

新世代高通量定序 — 打開生物基因體的潘朵拉盒子

呂美曄

高通量定序技術的重要性及其對生物科學的影響

在生物科學的世界裡，基因體定序一直是探索未知的重要工具。隨著科技的進步，新一代高通量定序技術（Next-Generation Sequencing, NGS）已成為打開生物基因體秘密的「潘朵拉盒子」。但這個盒子裡究竟隱藏了什麼神秘內容呢？

首先，高通量的平行定序技術使我們能夠快速、準確地閱讀和分析數百萬甚至數十億個DNA序列。這相較於傳統桑格定序方法，不僅大幅度提升了定序的速度和規模也降低了成本，並演發出許多相關的生物資訊科學。因此，從學術研究到醫學診斷，NGS都開啟了新的可能性。

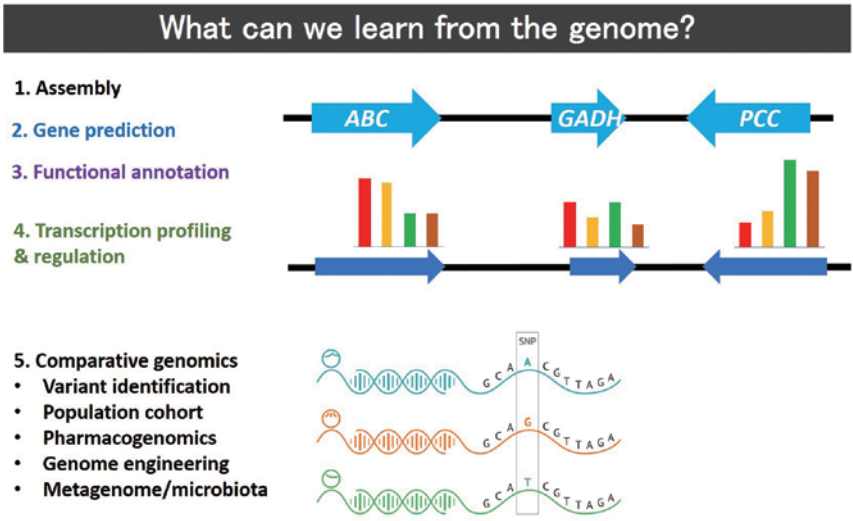
這項技術讓我們更深入地理解生物多樣性和演化歷程。例如，通過對不同物種基因體的比較分析，科學家可以揭示物種間的關聯性，甚至追溯到遠古共同祖先的遺傳信息。不僅能夠組裝出完整的基因體，還能深入瞭解基因

如何調控生物的生長、發育和適應環境。研究比較基因體學，我們可以發現那些在演化過程中保持穩定的基因，這些基因往往對物種的生存至關重要。

高通量定序所帶來的另一大優勢是其對疾病研究的影響。例如，在遺傳疾病和癌症研究領域，生物資訊學分析可以幫助我們識別出與疾病相關的基因變異，從而對病因進行更精確的定位，開發更為針對性的治療方法。如心血管疾病、某些類型的癌症，以及遺傳性疾病等，尤為重要。此外，透過分析病原體的基因體，我們能更迅速地應對傳染病爆發，並開發有效的疫苗。

在農業領域，高通量定序技術同樣發揮了重要作用。通過解析作物的基因體，科學家能夠培育出更為抗病、耐旱或產量更高的品種，這對於全球糧食安全具有深遠的意義。在環境生物學領域，基因體定序同樣發揮著關鍵作用。透過分析不同環境條件下生物的基因表達模式，科學家能更理解生物是如何適應各種環境壓力的機制。這不僅對於保護生物多樣性和生態系統具有重要意義，同時也為農業提供了改善作物耐逆性的方向。最後，我們還可以通過基因體定序深入了解生物體內的微生物群落。微生物對人類健康和環境都有著深遠的影響，通過基因體分析，我們可以更理解它們的功能，並在醫學、農業甚至是環境保護等領域找到新的應用。

