ESERCITAZIONE 1 – COVID-19

**CONSEGNA**

Questa esercitazione è organizzata per punti, ovvero per semplici operazioni che potrete effettuare utilizzando un semplice internet browser, sfruttando un po’ tutte le conoscenze acquisite durante questo corso. Per semplicità, potete provare a compilare questo file con i risultati delle vostre ricerche. Per quanto questo laboratorio non richieda alcuna compilazione obbligatoria, ma scrivere le risposte potrebbe tornarvi molto utile per confrontarvi con compagni di corso, fissare meglio alcuni conetti utili per lo studio oppure semplicemente inviarli a [mgerdol@units.it](mailto:mgerdol@units.it) per eventuali chiarimenti. Prendetela come una sorta di gioco e come uno strumento di autovalutazione. Siate curiosi, è sempre possibile partire da questa traccia per prendere vie alternative ed espandere questa ricerca a vostro piacimento. L’intero svolgimento di queste attività di laboratorio con le “soluzioni” è disponibile in un video il cui link è disponibile su Moodle nella sezione apposita, ma il mio consiglio è quello di provarci in prima battuta da soli, magari riprendendo i contenuti delle slides viste a lezione, ed affidarsi al video solo in casi di obiettiva difficoltà, oppure per controllare se ciò che è stato fatto è corretto.

