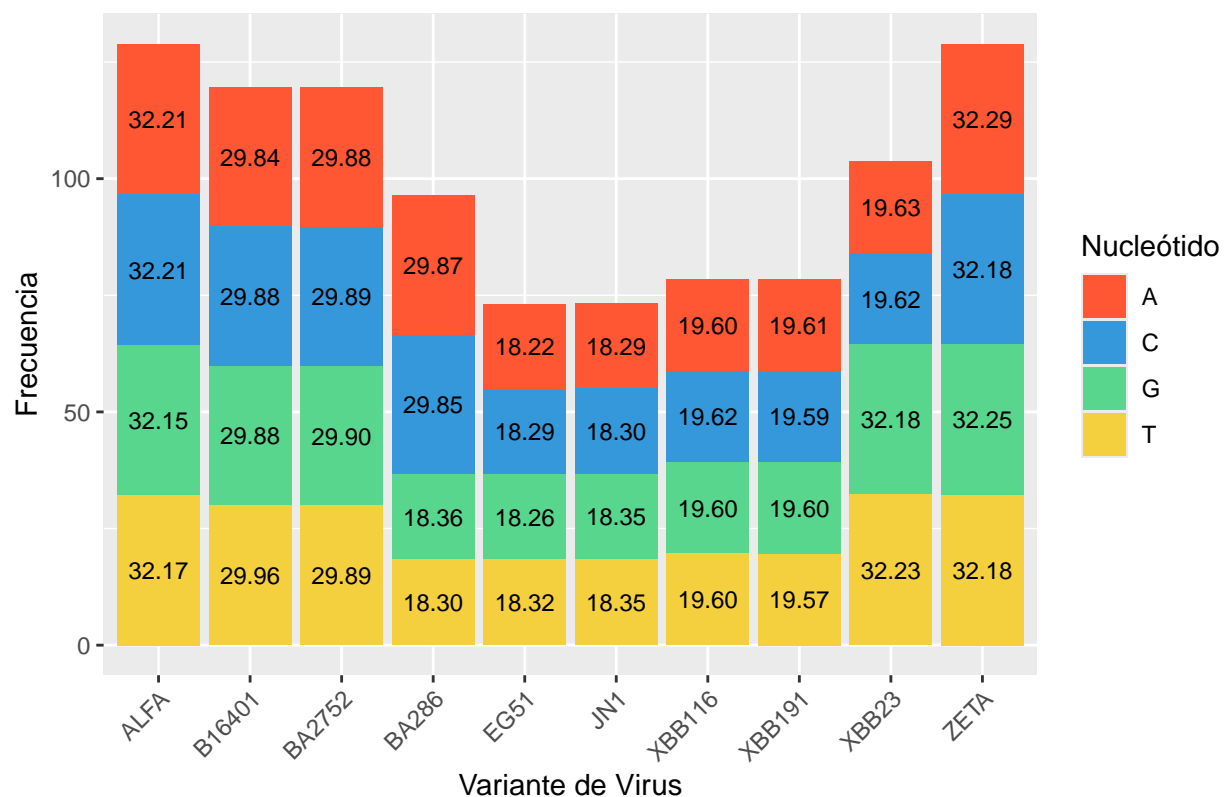


```
## [1] "Como conclusion, puedo decir que la tendencia es que cuan más grande es la secuencia,"
## [1] "el contenido de GC aumenta, lo cual suena razonable. Sin embargo como toda regla,"
## [1] "tiene su excepcion como lo arroja la variante XBB.1.9.1 y la variante BA.2.75.2. "
## [1] "Aunado a esto, me gustaria recalcar que las variantes VOI tienden a tener una "
## [1] "longitud mayor a las que tienen otra clasificacion asignada por la SIG. "
```

Longitud de las secuencias de ADN



Frecuencia de Nucleótidos por Variante de Virus



[1] "Por último, me gustaria agregar que las variables estudiadas ALFA, ZETA "

[1] "son en las que debemos enfocarnos mas, estas concuerdan en estar en la clasificacion de VOC,"

[1] " por lo que su alto porcentaje en "

[1] "nucleotodidos A puede tener relacion"