

◎ 張培均

利用基因體資訊 預測最佳用藥

疾病的發生機率不僅與個人的基因體特徵有關，
治療疾病的療法跟治療成效也跟個人的基因體特徵有關。
因此，對於疾病的預防、治療及預後，都可利用基因體資訊來評估，
以朝向精準化醫療發展。



基因體學的興起

細胞的運作與功能都是由細胞內染色體上儲存的訊息所主導。染色體是一條 DNA 大分子，人類細胞內共有 23 對染色體，其中包括 22 對體染色體，以及一對 XY 或 XX 染色體。這些染色體通常以細絲狀存在於細胞核內，若把染色體拉成直線，大約有 1.83 公尺長。

染色體內約含有 30 億個 DNA 鹼基對，鹼基對是由兩個含氮鹼基以氫鍵結合的結構，由胸腺嘧啶（T）、腺嘌呤（A）、胞嘧啶（C）和鳥嘌呤（G）4 種鹼基排列成鹼基序列，其中 A 與 T 之間由兩個氫鍵連接，G 與 C 之間由 3 個氫鍵連接。這些鹼基序列並不是隨機排列，而是有特定的樣式跟訊息編碼。這些樣式包括所有的基因與非基因，藉此構成了整個生命的藍圖。

基因體學是一個跨領域的科學，專注於生命藍圖的序列、結構、功能、演化及註解，因此基因體科學的興起，首先必須克服的障礙就是 DNA 鹼基序列的定序與分析。也就是說，基因體科學的興起是建立在 DNA 定序技術的突破及資訊技術的應用，也因此基因體學是生物資訊學的一個重要領域。

基因體學的目標

針對不同的細胞狀態建立完整的基因體資訊，是現今基因體學的主要目標，資訊的內容主要有基因表現、基因變異、表觀基因體、結構基因體及功能基因體。

基因表現是指細胞內染色體轉錄產生各種功能性 RNA 的產量，其下游產物通常是蛋白質，也包含非蛋白質基因如轉移 RNA（tRNA）、微 RNA（microRNA）等基因。不同的細胞狀態有不一樣的基因表現

樣式，因此透過基因表現的樣式比對或資料探勘，可以分辨不同的細胞狀態，例如癌細胞跟正常細胞的分辨，以及癌症相關基因的探勘。

基因變異是指群體中個體之間 DNA 序列的差異，變異發生在生殖細胞或體細胞中，突變是變異的主要來源，是 DNA 序列的永久性改變。當 DNA 複製過程中出現錯誤而未被修復糾正時，會發生新的突變，突變可能對生物體有害，例如增加致癌機率。單一核苷酸多型性（SNP）是基因變異的一個主要類別，在人類基因體中，每隔 100 至 300 個鹼基就會有一個 SNP 位點，因此單一核苷酸多型性是造成個體間差異的主要來源，也是未來疾病治療朝向個人化醫療的關鍵。

表觀基因體是指細胞中 DNA 或組蛋白的修飾，DNA 甲基化和組蛋白修飾是兩種最具特徵的表觀基因修飾。表觀基因修飾並不會改變 DNA 序列，但是在基因表現和調控上有重要的影響，參與了許多細胞過程，如細胞分化、個體發育、腫瘤的發生等。

結構基因體在於探討所有基因的蛋白質三維結構，利用全基因體序列，透過高通量的實驗和建模方法，可以更快地完成基因體蛋白質的結構測定。結構基因體採用大量的建模方法，由於相似的蛋白質序列有相似的蛋白質結構，相似的蛋白質序列稱為同源蛋白，因此利用已知結構的同源蛋白質可以預測未知的蛋白質結構。也有僅基於蛋白質的物理化學特性的計算方法，經由結構基因體的研究來確定蛋白質的結構，而後就能從其結構確定蛋白質功能。

