



Tecnológico de Monterrey

Evidencia 1

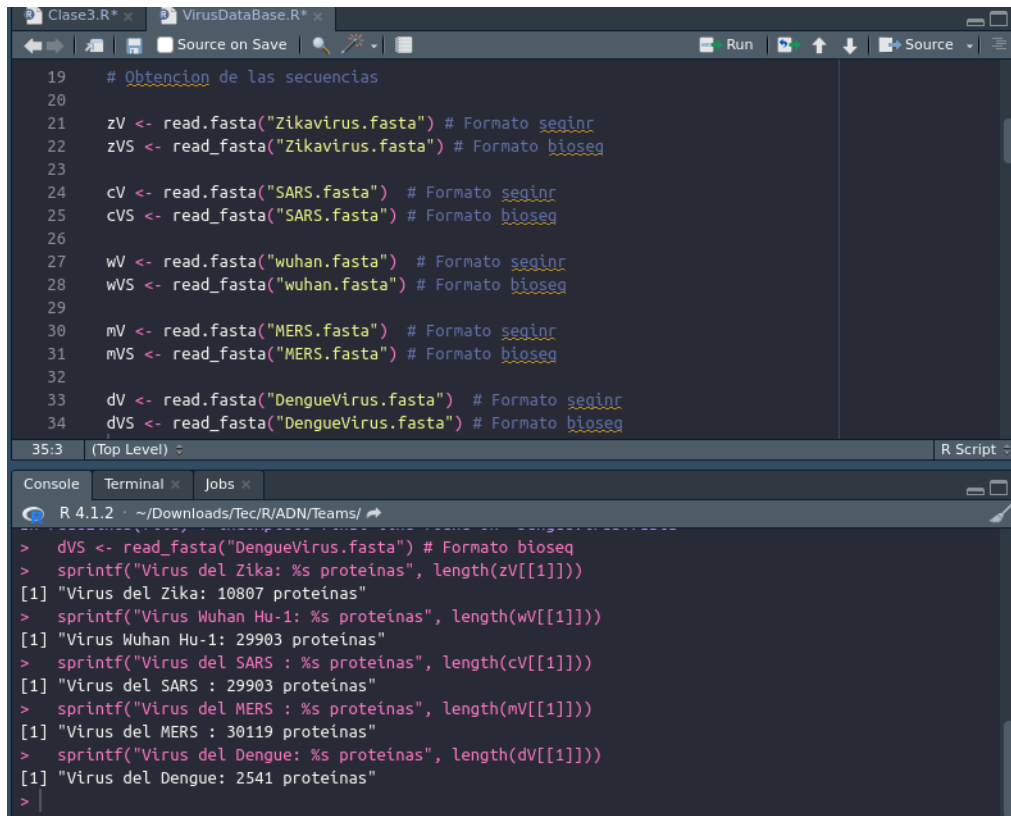
Guillermo Esquivel Ortiz A01625621

06 de Mayo del 2022

Análisis de biología computacional

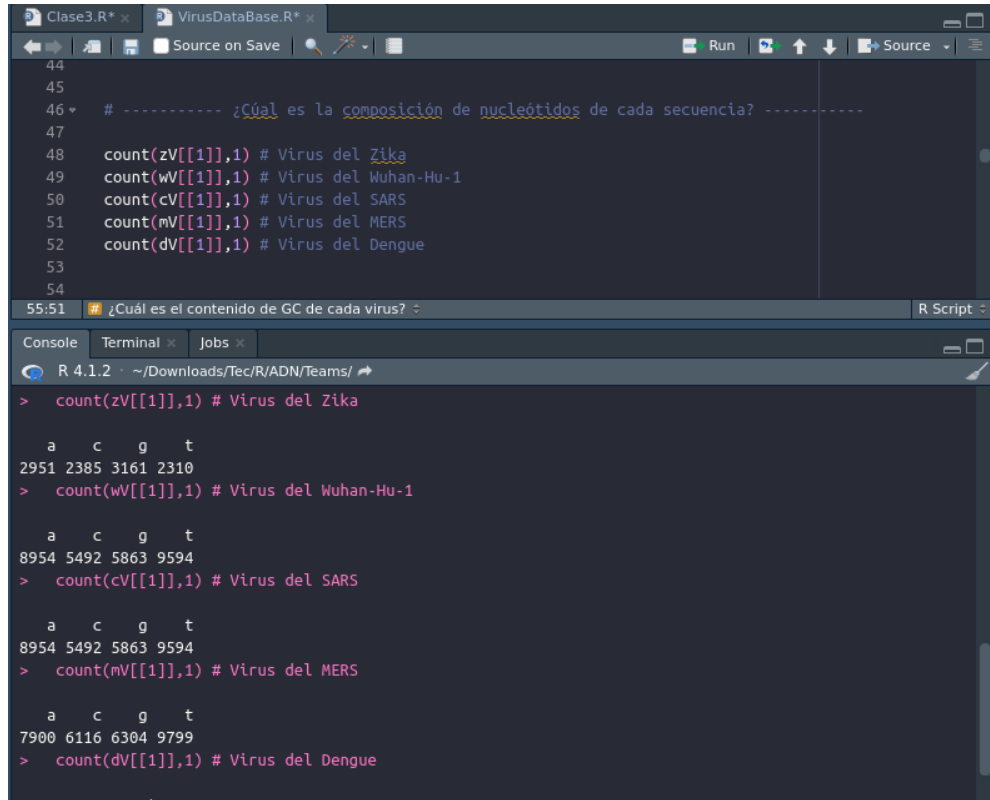
Carlos Pérez Almazán

Proceso de obtención de las secuencias y cuenta de el tamaño del número de proteínas de cada una



```
19 # Obtencion de las secuencias
20
21 zV <- read.fasta("Zikavirus.fasta") # Formato seginr
22 zVS <- read_fasta("Zikavirus.fasta") # Formato bioseq
23
24 cV <- read.fasta("SARS.fasta") # Formato seginr
25 cVS <- read_fasta("SARS.fasta") # Formato bioseq
26
27 wV <- read.fasta("wuhan.fasta") # Formato seginr
28 wVS <- read_fasta("wuhan.fasta") # Formato bioseq
29
30 mV <- read.fasta("MERS.fasta") # Formato seginr
31 mVS <- read_fasta("MERS.fasta") # Formato bioseq
32
33 dV <- read.fasta("DengueVirus.fasta") # Formato seginr
34 dVS <- read_fasta("DengueVirus.fasta") # Formato bioseq
35:3 (Top Level) : R Script
Console Terminal Jobs
R 4.1.2 ~./Downloads/Tec/R/ADN/Teams/
> dVS <- read_fasta("DengueVirus.fasta") # Formato bioseq
> sprintf("Virus del Zika: %s proteínas", length(zV[[1]]))
[1] "Virus del Zika: 10807 proteínas"
> sprintf("Virus Wuhan Hu-1: %s proteínas", length(wV[[1]]))
