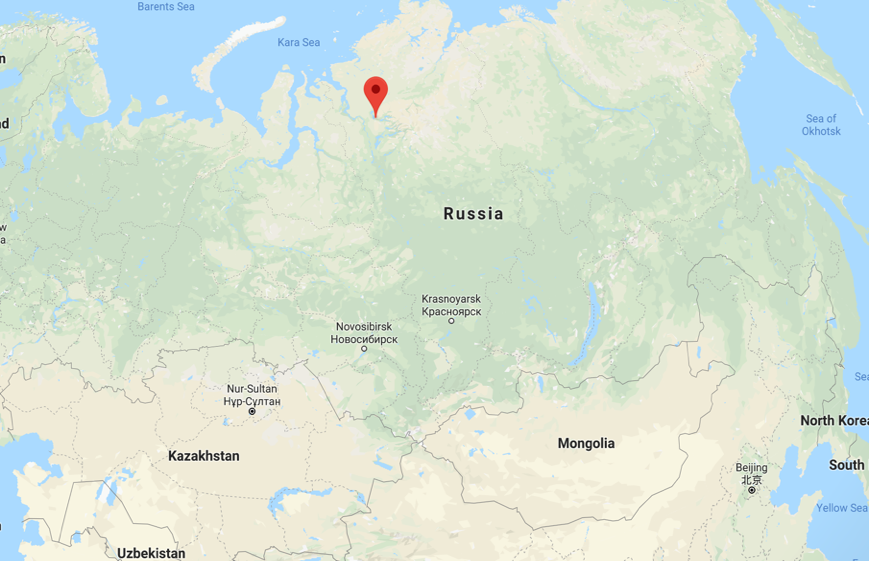
**Proyecto Final Curso de Bioinformática y Biología Computacional**

**Código DRBC A001**

**Semestre I, 2025**

**Introducción**

El verano de 2020 en el hemisferio norte ha sido catalogado como uno de los más calurosos de los que se tenga registro en la historia reciente. Más allá de las consideración éticas y morales, no cabe duda de que los diversos fenómenos relacionados con el cambio climático global han ido en aumento durante la última década. Es así como, año a año, las temperaturas promedio en el ártico y en particular en la zona de Siberia, han aumentado un promedio de 20℃ durante los veranos, provocando el derretimiento de las zonas cubiertas por el denominado *permafrost*. El derretimiento de estas regiones provoca la liberación del CO2 generado por la descomposición de la materia orgánica que se encuentra congelada, generado una retroalimentación positiva en el ciclo atmosférico de este gas. Además, del grave efecto generado por este gas sobre la regulación del clima en la tierra, se ha discutido la posibilidad que el derretimiento del permafrost libere agentes patógenos que han estado congelados por miles de años, incluso desde antes del último período glacial.



En este contexto, en el mes de Agosto recién pasado, una serie de localidades cercanas a la Ciudad de Norilsk, Siberia, Rusia, (Fig. 1) fueron afectadas por el brote de una enfermedad altamente contagiosa con una letalidad (CFR) de entre el 30 al 60%, cuyo agente etiológico es desconocido.

Figura 1. Posición de la ciudad de Norilsk, Siberia, Rusia.

Los síntomas de la llamada “fiebre de Norilsk” están agrupados en tres etapas. La primera etapa de la enfermedad se manifiesta 5 a 7 días desde el contagio. Durante esta etapa ocurren ciclos febriles cada 24 a 36 horas con temperaturas cercanas a los 41°, los que vienen asociados con vómitos, diarreas, fuertes dolores musculares e intensas cefaleas. En la segunda etapa, habitualmente entre 7 a 10 días de manifestados los primeros síntomas, los pacientes presentan, además, espasmos musculares, junto con dificultad para respirar. En la tercera etapa, que comienza entre 10 a 12 días de los primeros síntomas, los pacientes en su mayoría requieren de ventilación mecánica, y manifiestan un sangrado difuso proveniente de las diversas mucosas, incluyendo el tracto digestivo, los ojos, boca, nariz y ano. En esta etapa, además, los pacientes muestran una inflamación aguda de labios y ojos. La muerte sobreviene por falla multisistémica de la mayoría de los infectados sintomáticos (alrededor del 70%), no más allá del día 15. Dado que la enfermedad es aún reciente, no ha sido posible establecer su número reproductivo básico (R0), sin embargo, debido a la alta tasa de contagio, se asume que esta enfermedad se transmite de manera indirecta. Las autoridades de la región han impuesto la cuarentena, impidiendo el libre tránsito de personas, tanto hacia, como desde la región. Sin embargo, dado que el tren trans-siberiano pasa por las cercanías de esta ciudad, y debido a un período de incubación cercano a los 5 días, es altamente probable que personas infectadas hayan dejado la ciudad.