高通量基因體定序為我們打開了一扇窗，讓我們能夠更全面、更深入地瞭解生物世界的多樣性和複雜性，為



圖一、我們可以從基因體定序中學到什麼？生物學家通過獲取高通量定序資料，可以進行生物資訊學分析，這不僅有助於實現高品質的基因體組裝，還能預測潛在的基因模型、推斷功能註釋，以及進行轉錄體分析。此外，高通量定序還促進了多種其他基因體分析應用的發展。

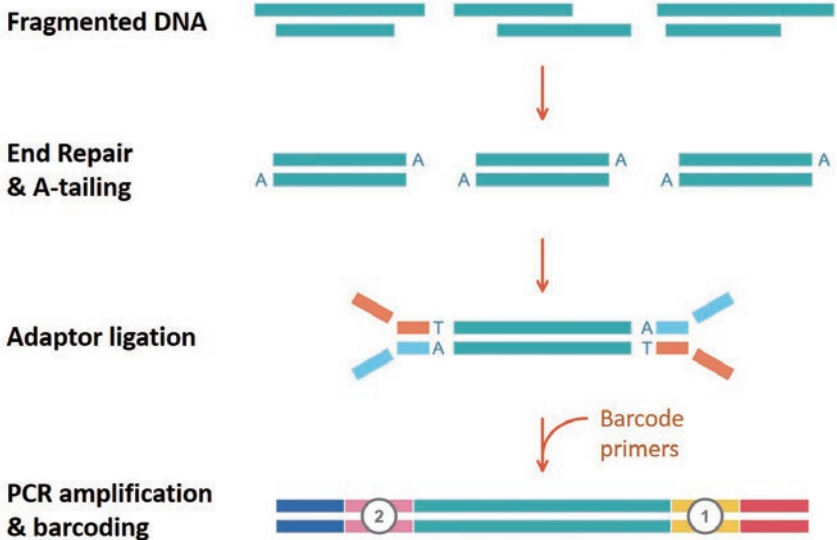
科學研究和人類社會的進步貢獻了巨大的力量。然而，這項技術也帶來了新的挑戰。其中一個重要問題是數據的解讀。基因體數據極為複雜，需要專業的生物資訊學工具和專業知識來進行分析和解釋。

關鍵技術：NGS文庫準備的步驟和重要性

次世代定序（NGS）文庫（Library）製備是一個精細且關鍵的過程，它決定了定序結果的品質和準確性。流程步驟大致如下（圖二）：

1. 從生物樣本中萃取高品質的DNA或RNA，確保定序的準確性和完整度。
2. 進行DNA或RNA的切割，將其斷裂成更小的片段，以適應不同定序平台的需求。（三代讀長序或可省略切割的步驟）。
3. 對這些片段進行末端修復和A尾加工。末端修復確保DNA片段的兩端平整，而A尾加工則是在DNA片段的3'末端添加一個額外的腺苷酸（Adenosine）。
4. 將文庫轉接子（Adaptor）連接到DNA或RNA片段上。這些轉接子含有必要的引子序列，用於後續的PCR擴增和定序。
5. 接下來是PCR擴增，增加足夠量的DNA片段以進行定序。在這個過程中，也會加入特定的條碼，以便在定序後將數據分配給相應的樣本。（三代文庫無需最終PCR擴增）。
6. 對文庫進行定量及定性和長度檢測，確保其適合進行定序。

NGS的樣品製備工作流程雖然複雜，但對於確保定序數據的質量和可靠性至關重要，也對後端的生物資訊分析影響甚鉅。



圖二、次世代定序（NGS）文庫建置工作流程。（由上至下）獲取所需的DNA片段、DNA末端修復與A尾加工、定序文庫轉接子連接、PCR擴增與條碼標記（Library Barcoding）。