[1] "Virus Wuhan Hu-1: 29903 proteínas"
> sprintf("Virus del SARS : %s proteínas", length(cV[[1]]))
[1] "Virus del SARS : 29903 proteínas"
> sprintf("Virus del MERS : %s proteínas", length(mV[[1]]))
[1] "Virus del MERS : 30119 proteínas"
> sprintf("Virus del Dengue: %s proteínas", length(dV[[1]]))
[1] "Virus del Dengue: 2541 proteínas"
>
```

Composición de nucleótidos de cada secuencia



```
Clase3.R* VirusDataBase.R*
Source on Save Run Source
44
45
46 # ----- ¿Cuál es la composición de nucleótidos de cada secuencia? -----
47
48 count(zV[[1]],1) # Virus del Zika
49 count(wV[[1]],1) # Virus del Wuhan-Hu-1
50 count(cV[[1]],1) # Virus del SARS
51 count(mV[[1]],1) # Virus del MERS
52 count(dV[[1]],1) # Virus del Dengue
53
54
55:51 # ¿Cuál es el contenido de GC de cada virus?
R Script

Console Terminal Jobs
R 4.1.2 ~/Downloads/Tec/R/ADN/Teams/
> count(zV[[1]],1) # Virus del Zika

  a   c   g   t
2951 2385 3161 2310
> count(wV[[1]],1) # Virus del Wuhan-Hu-1

  a   c   g   t
8954 5492 5863 9594
> count(cV[[1]],1) # Virus del SARS

  a   c   g   t
8954 5492 5863 9594
> count(mV[[1]],1) # Virus del MERS

  a   c   g   t
7900 6116 6304 9799
> count(dV[[1]],1) # Virus del Dengue
```