Pese a los esfuerzos de las autoridades locales, tras 3 semanas de desarrollo, la epidemia parece sin control, debido a la alta población flotante de la ciudad, siendo identificadas a la fecha más de 15.000 personas afectadas por la fiebre de Norilsk en un radio de 500 kms alrededor de esta ciudad, de acuerdo al siguiente cuadro:

Tabla : Evolución de la fiebre de Norlisk durante las primeras 3 semanas

La semana pasada, se presentaron en forma simultánea casos confirmados en Londres, Paris, New York, Los Ángeles, y Tokio situación que llevó recientemente a la OMS a decretar el estado de alerta sanitaria por pandemia.

Un equipo conjunto del CDC (Center for Disease Control) de USA y del Center for Infectious Diseases de la Organización Mundial de la Salud (CID-OMS), ha conducido un experimento de RNA-seq mediante la obtención de muestras de sangre y fluidos de mucosas de los pacientes infectados, durante la primera semana de desatada la enfermedad en Rusia. Este experimento tiene el objetivo de determinar que genes se ven diferencialmente expresados frente a la infección. La lista de estos genes aparece en el archivo “RNA-Seq-expression-Norilsk2019.csv” (https://tinyurl.com/RNA-seq-Norilsk2019) que se encuentra en repositorio del curso. En experimentos posteriores, se determinó que la mayoría de los pacientes afectados por esta enfermedad posee los siguientes transcritos sobre expresados, los cuales han sido secuenciados y son entregados a continuación en formato FASTA:

>unknown sequence 1

atttttaggagaggcctagagcctgagctttgtgaggacgtcctgtggaacctgtcccggccccctgccctgggatggggagaagtcagggggatagacagagtcaaggtgggggacagggcgggagtggggtccccagggctgggggcctttggtgcagtgaccagagtgtcaggagaggggagcaaagccctctagcctcatcctcataaaaggtctcatcatagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgaggagtatcagccctagataatataagtcaattaaacgaccaagccaaaattgtccctgccctgggatggggagaagtcagggggatagacagagtcaaggtgggggacagggcgggagtgggg

>unknown sequence 2

aaattatcatacaaatttgggggggggaggcctttgaagcaggtgttgaxxxxxttttcaagagagtgcggacagtttccttcgaggcctttgaagcaggtgagagcctgagctttgtgaggacgtcctgtggaacctgtcccggccccctgccctgggatggggagaagtcagggggatagacagagtccagtccaatcctaaggtgggggacagggcgggagtggggtccccagggctgggggcctttggtgcagtgaccagagtgtcaggagaggggagcaaagccctctagcctcatcctcataaaaggtctcatcattttccctccagcctcttatgcactggggatcatcatgctttgttgttgaxxxxxttttcaagagagtgcggacagtttccttcgaggcctttgaagcaggtgagagcctgagctttgtgaggacgtcctgtggaacctgtcccggcc

>unknown sequence 3

catgctttgtcttcgaaatacaaaattatgattctgcaaagtttaatgaatttcagagacatgttgaagaatatgaagtttctctcattttacaattgtgtaaaattccattgaaagctgaggttcttgcacagattaatgcaatgaactccgatattttagaaaatcatgcgtaccaaggcatcat

>unknown sequence 4

taaggaaactgcactcagcgcctccctggccaaaaxxxxaagcatggagtttttaaaagtatgctcctttcatagatagatagatagatagtagatagatgatagtgcccgaagagagagagagagggaatgctgaaattgaagatctcatctttctggcacggtccgcactcattctgagaggatcaatggcccataagtcctgcctgcctgcttgtgtgtacggacttgccgtagccagtggatatgacttcgagagagaagggtactctttcagcagtccaatccttccaatcctagtcggaatagatcctttccgtctgctgcaaaacagccaggtcttcagtctaattagaccaaatgagaatccagcacataaaagtcaattggtgtggatggcatgccattctgcagcgtttgaagatctgagagtatcaagtttcatcagaggaacaagagtggtcccaagaggacagctatctaccagaggagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgagaagcagatattgggctataaggaccaggagtggag

