

Proyecto Genómica Computacional

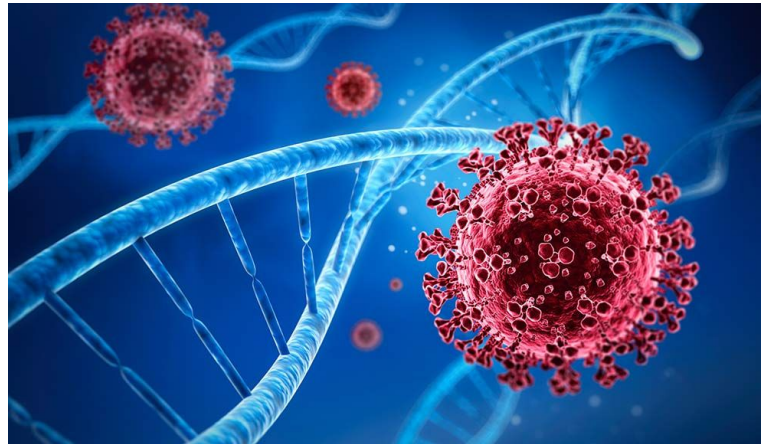
*Comparación entre genomas de
variantes de Coronavirus
(SARS-CoV-2)*

David Hernández Uriostegui
Noé Abraham Suaste Morales



Introducción

En medio de la pandemia de COVID-19, el SARS-CoV-2 ha evolucionado y dado lugar a diferentes variantes en todo el mundo. Estas variantes difieren en cierta medida de la secuencia original del virus, lo que plantea interrogantes sobre el alcance y las implicaciones de estos cambios genéticos





Pregunta de Investigación

¿Qué tanto difieren las secuencias
de las variantes de SARS-CoV-2 con
respecto
a la secuencia original?



Objetivos

El objetivo de éste trabajo es determinar las diferencias y similitudes entre las principales variantes de importancia médica del SARS-CoV-2 frente a la secuencia original

- **Comparar los genomas de diferentes cepas:** Comparar los genomas de diferentes cepas para identificar similitudes y diferencias en las mutaciones que se han producido.



Métodos

Se consideraron las secuencias de nucleótidos de las principales variantes de importancia del SARS-CoV-2: alfa (B.1.1.7), beta (B.1.3.51), delta (B.1.617.2), gamma (P.1) y ómicron (B.1.1.529) así como la variante original. [1]

- **Alineación de los genomas:** Comparación de las secuencias de los nucleótidos con respecto a otro. (BLAST y BioPython)
- **Alineación y filogenética:** Análisis filogenético de las secuencias para comprender la relación evolutiva entre ambos virus.

Obtención de datos

- NCBI GenBank

FASTA ▾

Send to: ▾

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

NCBI Reference Sequence: NC_045512.2

[GenBank](#) [Graphics](#)

>NC_045512.2 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

```
ATTAAAGGTTTATACCTTCCAGGTAACAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAA
CGAACTTTAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCAGCAGTATAATTAATAAC
TAATTACTGTCTGTGACAGGACACAGTAACCTGCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTG
TTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAAGTGGAGAGCCTTGTC
CCTGGTTTCAACGAGAAAAACACAGTCCAACCTCAGTTTGCCTGTTTACAGGTTTCGCAGCTGCTCGTAC
GTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCAGTCAACATCTTAAAGATGGCACTTGTGG
CTTAGTAGAAGTTGAAAAGGCGTTTTCCTCAACTTGAACAGCCCTATGTGTTTCAATCAACGTTCCGGAT
GCTCGAAGTGCACCTCATGGTCAATGTTATGGTTGAGTGGTAGCAGAACTCGAAGGCATTCAGTACGGTC
GTAGTGGTGAACACTTGGTGTCTTGTCCCTCATGTGGCGAAATACAGTGGCTTACCGCAAGGTTCT
TCTTCGTAAGAACGTTAATAAAGGAGTGTGCTGCCATAGTTACGCCGCCGATCTAAAGTCATTGACTTA
GGCGACGAGCTTGGCACTGATCCTTATGAAGATTTTCAAGAAAACGGAACACTAAACATAGCAGTGGTG
TTACCCGTGAAGCTATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTGATAAACAACCTTCTGTGG
CCCTGATGGCTACCTTCTTGAGTGCATTAAAGACCTTTCAGCACGTGCTGGTAAAGCTTCATGCACTTTC
TCCGAACAACTGGACTTTTATGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTG
CTTGGTACACGGAACGTTCTGAAAAGAGCTATGAATTGCAGACACCTTTGAAAATTAAATTGGCAAGAA
ATTTGACACCTTCAATGGGAATGTCCAAATTTGTATTTCCCTTAAATTCATTAATCAAGACTATTCAA
```