定序技術的演進：從傳統的桑格定序，到第二代Illumina，再到第三代PacBio和牛津奈米孔技術

定序技術的發展日新月異，從最初的桑格定序（Sanger Sequencing），到第二代的Illumina定序，再進化到第三代的PacBio（Pacific Biosciences）和牛津奈米孔（Oxford Nanopore）技術，每一步的演進都大大地擴展了我們對基因世界的認識。

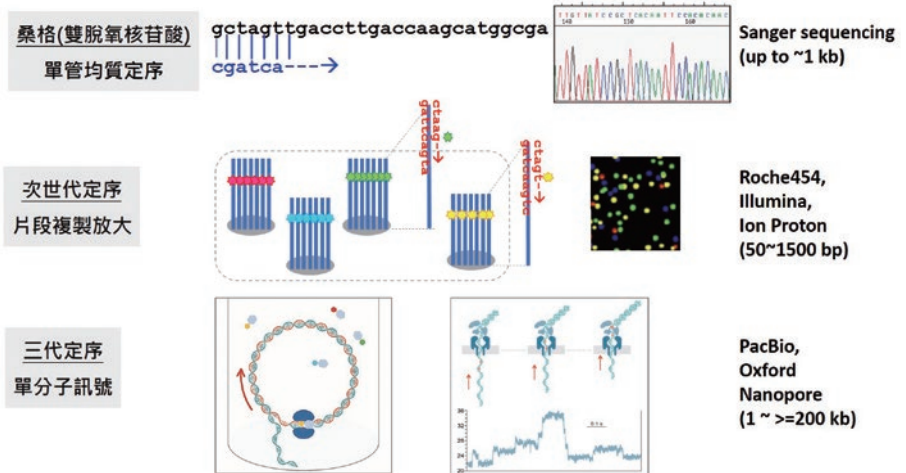
桑格定序，作為最早的DNA定序技術，其準確度高且可靠，為人類首次完成人類基因體計劃奠定了基礎。然

而，它的速度相對較慢，成本較高，不適合大規模的基因體定序。

進入21世紀，第二代定序技術，以Illumina為代表，大幅提高定序速度和降低成本。這類技術通過將DNA片段固定在固體表面上，並利用光學方法讀取生成的序列，實現了大規模平行定序。它在臨床診斷、癌症研究和個體化醫療等領域發揮了重要作用。

而第三代長片段定序技術，如PacBio和牛津奈米孔，則開啟了單分子即時定序技術的新紀元，可以更好地解析基因體中的重複序列和結構變異。PacBio技術除提供更長的讀取長度，其環形一致性定序（Circular Consensus Sequencing）產生的HiFi資料正確性遠大於99%，可準確地偵測各種單核苷酸至結構上的基因變異。牛津奈米孔技術則是一種無需預先放大DNA或RNA的直接定序平台，其超長（Ultra-long）定序能解決高度重複區序列組裝的困難。它能即時定序並快速提供數據，對於迅速應對環境變化和臨床診斷具有重大意義。這兩平台能助益單倍體相位分型（Haplotype Phasing）或基因體中難以定序的區域，並能偵測特定鹼基上的修飾變化，研究表觀基因的調控。