>unknown sequence 5

gctcctttcgcccgaagagagagagaagagcctgagctttgtgaggacgtcctgtggaacctgtcccggccccctgccctgggatggggagaagtcagggggatagacagagtcaaggtgggggacagggcgggagtggggtccccagggctgggggcctttggtagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgagaagcagatattgggctataaggaccaggagtggaggaaacaccaaccagcagagagcatctgcaggacagatcagtgtgcagcctactttctcagtacagagaaatcttccctttgaaagagcgaccattatggcggcattcacagggaatacggagggcagaacacagtccaatccttctgacatgaggactgaaatcataagaatgatggaaagtgccaaaccagaagacgtgtctttccaggggcggggcagtgaccagagtgtcaggagaggggagcaaagccctctagcctc

atcctcataaaaggtctcatcattttccctccagcctcttatgcactggggacttttgaacxxgaggccttatagatagatagatagatagtagatagatgatagttgaagcaggtgtttctggagxxxaaataatctgagagagagcatggtcttttccctca

>unknown sequence 6

agagagagagattgataaggctgagggagatgacggagccacaacaacgagtttgtgagctattttccattcaaaaacactgggggcatagatagatagataagagcctgagctttgtgaggacgtcctgtggaacctgtcccggccccctgccctgggatggggagaagtcagggggatagacagagtcacagtccaatcctaggtgggggacagggcgggagtggggtccccagggctgggggcctttggtgcagtgaccagagtgtcaggagaggggagcaaagccctctagcctcatcctcataaaaggtctcatcatagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgaggagtatcagccctagataatataagtcaattaaacgaccaagccaaaattgttcatatcccgctagcagcttaaaatataaatgaaataggagctatatctctgacagtattataatcaattgttattaagtaacccaaaccaaaaatgagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatatgaagattaagaaaaacctacctcgactgagagagtgtttttccattaaccttcatcttgtaaacgttgagcaaaattgttaagaatatgaggcgggttatattgcctactcagtccaatcctgctcctcctgaatatatggaggccatataccctgtcaggtcaaattcaacaattgctaggggtggcaacaacaatacaggcttcctgacaccggagtcagtcaatggagacactccatcgaatccactcaggccaattgctgatgacaccatcgaccatgctagccacacaccaggcagtgtgtcatcagcattcatccttgaagctatggtgaatgtcatatcgggccccaaagtgcaagcagatattgggctataaggaccaggagtggaggaaacaccaaccagcagagagcatctgcagacagatcagtgtgcagcctactttctcagtacagagaaatcttccctttgaaagagcgaccattagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatatggcggcattcacagggaatacggagggcagaacatctgacatgaggactgaaatcataagaatgatggaaagtgccaaaccagaagacgtgtctttccaggggcgggtttccctccagcctcttatgcactcagtccaatcctggggagatagtagatagatgatagtatcctatatattccctgtccctggccaaaaxxxxaagcatgtccctggccaaaaxxxxaagcatggccaaaaxxxxaagctccctggccaaaaxxxxaagcatgatggagtttttaaaagtaatgtgctacatagatagataagtgaaacagxxxxxxxcaatccaagaggcctttgaagcaggtgaattttttttttagttaaaaaacatgttccacgttttgctaaaagc

>unknown sequence 7

gctgagggagatgacggagccacaacaacgagtttgtgagctattttccattcaaaaacactgggggcatcctatatcatgactgtxxxxaaatcagaattttccctggccaaaaxxxxaagcatgtccctggccaaaaxxxxaagcatgtttacctacatcctcacagcagaaagcggattggataccattagcattgacgatcaagggtctcaatccaacagctatttttctaacaaccctttcaagaaccaacaagaaaaggagctggccactaaatgaggctatcatggcagtcgggatggtgagcattttggccagttcactcctaaagaatga

