Basesdatosvirus

Pedro Sotelo Arce y Leonardo Pequeño Moreno

2023-04-24

Zika virus

library(seqinr)  
library(ggplot2)  
  
cat("Zika")

## Zika

zika = read.fasta("zika.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(zika[[1]])

## [1] 10794

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(zika[[1]],1)

##   
## a c g t   
## 2991 2359 3139 2305

cat("Contenido GC: ")

## Contenido GC:

GC(zika[[1]],2)

## [1] 0.5093571

cat("virus original(zika): ",head((zika[[1]]),40))

## virus original(zika): a g t t g t t g a t c t g t g t g a g t c a g a c t g c g a c a g t t c g a g t

cat("virus original (zika, final): ",tail((zika[[1]]),40))

## virus original (zika, final): c c g a a c t t c g g c g g c c g g t g t g g g g a a a t c c a t g g t t t c t

cat("sequencia complementaria(zika):",head(comp(zika[[1]]),40))

## sequencia complementaria(zika): t c a a c a a c t a g a c a c a c t c a g t c t g a c g c t g t c a a g c t c a

cat("sequencia complementaria(zika, final):",tail(comp(zika[[1]]),40))

## sequencia complementaria(zika, final): g g c t t g a a g c c g c c g g c c a c a c c c c t t t a g g t a c c a a a g a

SARS coronavirus

cat("Sarscovid")

## Sarscovid

sarscovid = read.fasta("sarscovid.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(sarscovid[[1]])

## [1] 29903

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(sarscovid[[1]],1)

##   
## a c g t   
## 8954 5492 5863 9594

cat("Contenido GC: ")

## Contenido GC:

GC(sarscovid[[1]],2)

## [1] 0.3797278

cat("virus original(sarscovid): ",head((sarscovid[[1]]),40))

## virus original(sarscovid): a t t a a a g g t t t a t a c c t t c c c a g g t a a c a a a c c a a c c a a c

cat("virus original (sarscovid, final): ",tail((sarscovid[[1]]),40))

## virus original (sarscovid, final): g a a t g a c a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a

cat("sequencia complementaria(sarscovid):",head(comp(sarscovid[[1]]),40))

## sequencia complementaria(sarscovid): t a a t t t c c a a a t a t g g a a g g g t c c a t t g t t t g g t t g g t t g

cat("sequencia complementaria(sarscovid, final):",tail(comp(sarscovid[[1]]),40))

## sequencia complementaria(sarscovid, final): c t t a c t g t t t t t t t t t t t t t t t t t t t t t t t t t t t t t t t t t

Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1

cat("Wuhan")

## Wuhan

wuhan = read.fasta("wuhan.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(wuhan[[1]])

## [1] 30473

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(wuhan[[1]],1)

##   
## a c g t   
## 9115 5605 5951 9802

cat("Contenido GC: ")

## Contenido GC:

GC(wuhan[[1]],2)

## [1] 0.3792209

cat("virus original(wuhan): ",head((wuhan[[1]]),40))

## virus original(wuhan): c g g t g a c g c a t a c a a a a c a t t c c c a c c a t a c c t t c c c a g g

cat("virus original (wuhan, final): ",tail((wuhan[[1]]),40))

## virus original (wuhan, final): c t t t t t g t a a a a a t a a g g a a a g c a a g g t t t t t t g a t a a t c

cat("sequencia complementaria(wuhan):",head(comp(wuhan[[1]]),40))

## sequencia complementaria(wuhan): g c c a c t g c g t a t g t t t t g t a a g g g t g g t a t g g a a g g g t c c

cat("sequencia complementaria(wuhan, final):",tail(comp(wuhan[[1]]),40))

## sequencia complementaria(wuhan, final): g a a a a a c a t t t t t a t t c c t t t c g t t c c a a a a a a c t a t t a g

Middle East respiratory syndrome coronavirus

cat("MERS")

## MERS

mers = read.fasta("mers.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(mers[[1]])

## [1] 30108

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(mers[[1]],1)

##   
## a c g t   
## 7890 6087 6301 9830

cat("Contenido GC: ")

## Contenido GC:

GC(mers[[1]],2)

## [1] 0.4114521

cat("virus original(mers): ",head((mers[[1]]),40))