Contenido de Guanina-Citosina de cada virus

```
Clase3.R* x VirusDataBase.R* x
Source on Save Run Source
51 count(mV[[1]],1) # Virus del MERS
52 count(dV[[1]],1) # Virus del Dengue
53
54
55 # ----- ¿Cuál es el contenido de GC de cada virus? -----
56
57 print(GC(zV[[1]])) # Virus del Zika
58 print(GC(wV[[1]])) # Virus del Wuhan-Hu-1
59 print(GC(cV[[1]])) # Virus del SARS
60 print(GC(mV[[1]])) # Virus del MERS
61 print(GC(dV[[1]])) # Virus del Dengue
62
63 # Función para obtener la secuencia en complementaria e imprimirla por cada secuencia
66:37 # ¿Cuál es el contenido de GC de cada virus? R Script
Console Terminal Jobs
R 4.1.2 ~ /Downloads/Tec/R/ADN/Teams/
> count(dV[[1]],1) # Virus del Dengue

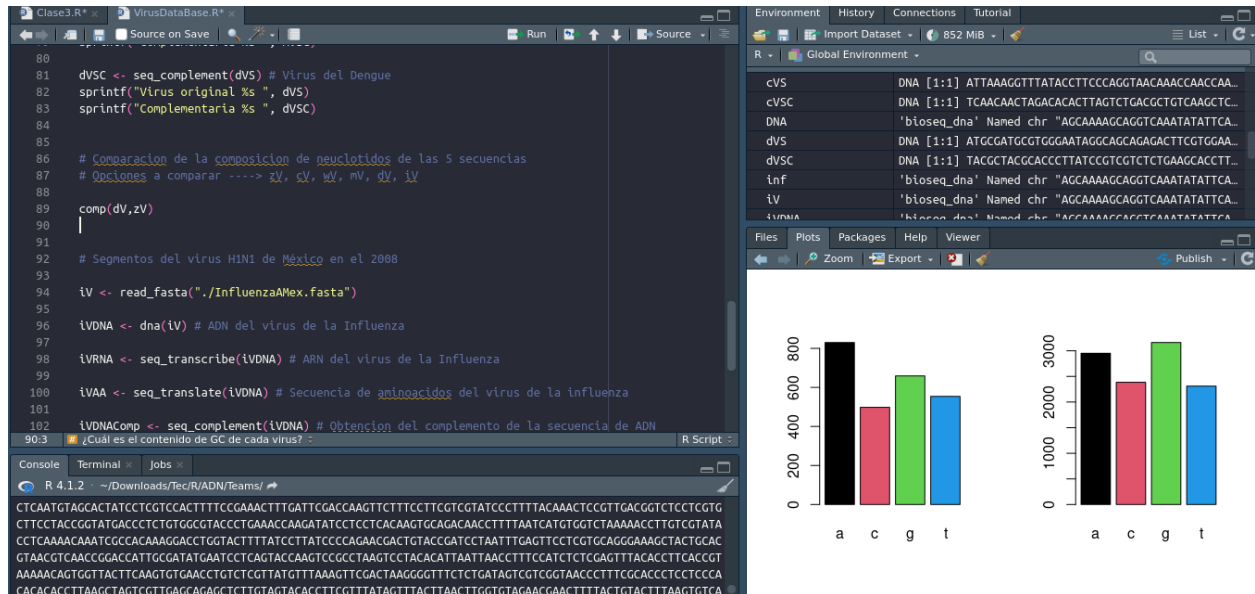
a c g t
830 498 659 554
> print(GC(zV[[1]])) # Virus del Zika
[1] 0.5131859
> print(GC(wV[[1]])) # Virus del Wuhan-Hu-1
[1] 0.3797278
> print(GC(cV[[1]])) # Virus del SARS
[1] 0.3797278
> print(GC(mV[[1]])) # Virus del MERS
[1] 0.4123643
> print(GC(dV[[1]])) # Virus del Dengue
[1] 0.4553325
>
```