cattcccatgacaggaccattagtggctgtcactggacgatcggccgatttggaactggagagagccgccgatgtcaaatgggaagatcaggcagagatatcaggaagcagtccaatcctgtcaataagtccctggccaaaaxxxxaagcatgcgagatagatagatagatagatagtagatagatgattaaatgaaataggagctatatctctgacagtattataatcaattgttattaagtaacccaaaccaaaaatgatgaagattaaagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatgaaaaacctacctcgactgagagagtgtttttccattaaccttcatcttgtaaacgttgagcaaaattgttaagaatatgaggcgggttatattgcctactgctcctcctgaatatatggaggccatataccctgtagtaacatcctcacagcagaaagcggattggataccattagcattgacgatcaagggtctcaatccaacagctatttttctaacaaccctttcaagaaccaacaagaaaaggagctggccactaaatgaggctatggaggctcgccaggataaacgttgcaaaaaggatxxxt

>unknown sequence 8

Gggggaggcctttgaagcagtccctgagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgagaagcagatattgggctataaggaccaggagtggaggaaacaccaaccagcagagagcatctgcaggacagatcagtgtgcagcctactttctcagtacagagaaatcttccctttgaaagagcxxxxxxxxxxxxxxxxxxxxxxxxxxxxggcattcacagggaatacggagggcagaacatctgacatgaggactgaaatcataagaatgatggaaagtgccaaaccagaagacgtgtctttccaggggcggggccaaaaxxxxaagcatgagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatattccctggccaaaaxxxxaagcatgtccctggccatagatagatagatagatagtagatagatagagcctgagctttgtgaggacgtcctgtggaacctgtcccggccccctgccctgggatggggagaagtcagggggatagacagagtcaaggtgggggacagggcgggagtggggtccccagggctgggggcctttggtgcagtgaccagagtgtcaggagaggggagcaaagccctctagcctcatcctcataaaaggtctcatcattttccctccagcctcttatgcactggggagatagtaaaaxxxxaagcatggtgttgaxxxxxttttcaagagagtgcggacagtttccttcgaggcctttgaagcaggtgtccagaattttttaccttcttacaatatagc

>unknown sequence 9

Atagatagatagatagaxxxxxxxxxxxxxxxxtagtagatagatgatagtgagtatcagccctagataatataagtcaattaaacgaccaagccaaaattgttcatatcccgctagcagcttaaaatataaatgaaataggagctatatctctgacagtattataatcaattgttattaagtaacccaaaccaaaaatgatgaagattaagaaaaacctacctcgactgagagagtgtttttccattaaccttcatggctgctgtcctcctacggggagctcctggtgagagtctctccttgctgcagcccccagcagaggggcagggctgggggacggtgcagggaggggacaggctcccagtgggaggaaactgaggcctggacctccaggactcaggctctgtttgggagaaggcttgtctctgcccagtcctcaccccattatcccaggcctccgaaggcccggcgggggagatgggggtgactctacccaaggaacttgtaaacgttgagcaaaattgttaagaaxxxxxxxxxxxxxxxxtatgaggcgggttatattgcctactgctcctcctgaatatatggaggccatataccctgtcaggtcaaattcaacaattgctaggggtggcaacaacaatacaggcttcctgacaccggagtcagtcaatggagacactccatcgaatccactcaggccaattgctgatgacaccatcgaccatgctggctgctgtcctcctacggggagctcctggtgagagtctctccttgctgcagcccccagcaagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatgaggggcagggctgggggacggtgcagggaggggacaggctcccagtgggaggaaactgaggcctggacctccaggactcaggctctgtttgggagaaggcttgtctctgcccagtcctcaccccattatcccaggcctccgaaggcccggcgggggagatgggggtgactctacccaaggaaagccacacaccaggcagtgtgtcatcagcattcatccttgaagctatggxxxxxxxxxxxxxxxxtgaatgtcatatcgggccccaaagtgc