## virus original(mers): g a t t t a a g t g a a t a g c t t g g c t a t c t c a c t t c c c c t c g t t

cat("virus original (mers, final): ",tail((mers[[1]]),40))

## virus original (mers, final): g c t a t t a t t c a a t t a g a t t a g g c t a a t t a g a t g a t t t g c a

cat("sequencia complementaria(mers):",head(comp(mers[[1]]),40))

## sequencia complementaria(mers): c t a a a t t c a c t t a t c g a a c c g a t a g a g t g a a g g g g a g c a a

cat("sequencia complementaria(mers, final):",tail(comp(mers[[1]]),40))

## sequencia complementaria(mers, final): c g a t a a t a a g t t a a t c t a a t c c g a t t a a t c t a c t a a a c g t

DENGUE TIPO 1

cat("Dengue tipo 1")

## Dengue tipo 1

dengue = read.fasta("dengue.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(dengue[[1]])

## [1] 10735

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(dengue[[1]],1)

##   
## a c g t   
## 3421 2194 2801 2319

cat("Contenido GC: ")

## Contenido GC:

GC(dengue[[1]],2)

## [1] 0.4653004

cat("virus original(dengue): ",head((dengue[[1]]),40))

## virus original(dengue): a g t t g t t a g t c t a c g t g g a c c g a c a a g a a c a g t t t c g a a t

cat("virus original (dengue, final): ",tail((dengue[[1]]),40))

## virus original (dengue, final): c g c c a g a a a a t g g a a t g g t g c t g t t g a a t c a a c a g g t t c t

cat("sequencia complementaria(dengue):",head(comp(dengue[[1]]),40))

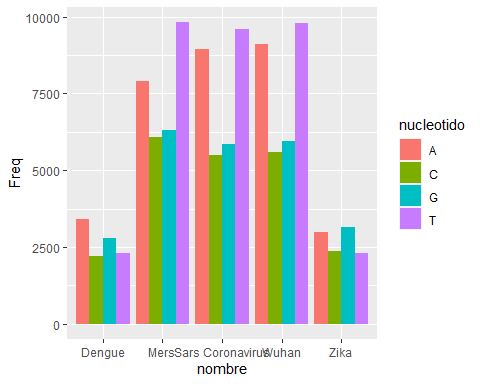
## sequencia complementaria(dengue): t c a a c a a t c a g a t g c a c c t g g c t g t t c t t g t c a a a g c t t a

cat("sequencia complementaria(dengue, final):",tail(comp(dengue[[1]]),40))

## sequencia complementaria(dengue, final): g c g g t c t t t t a c c t t a c c a c g a c a a c t t a g t t g t c c a a g a

GRAFICO

graph<-data.frame(  
 nombre=rep(c("Zika", "Sars Coronavirus", "Mers","Wuhan",  
 "Dengue"),each=4),  
 nucleotido=rep(c("A","C","G",  
 "T"),5),  
 Freq=c(count(zika[[1]],1), count(sarscovid[[1]],1),  
 count(mers[[1]],1), count(wuhan[[1]],1),  
 count(dengue[[1]],1))  
)  
  
ggplot(data = graph) +  
 geom\_bar(aes(x=nombre, y=Freq, fill=nucleotido),  
 position="dodge",stat="identity")



PB2

cat("PB2")

## PB2

PB2 = read.fasta("PB2.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(PB2[[1]])

## [1] 2293

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(PB2[[1]],1)

##   
## a c g t   
## 773 433 589 498

cat("Contenido GC: ")

## Contenido GC:

GC(PB2[[1]],2)

## [1] 0.4457043

cat("virus original(PB2): ",head((PB2[[1]]),40))

## virus original(PB2): t a t g g a g a g a a t a a a a g a a c t g a g a g a t c t a a t g t c g c a g

cat("virus original (PB2, final): ",tail((PB2[[1]]),40))

## virus original (PB2, final): c a a a a g a a t t c g g a t g g c c a t c a a t t a g t g t c g a a t t g t t

cat("sequencia complementaria(PB2):",head(comp(PB2[[1]]),40))

## sequencia complementaria(PB2): a t a c c t c t c t t a t t t t c t t g a c t c t c t a g a t t a c a g c g t c

cat("sequencia complementaria(PB2, final):",tail(comp(PB2[[1]]),40))

## sequencia complementaria(PB2, final): g t t t t c t t a a g c c t a c c g g t a g t t a a t c a c a g c t t a a c a a

PB1

cat("PB1")