Obtencion de la secuencia complementaria

```
Clase3.R* x VirusDataBase.R* x
Source on Save Run
67 sprintf("Complementaria %s ", zVSC)
68
69 wVSC <- seq_complement(wVS) # Virus del Wuhan-Hu-1
70 sprintf("Virus original %s ", wVS)
71 sprintf("Complementaria %s ", wVSC)
72
73 cVSC <- seq_complement(zVS) # Virus del SARS
74 sprintf("Virus original %s ", cVS)
75 sprintf("Complementaria %s ", cVSC)
76
77 mVSC <- seq_complement(zVS) # Virus del
78 sprintf("Virus original %s ", mVS)
79 sprintf("Complementaria %s ", mVSC)
74:37 # ¿Cuál es el contenido de GC de cada virus? R Script

Console Terminal Jobs
R 4.1.2 ~./Downloads/Tec/R/ADN/Teams/
CTTTCCCTCTTACCTACTGGTGACTTCTGTACGAACACCACACCTTGTCTCACACCTAACTCCTCTTGCTGGTGTACCTTCTGTTCTGGGGTCAATGCTTTACC
TGTCTGTAAGGGATAAACCTTTTTCCCTTCTGAACACCACACCCAGAGATATCCCGTGTCTGGCGCGTGGTGGACCCGACTCTTGTAAATTTTGTGTCAGTT
GTACTACGCGTCTAGTATCCACTACTTCTTTTCATGTACCTGATGGATAGGTGGGTTCAAGCGATGAACCCACTTCTCCAGGTGTGGACCTCACGATATTC
GTGGTCAGAATCACAACAGTCCGGACGATCAGTCGGTGTGCAACCCCTTTCGACACGTCGGACACTGGGGGGTCTCTTCGACCCTTTGGTTCGGGTATCAGT
CCGGCTCTTGGGTACCGTGCCTTCTCGGTACGACGGACACTCGGGGAGTCTCCTGTGACTCAGTTTTTTGGGGTGGCGGAACCTCCGCGTCTACCCCTTTTC
TTCCACCGCTGGAAGGGTGGGAAGTTAGACCCCGGACTTGACCTCTAGTCGACACCTAGAGGTCTTCTCCCTGATCACCATCTCCTCTGGGGGGCTTTTGC
GTTTTGTCGTATAACTGCGACCTTTCTGGTCTCTGAGGTACTCAAAGGTGGTGGCACC GGCGGTCCGTGTCTAGCGGCTTATCGCCGCGGCCACACCCCTTT
AGGTACCCAGA "
>
> dVSC <- seq_complement(dVS) # Virus del Dengue
> sprintf("Virus original %s ", dVS)
[1] "Virus original ATGCGATGCGTGGGAATAGGCAGCAGAGACTTCGTGGAAGGACTGTGAGGAGCAACTTGGGTAGATGGTGTGGAACATGGA
AGTTGCGTCACTACCATGGCAAAAGATAAACCAACATTGGACATTGAACCTTGAAGACGGAAGTCACAGACCTGCCGTCTGCGCAAACTGTGTATTGAAGC
TAAATATCAAAACACCACCACTGACTCAAGATGTCCAACACAAGGAGAAGCCACATTGGTGGAGAACAAGACGCGAACTTTGTGTGTCGACGAACGTTTGTGG
ACAGAGGCTGGGGCAATGGCTGTGGGCTCTTCGAAAAGGAAGCCTCATAACGTGTGCTAAGTTCAAGTGTGTGACAAAACCTGGAAGGAAAAATAGTTCAATAT
GAAAACCTTGAATATTCAGTAATAGTACCGTCCACACTGGAGACCAGCACCAGGTGGGAAATGAAAGCACAGAACATGGGACAACCTGCAACCATAACACCTCA
AGCTCCTACGACAGAAATACAGCTGACCGACTACGGAGCTTTACATTGGATTGTTACCTAGAACAGGACTGGACTTCAATGAAATGGTGTGTTGACAATGA
```

Comparación de nucleótidos de las 5 secuencias



Segmentos del virus H1N1 en México

```
Clase3.R* x VirusDataBase.R* x
Source on Save Run Source
94 iv <- read.fasta(./influenzaA/Mex/1999/1999_H1N1.fasta)
95
96 iVDNA <- dna(iv) # ADN del virus de la Influenza
97 iVDNA
98
99 iVRNA <- seq_transcribe(iVDNA) # ARN del virus de la Influenza
100 iVRNA
101
102 iVAA <- seq_translate(iVDNA) # Secuencia de aminoácidos del virus de la influenza
103 iVAA
104 iVDNAComp <- seq_complement(iVDNA) # Obtencion del complemento de la secuencia de ADN
105 iVDNAComp
106 iVDNACompRev <- seq_reverse(iVDNAComp) # Secuencia complemento inversa
107 iVDNACompRev
108
104:38 # ¿Cuál es el contenido de GC de cada virus? R Script
Console Terminal x Jobs x
R 4.1.2 ~ /Downloads/Tec/R/ADN/Teams/
DNA vector of 1 sequences
KY575193.1 Influe... AGCAAAAGCAGGTCAAATATATTCAATATGGAGAGAATAAAGAAGAGATCTAATG... + 2281 bases
> iVRNA
RNA vector of 1 sequences
> AGCAAAAGCAGGUCAAAUUAUUCAAUAUGGAGAGAAUAAAAGAACUAAGAGAUUUAUG... + 2281 bases
> iVAA
AA vector of 1 sequences
KY575193.1 Influe... SKSRNIFNMERIKELRDLMSQSRTREILTCTTVDHMAIIKKYTSGRQEKNPALRMKMM... + 720 amino acids
> iVDNAComp
DNA vector of 1 sequences
KY575193.1 Influe... TCGTTTTTCGTCCAGTTTATATAAGTTATACCTCTCTTATTTCTTGATTCTCTAGATTAC... + 2281 bases
> iVDNACompRev
DNA vector of 1 sequences
KY575193.1 Influe... AGTAGAAACAAAGGTCGTTTTTAAACAATTCGACACTAATTGATGGCCATTGGAATTCCTT... + 2281 bases
>
```