>unknown sequence 10

ccaacgagtttgtgagctattttccattcaaaaacactgggggcatcctatatcatgactgtxxxxaaatcagaattttccctggccaaaaxxxxaagcatggcagtggagtcagtgatgccattggcctcctgccagcactgcctgtctgaggagcctgagattcgggcctagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgagaagcagatattgggctataaggaccaggagtggaggaaacaccaaccagcagagagcatctgcaggacagatcagtgtgcagcctactttctcagtacagagaaatcttccctttgaaagagcgaccattatggcggcattcacagggaatacggagggcagaacatctgacatgaggactgaaatcataagaatgatggaaagtgccaaaccagaagacgtgtctttccaggggcgggtcgaccctgaggctgcggccgtgcagccctaccaagaccagacgtaccagtcagtctacttcgtgtctgagagcttcagtgacgccaaggacaagctcaggtgggctaggctgctagggcaagccccccatggtgcccccaaactgggccagccaggccttccttctggccttgagcagggctggacctgtgagcccaggtcacagatgagaagaaaagatagcaactcattggatatcattcatgctgagttccaggccagcctggctgaaggagactctcctcaatgtgccgcgatgaatgagcactgcctaattcaaattacaaaaagagttccaatcttccaagatgctgctccacctgtcatccacatccgctctcgaggtgacattccccgagcttgccagaaaagcttgcgtccagtcccgccatcacccaagattgatcgaggttgggtatgtgttttccagcttcaagatggtaaaacacttggactcaaaatttgagccaatctcccttccctccgaaagaggcgaccaataxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxgcagaggcttcaactgctgaactacagggtacgttacattaatgatacacttgtgacccggcttccagcgcgaacatgagcgtcggtttcgacgtgttgggcgcagccgtcacacccgttgacggcacgttgctgggcgatgtggtatccgttgaagcagcggaccatttccgtctgcataacctggggcgatttgccgataaactgccgccggagccgcgtgaaaatattgtttatcagtgctgggaacgtttttgccacgcgatgaatgagcactgggcattggggaaaaccatcccggtggcgatgacgctggaaaaaaatatgccgattggttccgggttagggtccagcgcctgttccgtcgtcgccgcgctggtcgcgatgaatgagcactgcggcaaaccgttaaacgacacgcgtctgttggcgctgatgggcgagctggaaggccgtatctccggcagcatccattacgataacgtcgcgccgtgctttcttggcggtatgcagttgatgattgaagaaaacggcattattagtcagcaggtgccgggctttgatgagtggctatgggtactggcttatccgggcattaaagtttctaccgcagaagcacgggccattttgcctgcgcagtatcgccgtcaggtatcagccctagataatataagtcaattaaacgaccaagccaaaattgttcatatcccgctagcagcttaaaatataaatgaaataggagctatatctctgacagtattataatcaattgttattaagtaacccaaaccaaaaaaaccgacccctggttgcagagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatcagcccccacacagcagggacaccatccgtgagaaggaccccagcgtctggggaggggcagacctacaggactgggggctgctgggtggccgggtcaaggccagtcttggaggtgctgacagagcctgagctttgtgaggacgtcctgtggaacctgtcccggccccctgccctgggatggggaggaatgctgaaattgaagatctcatctttctggcacggtccgcactcattctgagaggatcaatggcccataagtcctgcctgcctgcttgtgtgtacggacttgccgtagccagtggatatgacttcgagagagaagggtactctttagtcggaatagatcctttccgtctgctgcaaaacagccaggtcttcagtctaattagaccaaatgagaatccagcacataaaagtcaattggtgtggatggcatgccattctgcagcgtttgaagatctgagagtatcaagtttcatcagaggaacaagagtggtcccaagaggacagctatctaccagaggagttcagattgcttcaaatgagaacatggaaacaatggactccagcactcttgaactgagaagcagatattgggctataaggaccaggagtggagaagtcagggggatagacagagtcaaggtgggggacagggcgggagtgcttccacccacgaattgtgaccctcaccctggcctgcctgcatcctggcctggcctccctgggggtggtatcctggtcacgggtgaccaggggctgcccggtgggcggcagctgtctctgggctgatgctgcccggcttccccgcagctgtactggttcacggtggagttcgggctgtgtaagcagaacggggaggtgaaggcctatggtgccgggctgctgtcctcctacggggagctcctggtgagagtctctccttgctgcagcccccagcagaggggxxxxxxxxxxxxxxxxxxxxxxxxcagggctgggggacgcccggcttccagcgcgaacatgagcgtcggtttcgacgtgttgggcgcagccgtcacacccgttgacggcacgttgctgggcgatgtggtatccgttgaagcagcggaccatttccgtctgcataacctggggcgatttgccgataaactgccgccggagccgcgtgaaaatattgtttatcagtgctgggaacgtttttgccaggcattggggaaaaccatcccggtggcgatgacgctggaaaaaaatatgccgattggttccgggttagggtccagcgcctgttccgtcgtcgccgcgctggtcgcgatgaatgagcactgcggcaaaccgttaaacgacacgcgtctgttggcgctgatgggcgagctggaaggccgtatctccggcagcatccattacgataacgtcgcgccgtgctttcttggcggtatgcagttgatgattgaagaaaacggcattattagtcagcaggtgccgggctttgatgagtggctatgggtactggcttatccgggcattaaagtttctaccgcagaagcacgggccattttgcctgcgcagtatcgccgtcaggtgcagggaggggacaggctcccagtgggaggaaactgaggcctggacctccaggactcaggctctgtttgggagaaggcttgtctctgcccagtcctcaccccattatcccaggcctccgaaggcccggcgggggagatgggggtgactctacccaaggaacccacccagcgtcaggccacggtgccccagttccctcggggacctgggtagggagatgacggcgcgatgaatgagcactgagccacaagggtccccagggctgggggcctttggtgcagtgaccaxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxgagtgtcaggagaggggagcaaagccctctagcctcatcctcataaaaggtctcatcattttccctccagcctcttatgcactggggaaactgaggccaggggctatgtgtccagcggacaggggtgctgaattccacccacaggcttagggatatggtcaaggaaagcttcctggaggaggcccagtggaggttcagggagggatggggtgcccggcagtctctagtggaaaaggcgcctagcctatctcccccatgaaccccctcacccagccctggaagaggcctcagtgtcccgcctgtgaccagttggctcagaaaagccctgggagctctgagccactgtgaaggtggaaacgcggcccctggcctcccctctcctggaggctgcagactctgcccgccagttgacgagggctctgccgctctcct

