ESERCITAZIONE 3 – GENOME BROWSERS

**CONSEGNA**

Questa esercitazione è organizzata per punti, come quella precedente, quindi restano valide le stesse indicazioni. A differenza delle due precedenti esercitazioni, in questo caso faremo largo utilizzo dei genome browser (in questo caso Ensembl) e degli strumenti collegati come OMIM e COSMIC

**LA CERAMIDASI ACIDA**

Iniziamo la nostra ricerca dal database NCBI Gene. Spesso i geni sono caratterizzati da nomenclatura inconsistente, con l’utilizzo di sinonimi.

1. Provate ad effettuare una ricerca del gene relativo al gene di interesse, che è membro della famiglia delle ceramidasi, su NCBI gene. **Quale è il nome ufficiale del gene? Quali sono i suoi sinonimi? Quale è il suo gene symbol**?
2. **Utilizzando i link presenti su questa pagina, che ormai avrete imparato ad utilizzare efficientemente, recuperate la sequenza proteica in formato FASTA della variante X1, assieme al suo accession ID ed alla sua lunghezza**.
3. Spostiamoci su GeneCards (<https://www.genecards.org>), un portale molto carino di cui non abbiamo parlato a lezione, ma che è molto comodo in quanto comprende numerosi link a molte risorse che abbiamo studiato. **Cercate il gene di nostro interesse e recuperate da questa scheda il link ad OMIM**. Lasciate aperta la scheda di Genecards, ci tornerà utile più tardi.
4. **A quali malattie genetiche ereditarie sono collegate mutazioni genetiche a carico di questo gene?**
5. **Da che tipo di ereditarietà sono contraddistinte?**
6. Scegliete una delle malattie prese in esame e **cercate di ottenere tramite una ricerca in un database di letteratura scientifica una o due «review» (articoli generali e solitamente molto citati) che descrivano bene la sue caratteristiche, la sua diffusione ed eventualmente le possibili terapie**
7. Tornando ora alla scheda di GeneCards, **provate a recuperare l’Ensembl Gene Id e collegatevi alla scheda ad esso relativa in questo database.**
8. **A quante varianti di splicing alternativo da origine questo gene**?
9. **Quali sono le varianti che danno origine alla molecola di mRNA ed alla proteina di maggiori dimensioni? Quale di queste corrisponde alla variante X1 che avete visto prima?** Segnatevi l’accession ID del trascritto, ci servirà più tardi.
10. **Su quale cromosoma è localizzato questo gene e quali sono le sue coordinate genomiche?**
11. **Quante Kb di sequenza copre in tutto questo gene? Si tratta di un gene codificato sullo strand positivo o negativo?**
12. **Quali sono i geni fiancheggianti al 5’ ed al 3’?**
13. **Per quanto riguarda l’annotazione funzionale di questo gene a quali termini Gene Ontology Molecular Function è legato?**
14. nella sezione «variant table», è possibile recuperare informazioni relative alle mutazioni deleterie relative al gene ASAH1 finora descritte. Centriamo la visualizzazione su un singolo esone, specificando le seguenti coordinate genomiche: 18069750-18069900: **quali/quante varianti genetiche marcate come benigne riuscite ad identificare**?
15. Prendiamo in esame la variante rs1049874. **Si tratta di una mutazione frequente nella popolazione mondiale? C’è qualche gruppo etnico che vi sembra mostrare una frequenza decisamente minore rispetto alle altre?**
16. **Si tratta di una sostituzione sinonima o non sinonima? Se si tratta di una sostituzione non sinonima, a che sostituzione amino acidica corrisponde?**
17. Recuperate ora l’accession ID del trascritto di maggiori dimensioni dal punto 9. **Sapete localizzare la posizione di questa mutazione (cioè le sue coordinate) nel mRNA, nel CDS e nella proteina?**
18. **Di che tipo di mutazione è la rs750374592? Quale è la sua frequenza nella popolazione mondiale? Che effetto potrebbe avere questa mutazione sulle proteine codificate da questo gene?**
19. Collegatevi ora a COSMIC (<https://cancer.sanger.ac.uk/cosmic/>) ed **identificate le mutazioni somatiche associate a questo gene. Si tratta principalmente di SNPs, di inserzioni o di delezioni?** SUGGERIMENTO: aiutatevi con il tab «mutation distribution»
20. **Riuscite ad individuare se è presente in catalogo una variante somatica in posizione 60 (rispetto agli amino acidi della proteina codificata). A quale sostituzione amino acidica corrisponde?**
21. **A che tessuto e a quale tipo di cancro è stata associata questa mutazione?**
22. **Quale è la percentuale di campioni di tumore al pancreas, tra quelli testati per questo gene, dove sono state identificate mutazioni somatiche?**
23. **In che percentuale dei soggetti affetti da tumore allo stomaco si nota una sovraespressione del gene?**
24. Collegatevi ora ad Human Protein Atlas. **A livello proteico, in quali tessuti sembra essere maggiormente espresso il gene ASAH1?**
25. **E a livello trascrittomico? Vi sembra che i dati HPA GTEx e FANTOM5 siano concordi tra loro?**
26. **Se dovessi scegliere una linea cellulare dove questo gene è particolarmente espresso, quale dovrei selezionare?** SUGGERIMENTO: aiutatevi con il cell atlas.
27. Non ci sono dato relativi alla localizzazione subcellulare (presumibilmente per la mancanza di un anticorpo), ma questi dati sono disponibili per il suo partner molecolare saposin D (gene ID: PSAP). **Aiutandovi con il segnale dell’anticorpo fluorescente, riuscite ad identificare in quale compartimento cellulare sia localizzata principalmente la proteina?**
28. Torniamo ora su Ensembl. **Riuscite ad identificare se il nostro gene di interesse ha un gene paralogo? Se sì, quale è il suo nome?**
29. **Su quale cromosoma è localizzato questo gene e quali sono le sue coordinate genomiche?**
30. Proviamo ora ad effettuare una piccola analisi di sintenia. **Su quali cromosomi dell’orangutan sono presenti il gene della ceramidasi acida e quello del suo paralogo?**
31. **Se invece prendessi in considerazione l’opossum?**
32. **Proviamo ora a vedere che cosa accade con un’analisi di sintenia in zebrafish.** SUGGERIMENTO: non basatevi soltanto sul grafico! Nelle tabelle sottostanti di solito sono contenute mole più informazioni.
33. Come ultimo task collegatevi a Genomicous (<https://www.genomicus.biologie.ens.fr/genomicus-98.01/>). Facciamo una semplicissima analisi di genomica comparata. **Cercate il nostro gene di interesse in uomo: otterrete una grafica comparativa dell’organizzazione genomica dei blocchi di sintenia in cui il gene è presente. Tra i due geni fiancheggianti NAT1 e PCM1, quale dei due vi sembra maggiormente conservato evolutivamente (cioè quale dei due tende a mantenere la stessa posizione andando “a ritroso” nell’albero evolutivo)?**
34. **Notate qualcosa di particolare per quanto riguarda questo gene (ASAH1) nel cavallo?**