功能基因體著重於基因的功能以及基因之間的互動，而不是基因體資訊的靜態儲存，例如基因轉錄的調節、基因與基因之間的互動、蛋白質與蛋白質之間的相互作用等。以基因體整體的角度來探討細胞內的各種功能，是功能基因體學的一個主要特徵。

藥物基因體學的目標在於針對患者的基因體特徵，
開發優化藥物或評估最佳的治療藥物，以確保最大的效率和最小的副作用。

藥物基因體

近幾年的研究顯示，疾病的發生機率不僅與個人的基因體特徵有關，治療疾病的療法和治療成效也跟個人的基因體特徵有關。因此，對於疾病的預防、治療及預後，開始朝向個人化與精準化發展。加上基因定序成本的降低，藥物基因體學的研究隨之興起。藥物基因體是從基因體特徵的角度，分析個體如何影響其對藥物的反應，通常是透過基因表現或基因變異特徵（例如單一核苷酸多型性、序列片段重複、序列遺失、序列剪接等）。

藥物基因體學這個詞彙最早出現於 1959 年的德國，主要在探討人體細胞的基因表現或基因變異，對於藥物活性（藥物與作用標靶的作用效果）與藥物動力學（藥物吸收、分布、代謝、消除、毒性）的影響。基因體特徵差異影響藥物反應的可能機制有二種，一是改變體內的藥物代謝動力學性質，二是影響與藥物標靶蛋白的作用力。例如胺基酸突變改變藥物與標靶蛋白的結合方式，最終改變藥物的活性。

藥物基因體與臨床治療息息相關，藥物基因體學的目標在於針對患者的基因體特徵，開發優化藥物或評估最佳的治療藥物，以確保最大的效率和最小的副作用，減少治療處方的反覆試驗。未來醫生可以根據每個病人的基因體特性，獲得更好的治療療效，以及最小化藥物毒性的發生和藥物不良反應，朝向精準醫療甚至是個人化醫療的理想境界。

近來已有許多以基因特性為基礎，對藥物反應或副作用進行評估，以得到較為明確的用藥選擇的例子。例如美國食品暨藥物管理局就建議在使用某些會受基因特徵影響的抗癌藥物前，必須進行基因型測試，以確保治療的有效性與安全性。

近來已有許多以基因特性為基礎，對藥物反應或副作用進行評估，
以得到較為明確的用藥選擇的例子。

美國食品暨藥品管理局建議在使用前做 特定基因型測試的化療藥品摘錄

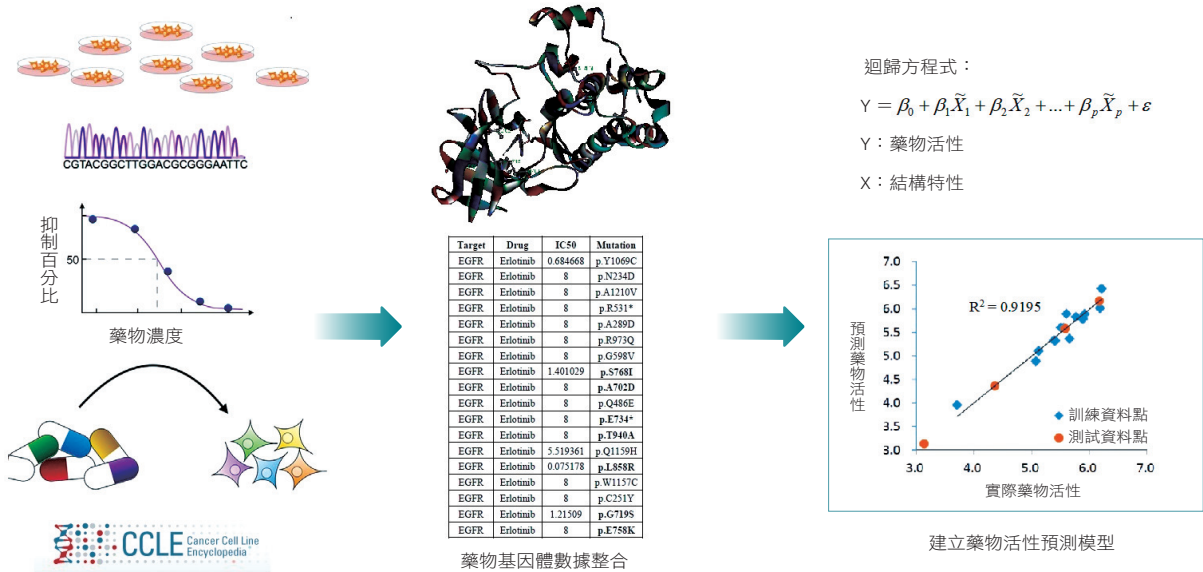
藥物名稱	受測基因	治療的癌症與可能的副作用
Trastuzumab	HER2 基因是否過度表現	乳癌
Lapatinib	HER2 基因是否過度表現	乳癌
Cetuximab	EGFR 基因是否過度表現、KRAS 基因是否突變	直腸癌
Erlotinib, gefitinib	EGFR 基因是否突變	EGFR 基因相關癌症
Irinotecan	UGT1A1 基因是否突變	毒性