>unknown sequence 11

cgacggtatgcgcacgctacgcggcatttcagcgaaattcttcgccgcgctggcgcgggctaatatcaatatcgtggcgatcgctcagggatcttctgagcgttctatttctgtggtggtgaataacgacgatgccaccaccggcgtgcgggttagatcctttccgtctgctgcaaaacagccaggtcttcagtctaattagaccaaatgagaatccagcacataaaagtcaattggtgtggatggcatgccattctgcagcgtttgaagatctgagagtatcxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxaagtttcatcagaggaacaagagtggtcccaagaggacagctatctaccagaggagttcagattgcttcaaatgagaacatggaaacaataacgcaccagatgctgttcaataccgatcaggtgattgaagtgtttgtcattgggctgctgtcctcctacggggagctcctggtgagagtctctccttgctgcagcccccagcagaggggcagggctgggggacggtgcagggaggggacaggctcccagtgggaggaaactgaggcctggacctccaggactcaggctctgtttgggagaaggcttgtctctgcccagtcctcaccccattatcccaggcctccgaaggcccggcgggggagatgggggtgactctacccaaggaagcgtcggcggcgtcggcggcgcgctactggaacagcttaaacgtcagcaaacctggctgaagaacaagcacatcgatctacgcgtgtgcggcgtggcgaactcaaaggcgttgctaaccaatgtgcatggcctgaatctggacaactggcaggcggaactggxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxcgcaagcgaacgcgccgttcaatctgggacgcttaattcgcctggtgaaagaatatcatctactcaatccggtgattgttgattgtacctccagtcaggcggtggccgaccagtatgccgacttcctgcgtggaagtcagggggatagacagagagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatattcaaggtgggggacagggcgggagtggggtccccagggctgggggcctttggtgcagtgaccagagaagggttccatgtggtgacgccaaacaagaaagcgaacacctcgtcaatggactactaccatcagctacgtttcgccgccgcgcaatcacggcgcaaattcttgtatgacaccaacgtcggcgccggtttgccggta

>unknown sequence 12

Ggctagtgtaattacaaagggaaccaaggaaaggcaatacttagatgaagagtttgttcttcgagtgatgactcagttgactctggccctgaaggaatgccacagacgaagtgatggtggtcataccgtattgcatcgggatctgaaaccagccaatgttttcctggatggcaagcaaaacgtcaagcttggagactttgggctagctagaatattaaaccacgacacgagttttgcaaaaacatttgttggcacaccttattacatgtctagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatcctxxxxtacaatgagaaatcagatatctggtcattgggctgcttgctgtatgagttatgtgcattaatgcctccatttacagcttttagccagaaagaactcgctgggaaaatcagagaaggcaaattcaggcgaattccataccgttactctgatgaattgaatgaaattattacgaggatgttaaacttaaaggattaccatcga

