

# 企劃書綱要

近年來的基因科學發展非常快速，尤其在 DNA 甲基化的議題以及研究上，在現今備受關注，科學界已經明確的認定，DNA 的『甲基化』，是一個控制基因表現與否的重要的關鍵，所以，『甲基化』也成為基因表觀調控的重要標誌[1]。在做 DNA 相關實驗時，會使用管家基因做為參考基準，因為管家基因擁有在生物上的表現需穩定以及高表現的特性，以致許多 DNA 實驗會使用多個管家基因的表現量作為其他基因表現的參考，管家基因被廣泛用作實驗和計算研究的內部對照[2]，所以會挑選管家基因來作為 DNA 實驗的基準參考，使得在做 DNA 實驗時可以避免對獲得的結果進行誤解，從而導致錯誤的假說[3]。

因特定基因在不同病狀的表現量不同，導致不同實驗條件下的差異，可能會造成所使用的管家基因無法去作為通用參考的基準，且在現今 DNA 甲基化的蓬勃發展，但在目前在 DNA 甲基化可泛用的管家基因資料非常少許，所以本研究的目的是為了找出 DNA 甲基化中通用的管家基因。計畫方向是將透過全基因體甲基化晶片取得慢性病患者的甲基化表現數據，綜合病患的共病資料分析，以 K-means 等非監督式學習的方式，把資料集做分群，再分別挑出各群中高表現且差異性小的甲基化基因位點，與現在醫界常運用的管家基因去做對比，找出有交集的集合，最後再用 K Nearest Neighbor (KNN) 等監督式學習的迴歸分析技術，挑選出對於 DNA 甲基化分析最佳、最適用的管家基因，可作為未來對於 DNA 甲基化分析時可以有管家基因的參考基準，使得實驗結果降低錯誤假設的機率。

[1]

莊樹諄 ( 2013 ) 。就算是甲基化，也要看位置！。中央研究院 基因體研究中心

[2]

Eli Eisenberg & Erez Y Levanon (2013). Human housekeeping genes, revisited

[3]

Veronika Kloibert & Lothar Rink (2019). Selection of an inadequate housekeeping gene leads to misinterpretation of target gene expression in zinc deficiency and zinc supplementation models. Journal of Trace Elements in Medicine and Biology, 56, 192-197