以下根據美國食品暨藥品管理局的資料，簡介幾個相關的抗癌藥物。

Trastuzumab 一是利用單株抗體對抗致癌基因 **HER2** 的藥品。**HER2** 是一種生長激素受器蛋白，接受訊息而加速細胞的生長與分裂，當 **HER2** 基因過度表現時，會造成細胞快速生長而失控。臨床上發現，在原發性乳癌病人中約有 25% 到 30% 其癌細胞有 **HER2** 基因過度表現的現象，而這些乳癌患者比無 **HER2** 過度表現的患者有較差的預後，病人存活時間也較短。

Trastuzumab 可專一地作用在腫瘤細胞表面的 **HER2** 蛋白上，阻礙 **HER2** 的訊號過度傳遞，進而減少腫瘤細胞增生，但在使用前須先確認是屬於 **HER2** 基因過度表現的癌症類型才適合投藥。

Lapatinib 一是小分子酪胺酸激酶抑制劑，可同時抑制 **HER2** 及 **epidermal growth factor receptor (EGFR)** 的酪胺酸激酶活性，阻斷受體的磷酸化反應，因而阻斷下游訊息



藥物基因體學探討基因變異與藥物活性的關係

傳遞的進行，可用於治療對 herceptin 具有抗藥性的乳癌。但在使用前須先確認是 HER2 基因過度表現的癌症類型才適合投藥。

Panitumumab 和 cetuximab 一兩者都是可與腫瘤細胞的表皮生長因子受體結合的單株抗體，藉由抑制 EGFR 的作用進而抑制癌細胞的分裂與增生，可用於標準化療藥物治療無效的 EGFR 型轉移性大腸直腸癌。研究顯示大約有近 4 成的大腸直腸癌病人會出現 KRAS 基因突變，導致化療藥物 cetuximab 或 panitumumab 治療失效，因此 cetuximab 和 panitumumab 僅可用於 KRAS 基因未突變的大腸直腸癌。

基於基因體資訊的抗癌藥物開發

藥物開發的臨床試驗往往複雜且昂貴，若能預先建立預測模型，並盡可能捕捉癌症的特徵與可能的藥物反應，將有助於抗癌藥物開發的成功。因此，大量且多樣的基因體

與藥物活性相關的實驗無法避免，癌細胞株百科資料庫（Cancer Cell Line Encyclopedia, CCLE）就是為了這個目的而建立的。

CCLE 資料庫總共收集 947 種人類癌細胞，並使用其中約 500 種細胞與 24 個抗癌藥物進行藥物基因體實驗，這些細胞涵蓋 36 種癌症類型，所有的細胞都進行了基因體學的研究，並定序其中約 1,600 個變異基因，最後分別蒐集這些細胞株的基因體變異及基因表現資料。