1. **Utilizziamo Pubmed. Quanti articoli sono stati ad oggi pubblicati che contengono la parola “SARS-CoV-2” nel titolo o nell’abstract?**
2. **Provate a tracciare quanto è variato il numero di articoli pubblicati sull’argomento, mese per mese, da gennaio ad aprile.** SUGGERIMENTO: ricordate le funzionalità dell’”advanced search”
3. **Quanti articoli sono stati pubblicati relativamente all’utilizzo dei seguenti farmaci nella terapia da Covid-19? Prendiamo in esame “hydroxychloroquine”, “convalescent plasma”, “heparin”, “Tocilizumab” e “Avigan””.** L’Avigan è il farmaco antiinfluenzale reso celebre da un finto farmacista italiano che vive in Giappone, grazie ad un video “virale” lanciato su Facebook, senza alcuna evidenza scientifica.
4. **Questi risultati dovrebbero darci un’idea di quanto siano stati esplorati i suddetti farmaci come potenziali agenti terapeutici per Covid-19. Ora, per curiosità, colleghiamoci a Google Trends (**[**https://trends.google.com/trends/explore**](https://trends.google.com/trends/explore)**) e verifichiamo quanto a questo interesse scientifico sia corrisposto un effettivo interesse dei media e dell’opinione pubblica.** Possiamo fare questo confronto sia per quanto riguarda gli Stati Uniti (dove Trump ha largamente spinto verso un utilizzo dell’idrossiclorochina) e per quanto riguarda l’Italia (dove invece l’effetto del video virale si è fatto sentire. Ricordate di tradurre i termini in italiano per questa ricerca!).
5. **Chiediamoci ora quali sono gli studi che sono stati fino ad ora più citati in letteratura che riportino nel titolo la parola “COVID-19”. Individuate i 3 articoli più citati.** Suggerimento: esistono altri strumenti per la ricerca di letteratura scientifica più avanzati di Pubmed… ma potreste avere bisogno di connettervi alla rete universitaria via VPN per accedervi (seguite le istruzioni da <https://vpn-client.units.it>, vi tornerà molto utile anche per le prossime esercitazioni e più in generale per poter accedere anche da casa ad articoli altrimenti inaccessibili.
6. Il secondo studio, pubblicato sul New England Journal of Medicine, è molto interessante in quanto è stato il primo a riportare la sequenza genomica del virus. **Uno degli autori è George Fu Gao. Riuscite a recuperare il suo H-index? Di che cosa si è occupato principalmente questo ricercatore nella sua carriera?** (SUGGERIMENTO: provate ad ordinare le sue pubblicazioni per il numero di citazioni ricevute, oppure a navigare su altre schede dal profilo dell’autore).
7. Proviamo ora a cercare la sequenza del genoma del virus su NCBI Genomes. Il nome ufficiale dell’organismo è “Severe acute respiratory syndrome coronavirus 2”. **Quanto è grande il genoma? Quanti geni contiene? Quale è il suo contenuto in GC?**
8. Colleghiamoci quindi ad NCBI Taxonomy. **Come viene classificato questo virus? Quali sono i suoi più stretti “parenti”?**
9. Forse avrete sentito parlare del pangolino come una dei possibili serbatoi “intermedi” in cui questo virus abbia potuto svilupparsi prima di fare il “salto di specie”. **Tra i virus strettamente imparentati a quello responsabile del Covid-19 riuscite ad identificarne qualcuno riconducibile ai pangolini?**
10. **Riuscite ad identificare le sequenze nucleotidiche associate a questo virus? Siete in grado di recuperare l’accession ID di questa sequenza?** (SUGGERIMENTO: usate le cross-references!). **Quali sono le dimensioni del primo isolato del coronavirus del pangolino rispetto a quelle del SARS-CoV-2?**
11. Ritorniamo alla letteratura… Usiamo Pubmed. **C’è qualche pubblicazione scientifica che ci faccia pensare, in tempi non sospetti, ad una associazione tra pangolini e coronavirus?**
12. Se siete riusciti ad ottenere un risultato interessante, cliccate su PMC – Full text ed andiamo ad analizzare l’abstract un po’ più nel dettaglio. Fatto ciò, scendiamo alla Tabella 3. Questa riporta dei risultati di analisi di BLAST ottenute confrontando delle sequenze virali ottenute dal pangolino con quelle ritrovate come subject all’interno del database GenBank. **Come commentereste questi risultati in termini di e-value, percentuale di identità e lunghezza degli allineamenti? In altre parole vi sembrano risultati attendibili?**
13. Osserviamo ora la posizione della sequenza “Pangolin07\_contig 1420” nell’albero filogenetico riportato nella Figura 5. **Come potrebbe essere definito il suo grado di parentela con il clade che contiene molti betacoronavirus, tra cui anche uno umano (indicato dall’accession ID FJ415324.1). Si tratta di un raggruppamento ben supportato?**
14. Prendiamo proprio in considerazione ora la sequenza umana citata sopra e recuperiamo la scheda relativa alla proteina spike di questo virus (SUGGERIMENTO: utilizzate sempre le crossreferences!). **Quale è l’accession ID della sequenza? Quanto è lunga? Ricavate la sequenza in formato FASTA.** Lasciate aperta questa scheda, ci servirà più tardi!
15. Analizziamo la topologia della proteina. **Si tratta di una proteina secretoria? Contiene domini transmembrana? In caso affermativo, in che regione?**
16. **Analizziamo quali domini sono presenti nella proteina spike, ed esploriamo la funzione in particolare del dominio centrale.**
17. Proviamo ora ad esplorare nel dettaglio l’interazione tra il dominio di legame proteina spike (in questo caso del virus della SARS) ed il recettore ACE2 umano. **Utilizzate l’accession ID “2ajf” per recuperare le informazioni di interesse dal database che ritenete più opportuno.**
18. **Che struttura secondaria prevalente ha il dominio della proteina spike? Quale ha invece il recettore ACE2? Come faccio a capire quale è una e quale è l’altra?**
19. Proviamo ora a confrontare la sequenza della proteina spike del coronavirus umano da cui siamo partiti con quella di altri virus decisamente più temibili. **Che strumento dovremmo utilizzare?**
20. Selezioniamo quindi per la nostra ricerca le specie di interesse. Eseguiamo una prima ricerca selezionando “MERS”, una seconda digitando “SARS-CoV-2” ed una terza digitando “Bat SARS coronavirus HKU3-4”. **Analizzate e commentate i risultati, con riferimento al migliore hit.**
21. **Recuperate la sequenza in formato FASTA delle 4 proteine.**
22. **Proviamo ora ad effettuare un allineamento multiplo tra le 4 sequenze tramite Clustal Omega**.
23. **In quali regioni vi sembra ci sia un maggiore/minore grado di conservazione?** (SUGGERIMENTO: aiutatevi con l’utilizzo dei colori, che possono essere mostrati!)
24. Sempre utilizzando la sequenza del virus umano come sonda, chiediamoci ora se esistono proteine omologhe alla proteina spike anche in altri virus, che però non siano Coronaviridae. **Come fareste? Provate ad effettuare una ricerca e commentate i risultati.**
25. Ultima domanda. Se volessi individuare le coordinate del gene codificante la proteina spike nel genoma di SARS-CoV-2, partendo però dalla sequenza proteica della proteina spike del virus responsabile della MERS, come poteri fare? **Provate ad individuare queste coordinate utilizzando lo strumento che vi sembra più appropriato.**