> unknown sequence 13

ctttgactactggggccgaggaacccgggtcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggcaccctcctccaagagcacctctgggggcacagcagccctgggctgcctggtcaaggactacttccccgaaccggtgacggtgtcgtggaactcaggcgccctgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactctactccctcagcagcgtggtgacc

gtgccctccagcagcttggagaaatcagaaatcagatatagaaatcagaxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxtatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatxxxxxxxxxxcagatatagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatagaaatcagatatacacccagacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggacaagaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgcccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaaacccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtggtggtggacgtgagccacgaag

**Descripción del proyecto:**

Durante el transcurso del semestre, los estudiantes constituirán grupos. Cada grupo deberá presentar un informe final, y realizar una presentación de los resultados, mismos que serán evaluados por los profesores del curso. Para el informe, así como para la presentación, usted deberá cumplir con los siguientes objetivos. Recuerde que esto se trata de Ciencia: no se quede con los aspectos superficiales. Analice el texto entregado, compare con la bibliografía existente y trate de ir más allá de lo evidente.

**Objetivos:**

1. **Comportamiento de la enfermedad:**
   1. Calcular el R0 de la enfermedad
   2. Proyectar el número de infectados y muertos en las ciudades de la Tabla 1 a la semana 10 de desarrollo de la enfermedad

\*Modelos de dispersión de enfermedades

1. **Experimento RNA-seq** 
   1. Graficar los cambios en los niveles de expresión de los genes obtenidos en RNA-seq utilizando la metodología del Volcano plot. Justifique la elección del parámetro de corte para determinar el umbral de sobre expresión.
   2. Determinar el nombre y anotación en base a Gene Ontology de los genes sobre expresados en el RNA-seq
2. **Identificación del agente patógeno:**
   1. Realizar una anotación completa de todas las secuencias entregadas utilizando las diversas herramientas bioinformáticas revisadas en clases. Para cada secuencia deberá presentar un alineamiento múltiple usando CLUSTAL contra las principales secuencias obtenidas por una búsqueda BLAST. Debe presentar y justificar cada uno de los parámetros utilizados.
   2. A través de búsquedas BLAST y de la revisión de la literatura, determinar cuál es el agente patógeno en cuestión. Se deberán presentar los resultados BLAST, una justificación de sus parámetros, junto con los alineamientos y una justificación de la elección del agente patógeno.
   3. Una vez determinado el agente patógeno usted deberá bajar el genoma completo del organismo en cuestión y realizar una anotación completa de su genoma o de las porciones disponibles. Se deberán presentar la anotación genómica completa.
3. **Determinar potenciales blancos terapéuticos:**
   1. A partir de la anotación del genoma, determinar cuáles son los potenciales blancos terapéuticos, realizando una anotación completa de cada uno de estos. Se deberán identificar al menos dos blancos por grupo. (aquí usar péptidos antimicrobianos + antibióticos) (Anticuerpos?)(Antibióticos)(Antivirales)(Vacunas)
   2. Para cada blanco terapéutico se deberá determinar su estructura 3D, incluyendo una anotación detallada de su clase, arquitectura, topología y homología.
   3. Para cada blanco se deberán realizar un estudio de su(s) potencial(es) sitios de regulación.
   4. Para cada sitio se deberá realizar un estudio de docking de potenciales drogas con potencial actividad farmacológica sobre el blanco.
4. **Inferir la red de interacción génica:**
   1. Utilizando las secuencias humanas que resultaron sobre-expresadas, infiera una red de interacción génica y preséntela utilizando al menos dos algoritmos gráficos.
   2. Para la red, determine: coeficiente de clustering y grado promedio, camino largo promedio, geodésica, diámetro y probabilidad de conectividad
   3. Presente una anotación de todos los nodos que participan de la geodésica

Para la presentación de resultados, considere una exposición de 20 minutos en los que deberá incluir: introducción, objetivos, materiales y métodos, resultados y discusión. Las mismas secciones deben aparecer en el informe escrito, que ademas deberá tener una extension máxima de 10 paginas de texto, portada, indice, figuras y referencias aparte.

Buena suerte!!!