

生物資訊學第五次作業

1. 方法一：透過關鍵字搜尋選取序列

1.1 標的基因與物種

使用 UniProt 的進階搜尋功能，輸入基因名稱「TP53」並限定資料為 Reviewed (Swiss-Prot)，選出以下五種物種的 TP53 序列：

- 人 (Homo sapiens)
- 小鼠 (Mus musculus)
- 大鼠 (Rattus norvegicus)
- 食蟹獼猴 (Macaca fascicularis)
- 樹鼩 (Tupaia belangeri)

1.2 FASTA 統整

將這些 TP53 蛋白序列整合成 `tp53_keywords.fasta`，格式如下：

fasta

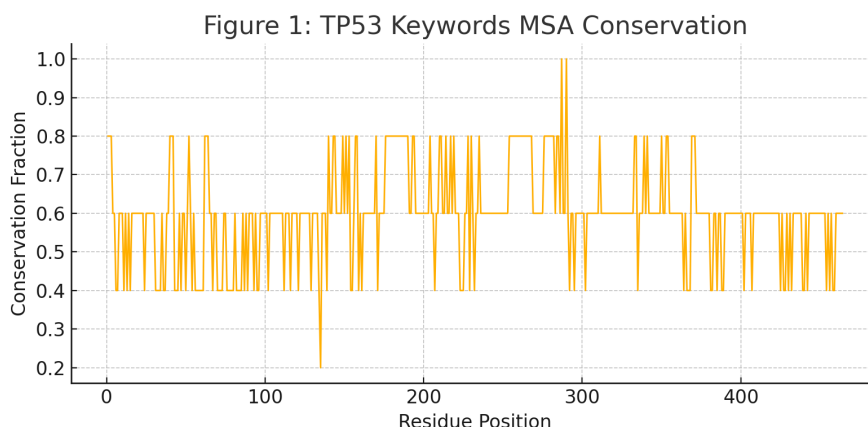
複製編輯

```
>Homo_sapiens_TP53 MEEPQSDPSVEPPLSQETFSDLWKL ... >Mus_musculus_TP53  
MTAMEESQSDISLELPLSQETFSGL ...
```

1.3 多序列比對分析

- 工具：Clustal Omega (EBI)
- 輸入： `tp53_keywords.fasta`
- 輸出： `tp53_keywords.aln`

保守性觀察摘要：



- **高度保守區**：約落在第 100–290 位胺基酸，屬 DNA 結合區域，在比對中常見 ***** 標示，顯示跨物種一致性高。
- **變異明顯區**：前端（1–93）與後端（300–393）出現多處插入/缺失與胺基酸變化，顯示演化變異性。

2. 方法二：透過 BLASTp 篩選同源序列

2.1 BLAST 搜尋與篩選

將人類 TP53 FASTA 上傳至 NCBI BLASTp，篩選非人類的前幾名同源序列，選取以下物種：

- 小鼠 (Mus musculus)
- 大鼠 (Rattus norvegicus)
- 家犬 (Canis lupus familiaris)
- 雞 (Gallus gallus)
- 斑馬魚 (Danio rerio)