- ## 7ZBCZVZZ114-Alignment-Descriptions

[illegible]

- **Uso del algoritmo Needleman-Wunsch (también implementado)**
- **Evaluar por regiones, de otra forma se acaban los recursos de nuestras computadoras**

```
In [5]: score = perform_sequence_alignment(ref_seq, alpha_seq, gap_penalty, match_score, mismatch_penalty, start=200, end=600,
```

TTCTGTCCTGGTTGTCAGCCGATCATCAGCACATCTAGGTTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCCCTGGTTTCAACGAGAAAAACACAGTCCAACCTCAGTT
TGCCTGTTTTACAGGTTCCGCGACGTGCTCGTACGTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACCTTGTGGCTTAGTAGAAGTTG
AAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTCCGGATGCTCGAACTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTC
AGTACGGTCGTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGG

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

TTTCGTCCGTGTTGCAGCCGATCATCAGCACATCTAGGTTTTGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCCTCGGTTTCAACGAGAAAAACACAGTCCAACCTCAGTT
TGCCTGTTTTACAGGTTTCGCGACGTGCTCGTACGTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACCTGTGGCTTAGTAGAAGTTG
AAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTCCGGATGCTCGAACTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTC
AGTACGGTCGTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGG

Score=797

Resultados obtenidos de alinear la secuencia original contra sus variantes

1. Se evaluó en intervalos de 900 bases. [(0, 900), (1000, 1900)...
2. Se promedió el resultado de cada variante

```
In [9]: for seq, name in variants:
        for start, end in query_ranges:
            score = perform_sequence_alignment(ref_seq, seq, gap_penalty, match_score, mismatch_penalty, start=start, end=end)
            variants_scores[name] += score

In [10]: for variant in variants_scores:
          variants_scores[variant] /= len(query_ranges)
```

☐ **Alpha: 1760**

☐ **Beta: 1691**

☐ **Delta: 1751**

☐ **Gamma: 1671**

☐ **Omicron: 1668**

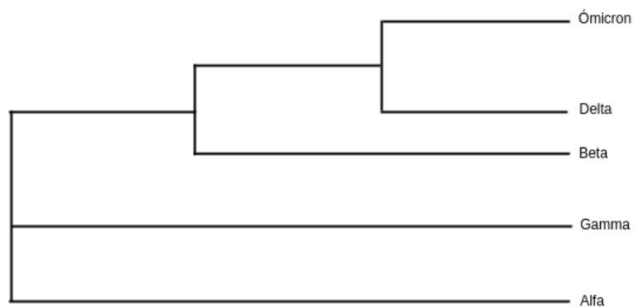


Conteo de GC

	Secuencia original	Alpha	Beta	Delta	Gamma	Omicron
G	5863	5844	5701	5778	5647	5615
C	5492	5462	5324	5393	5282	5256



CLUSTEL OMEGA



	original	alfa	beta	delta	gamma	ómicron
original	100	99.47	97.05	98.24	95.94	95.89
alfa		100	97.36	98.54	96.29	96.21
beta			100	96.10	93.87	93.78
delta				100	96.50	95.09
gamma					100	92.70
ómicron						100



Referencias

[1] Farhud, D. D., & Mojahed, N. (2022). SARS-COV-2 Notable Mutations and Variants: A Review Article. *Iranian Journal of Public Health*. <https://doi.org/10.18502/ijph.v51i7.10083>

[2] Mittal, A. (2021, February 23). *Sequence Alignment and the Needleman-Wunsch Algorithm*. Analytics Vidhya.
<https://medium.com/analytics-vidhya/sequence-alignment-and-the-needleman-wunsch-algorithm-710c7b1a23a4>

[3] *Bio.pairwise2 module — Biopython 1.75 documentation*. (n.d.). Biopython.org.
<https://biopython.org/docs/1.75/api/Bio.pairwise2.html>