在抗癌藥物的探勘上，雖然癌細胞是一個相對容易進行實驗的生物模型，並已廣泛用於藥物開發，但為了建立藥效的預測模型，獲取更多癌症的分子特徵和治療反應是必要的，也才能利用這些資料發展藥物活性的評估方法。

基於基因體資訊評估抗癌最佳用藥

這裡以標靶 EGFR 的抗癌藥物的評估為例，簡單介紹抗癌最佳用藥的評估。標靶蛋白

藥物開發的臨床試驗複雜且昂貴，若能預先建立預測模型，並盡可能捕捉癌症的特徵與可能的藥物反應，將有助於抗癌藥物的開發。

突變而引起藥物抑制活性改變的原因很多，其中一種是標靶蛋白的胺基酸突變位於與藥物結合的位置，由於胺基酸的物理化學性質改變，也改變了與藥物的結合特性。另一種原因是胺基酸突變位置不在藥物結合的位置上，卻間接改變了結合位置的構形，影響結合位置與藥物的結合特性。因此突變胺基酸的物理化學特性，以及在標靶蛋白中的空間位置，應該含有足夠的資訊來評估藥物的抑制活性。

CCLE 藥物基因體資料庫提供了所需要的資料，例如 p.X100Y 表示 EGFR 標靶蛋白序列位置第 100 的胺基酸由 X 突變為 Y，其對應的藥物抑制活性資料是 IC50，是指藥物誘導腫瘤細胞凋亡 50% 的濃度值，稱為 50% 抑制濃度 IC50，即凋亡細胞與全部細胞數的比等於 50% 時所對應的濃度。IC50 值可以用來衡量藥物抗癌的能力，即抗癌能力越強，這數值越低，當然也可以反向說明某種細胞對藥物的耐受程度。

影響標靶蛋白與藥物交互作用力的因素，主要有氫鍵、離子靜電力、疏水性作用力及凡得瓦力，因此可以採用這些胺基酸的物理化學性質來評估藥物抑制活性。當然若能考慮所有的影響因素是最好不過了，可以在資料探勘的計算過程中，找出真正影響抗癌藥效的因素，並利用迴歸模型建立藥效的評估公式。

因此可以利用基因演算法建立迴歸模型，利用這個迴歸模型就可以針對標靶蛋白不同的胺基酸變異的特性，評估藥物對於 EGFR 的抑制效果。若病人的 EGFR 變異經評估後，發現藥物的抑制效果甚差，就可建議病人不要使用該藥物。

展望

隨著電腦硬體計算能力的大幅提升，人工智慧 (artificial intelligence, AI) 已經可以普遍應用在科學研究、醫療診斷，甚至在日常生活中。藥物開發也大量使用 AI 來加快藥物的發現過程，降低研發費用和臨床試驗的失敗率，最終產生優質藥物。生命科學中大量數據的可用性和機器學習演算法的快速發展，也導致近年來專注於 AI 藥物開發初創公司的發展與結盟。

醫療決策上的應用更是深具潛力，目前的人工智慧在醫療照護的主要議題包括：促進健康的雲端諮詢平台—透過即時數據蒐集與 AI 的分析，適時提供病患資訊，從而改善患者的治療效果；精準化醫療—結合患者的基因體與醫療紀錄，利用 AI 的輔助診斷提出醫療決策，使病患有最佳的醫療結果。

例如微軟一直致力於推出一款名為漢諾威的機器，這個機器的目的是整合治療癌症的相關資料庫，從而預測對個體患者最有效的藥物組合。又例如史丹佛大學的研究人員開發了一種基於人工智能的演算法，可以識別皮膚癌，其精準度和專業醫生一樣好。應用人工智慧於基因體與醫療紀錄大數據的整合分析，使得智慧醫療的願景指日可待。

張培均

亞洲大學生物資訊與醫學工程學系