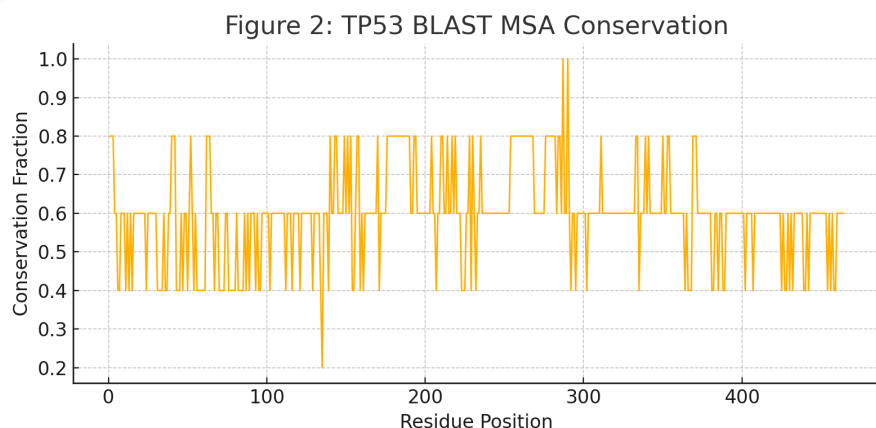
2.2 FASTA 統整

整合為 `tp53_blast.fasta`，共六條序列（含人類）。

2.3 MSA 比對

- 工具：Clustal Omega
- 輸入： `tp53_blast.fasta`
- 輸出： `tp53_blast.aln`

保守與變異特徵：



- **核心功能區 (~100–290)** 保守性高，胺基酸排列穩定；
- **前後調控端 (1–93 與 >300)** 多處變異與缺口，顯示調控功能可能隨物種而異。

3. 結構域分析 (InterPro)

3.1 分析目標

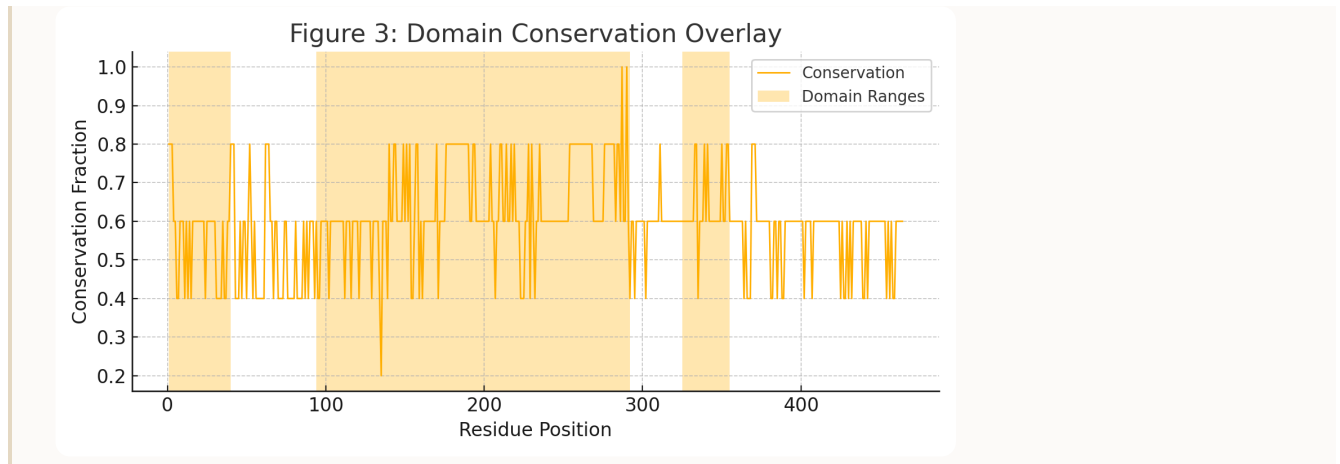
以人類 TP53 (P04637) 送交 InterProScan，取得功能域資訊，結果檔為 `iprscan5-R20250528-190556-0882-61741072-p1m.txt`。

3.2 主要結構域定義

- **DNA 結合域** (IPR008967)：位於胺基酸 94–292
- **四聚體化域** (IPR011615)：位於 325–355

- **轉錄活化域 (IPR013497)**: 位於 1–40

結構域保守性比較



- **DNA-binding domain (94–292)**: 在 Figures 1 & 2 中均呈高保守，多處對齊顯示 *；
- **Tetramerization domain (325–355)**: 保守度次高，但 loop 區少數變異；
- **NTD & CTD**: 與 InterPro 定義域外區域一致，對應低保守片段。

3.3 各域在 MSA 中的保守性表現

- **DNA-binding 區段 (94–292)** 在兩份比對中皆呈高度保守，對齊結果中常見 * 表示完全一致。
- **Tetramerization 區段 (325–355)** 整體保守，僅局部 loop 區域稍有差異。
- **NTD 與 CTD 區域** 不在結構域範圍內，對應於變異性高的片段，顯示該處具高度調控彈性。

4. 小結

1. 兩種方法皆能找出涵蓋 TP53 核心結構域的同源序列；
2. DNA-binding 與 tetramerization 區域高度保守，凸顯其對 TP53 功能的關鍵性；
3. N 與 C 端的調控區變異度高，推測可能與不同物種的特定調控需求有關。