## PB1

PB1 = read.fasta("PB1.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(PB1[[1]])

## [1] 2303

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(PB1[[1]],1)

##   
## a c g t   
## 820 443 521 519

cat("Contenido GC: ")

## Contenido GC:

GC(PB1[[1]],2)

## [1] 0.4185845

cat("virus original(PB1): ",head((PB1[[1]]),40))

## virus original(PB1): t t g a a t g g a t g t c a a t c c g a c t c t a c t t t t c c t a a a a a t t

cat("virus original (PB1, final): ",tail((PB1[[1]]),40))

## virus original (PB1, final): a g a c g g c a a a a a t a a t g a a t t t a a c t t g t c c t t c a t g a a a

cat("sequencia complementaria(PB1):",head(comp(PB1[[1]]),40))

## sequencia complementaria(PB1): a a c t t a c c t a c a g t t a g g c t g a g a t g a a a a g g a t t t t t a a

cat("sequencia complementaria(PB1, final):",tail(comp(PB1[[1]]),40))

## sequencia complementaria(PB1, final): t c t g c c g t t t t t a t t a c t t a a a t t g a a c a g g a a g t a c t t t

PA

cat("PA")

## PA

PA = read.fasta("PA.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(PA[[1]])

## [1] 2175

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(PA[[1]],1)

##   
## a c g t   
## 713 428 532 502

cat("Contenido GC: ")

## Contenido GC:

GC(PA[[1]],2)

## [1] 0.4413793

cat("virus original(PA): ",head((PA[[1]]),40))

## virus original(PA): t c c a a a a t g g a a g a c t t t g t g c g a c a a t g c t t c a a t c c a a

cat("virus original (PA, final): ",tail((PA[[1]]),40))

## virus original (PA, final): c c t c a c a c a t g c a c t g a a g t a g t t g t g g c a a t g c t a c t a t

cat("sequencia complementaria(PA):",head(comp(PA[[1]]),40))

## sequencia complementaria(PA): a g g t t t t a c c t t c t g a a a c a c g c t g t t a c g a a g t t a g g t t

cat("sequencia complementaria(PA, final):",tail(comp(PA[[1]]),40))

## sequencia complementaria(PA, final): g g a g t g t g t a c g t g a c t t c a t c a a c a c c g t t a c g a t g a t a

HA

cat("HA")

## HA

HA = read.fasta("HA.fasta")  
cat("Tamaño de la secuencia: ")

## Tamaño de la secuencia:

length(HA[[1]])

## [1] 1734

cat("Composicion de nucleotidos: ")

## Composicion de nucleotidos:

count(HA[[1]],1)

##   
## a c g t   
## 617 319 384 414

cat("Contenido GC: ")

## Contenido GC:

GC(HA[[1]],2)

## [1] 0.405421

cat("virus original(HA): ",head((HA[[1]]),40))

## virus original(HA): a a a a g c a a c a a a a a t g a a g g c a a t a c t a g t a g t t c t g c t a

cat("virus original (HA, final): ",tail((HA[[1]]),40))

## virus original (HA, final): a g t g t a g a a t a t g t a t t t a a c a t t a g g a t t t c a g a a g c a t

cat("sequencia complementaria(HA):",head(comp(HA[[1]]),40))

## sequencia complementaria(HA): t t t t c g t t g t t t t t a c t t c c g t t a t g a t c a t c a a g a c g a t

cat("sequencia complementaria(HA, final):",tail(comp(HA[[1]]),40))

## sequencia complementaria(HA, final): t c a c a t c t t a t a c a t a a a t t g t a a t c c t a a a g t c t t c g t a

NP

H1N15 = read.fasta("H1N15.fasta")  
print("Tamaño de la secuencia")

## [1] "Tamaño de la secuencia"

print(length(H1N15[[1]]))

## [1] 1523

print("Composicion de nucleotidos")

## [1] "Composicion de nucleotidos"

print(count(H1N15[[1]],1))

##   
## a c g t   
## 505 301 400 317

print("Contenido GC")

## [1] "Contenido GC"

print(GC(H1N15[[1]])\*100)

## [1] 46.02758

print("Secuencia")

## [1] "Secuencia"

print(H1N15[[1]][c(1:10,(length(H1N15[[1]])-10):length(H1N15[[1]]))])

