

Bioinformatics: 2016 Final Project

在一個探討轉譯調控機制的研究中，我們利用 RNA pull-down 的方法找出一些有趣的 RNA-binding protein。其中 LC/MS/MS 蛋白質體學工具分析出來的資訊顯示有 5 段 peptide 與 heterogeneous nuclear ribonucleoprotein 吻合。

1. MGANSLER
2. MMNGMKLSGR
3. MGQTMERIGSGVER
4. FNECGHVLYADIKMENGKSK
5. GIGMGNIGPAGMEGPFGGGMENMGR

為了進一步研究此基因在轉譯調控機制所代表的意義，請各小組進行討論分析試著回答以下的問題：

1. 請找此基因在染色體上的位置。請簡單描述此基因的大小及結構 (exon, intron, splicing site, 等)。已知此基因在 NCBI 的資料庫中有三不同蛋白質產物 (isoform)，請比較此基因的不同蛋白質產物，找出不同蛋白質產物產生的原因 (alternative splicing? Alternative promoter?)？請說明不同蛋白質產物其 exon 的組成。
2. 依據 LC/MS/MS 蛋白質體學工具分析出來的資訊能否知道 RNA pull-down 實驗是哪一個 (或多個) isoform? 請在資料庫中搜尋此基因在不同組織中 RNA 及 protein 的表現情形? 如何偵測此基因不同 isoform 的表現狀況? 利用 GEO 資料庫，是否可以找出此基因可能與癌症或其他疾病相關的證據? 有那些資料顯示其此基因表現量在何種情況下會發生變化?
3. 利用資料庫資料，分析此基因不同蛋白質 isoform 的 5'UTR，並分析 5'UTR 來推測其是否具有可能的調控因子來影響 isoform 的表現? 此外並請找出是否有 microRNA 會 target 到該基因 3'UTR? 分析的此基因的 5'-flanking sequence (-1000bp) 並推測其可能參與轉錄調控之因子 (promoter, transcription factor...)?
4. 請分析此蛋白質的特性 (IP、subcellular localization、functional domain、homolog proteins in other species.....)。不同 isoform 是否改變了蛋白質的特性?
5. 請利用資料庫或其他網路資源找出可能與此蛋白質相關的作用分子 (protein? DNA? RNA?....)。並試著利用蛋白質交互作用及 pathway analysis 工具來預測此蛋白質的生物功能。依據你的分析，你可以解釋此蛋白質如何參與轉譯調控嗎?

You will have approximately four weeks to finish your project. ***There is no restriction to the tools you use. You can apply any computational tools (online or desktop programs) to help you to complete your project. But you have to report the name and reference for the program you used in this project.*** After you have completed your project, upload the results to your web directory before 12/28/2016. There will be ~30 minutes project presentation /discussion for each group at the end of the semester. If you have any questions regarding this project, you are welcome to post your question on the forum, or ask any of the teaching assistants. Good luck and have fun!