## [1] "t" "c" "a" "c" "t" "c" "a" "a" "t" "g" "t" "g" "a" "c" "a" "g" "t" "t" "g"  
## [20] "a" "g"

print("Secuencia complementaria")

## [1] "Secuencia complementaria"

print(comp(H1N15[[1]])[c(1:10,(length(H1N15[[1]])-10):length(H1N15[[1]]))])

## [1] "a" "g" "t" "g" "a" "g" "t" "t" "a" "c" "a" "c" "t" "g" "t" "c" "a" "a" "c"  
## [20] "t" "c"

NA

H1N16 = read.fasta("H1N16.fasta")  
print("Tamaño de la secuencia")

## [1] "Tamaño de la secuencia"

print(length(H1N16[[1]]))

## [1] 480

print("Composicion de nucleotidos")

## [1] "Composicion de nucleotidos"

print(count(H1N16[[1]],1))

##   
## a c g t   
## 153 99 105 123

print("Contenido GC")

## [1] "Contenido GC"

print(GC(H1N16[[1]])\*100)

## [1] 42.5

print("Secuencia")

## [1] "Secuencia"

print(H1N16[[1]][c(1:10,(length(H1N16[[1]])-10):length(H1N16[[1]]))])

## [1] "t" "t" "g" "t" "c" "a" "t" "g" "a" "g" "c" "c" "c" "t" "a" "a" "t" "t" "a"  
## [20] "t" "c"

print("Secuencia complementaria")

## [1] "Secuencia complementaria"

print(comp(H1N16[[1]])[c(1:10,(length(H1N16[[1]])-10):length(H1N16[[1]]))])

## [1] "a" "a" "c" "a" "g" "t" "a" "c" "t" "c" "g" "g" "g" "a" "t" "t" "a" "a" "t"  
## [20] "a" "g"

MP

H1N17 = read.fasta("H1N17.fasta")  
print("Tamaño de la secuencia")

## [1] "Tamaño de la secuencia"

print(length(H1N17[[1]]))

## [1] 987

print("Composicion de nucleotidos")

## [1] "Composicion de nucleotidos"

print(count(H1N17[[1]],1))

##   
## a c g t   
## 289 208 254 236

print("Contenido GC")

## [1] "Contenido GC"

print(GC(H1N17[[1]])\*100)

## [1] 46.80851

print("Secuencia")

## [1] "Secuencia"

print(H1N17[[1]][c(1:10,(length(H1N17[[1]])-10):length(H1N17[[1]]))])

## [1] "t" "a" "a" "a" "g" "a" "t" "g" "a" "g" "a" "g" "c" "t" "a" "g" "a" "g" "t"  
## [20] "a" "a"

print("Secuencia complementaria")

## [1] "Secuencia complementaria"

print(comp(H1N17[[1]])[c(1:10,(length(H1N17[[1]])-10):length(H1N17[[1]]))])

## [1] "a" "t" "t" "t" "c" "t" "a" "c" "t" "c" "t" "c" "g" "a" "t" "c" "t" "c" "a"  
## [20] "t" "t"

NS

H1N18 = read.fasta("H1N18.fasta")  
print("Tamaño de la secuencia")

## [1] "Tamaño de la secuencia"

print(length(H1N18[[1]]))

## [1] 850

print("Composicion de nucleotidos")

## [1] "Composicion de nucleotidos"

print(count(H1N18[[1]],1))

##   
## a c g t   
## 271 167 201 211

print("Contenido GC")

## [1] "Contenido GC"

print(GC(H1N18[[1]])\*100)

## [1] 43.29412

print("Secuencia")

## [1] "Secuencia"

print(H1N18[[1]][c(1:10,(length(H1N18[[1]])-10):length(H1N18[[1]]))])

## [1] "a" "a" "c" "a" "t" "a" "a" "t" "g" "g" "t" "t" "t" "a" "a" "t" "g" "a" "t"  
## [20] "a" "a"

print("Secuencia complementaria")

## [1] "Secuencia complementaria"

print(comp(H1N18[[1]])[c(1:10,(length(H1N18[[1]])-10):length(H1N18[[1]]))])

## [1] "t" "t" "g" "t" "a" "t" "t" "a" "c" "c" "a" "a" "a" "t" "t" "a" "c" "t" "a"  
## [20] "t" "t"