從桑格到Illumina，再到PacBio和牛津奈米孔，定序技術的每一次演進都為科學家提供了更強大的工具，促進了生物科學和醫學的飛速發展（圖三）。

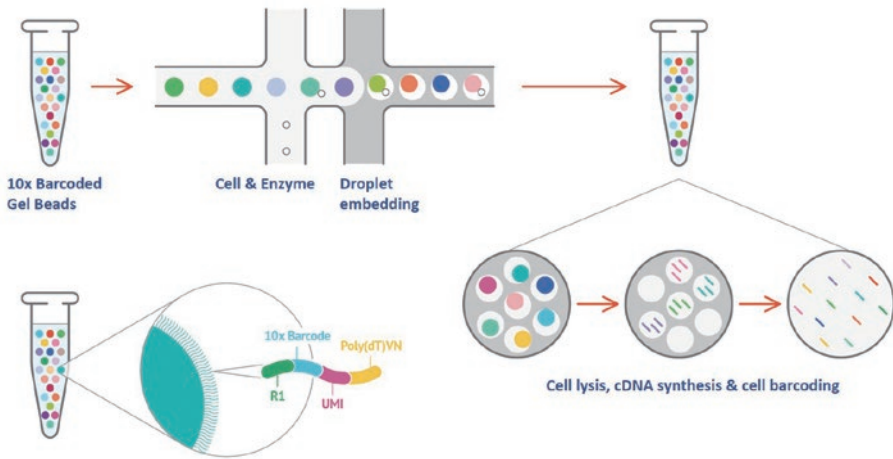


圖三、定序技術的世代演進。圖中顯示了目前主要平台的原理；從上到下依次為桑格（Sanger）、Illumina（中）、PacBio（左下）以及Oxford Nanopore定序（右下）。

先進基因體技術：單細胞定序和空間轉錄體定序技術

在基因體科學的最前沿，單細胞定序和空間轉錄體定序技術正開啟新的研究領域。這些突破性技術使我們能夠以前所未有的精確度和細節來探索生命的奧秘，用新的視角來理解生物體內的複雜互動。

單細胞定序：在於分析單一細胞的基因體或轉錄體，從而獲得更精確的生物學資訊。傳統的定序技術是在數百萬個細胞的平均數據上進行分析，而單細胞定序則能揭示細胞間的微妙差異。這對於瞭解腫瘤異質性、發育生物學以及神經科學等領域至關重要，利於探索癌症的發展機制、神經系統的功能以及細胞分化命運決定等問題（圖



(Reference: <https://www.10xgenomics.com/platforms/chromium>)

圖四、10x Genomics單細胞RNA轉錄體定序。使用油滴的凝膠和細胞條碼技術。(上圖)將分離的單細胞依序與攜帶不同細胞條碼的凝膠珠子包埋封裝，產出的cDNA文庫片段包含帶有細胞條碼和獨特分子指標(UMI)(下圖)，從而對單細胞的轉錄體進行分析。

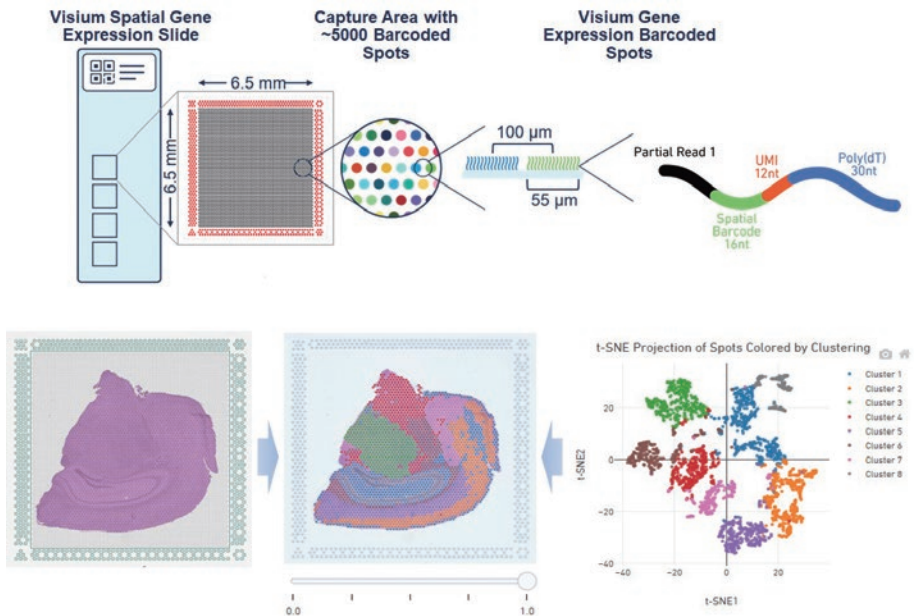
四)。

空間轉錄體定序：這技術將研究推向了一個新的層次，允許我們在保持原有空間結構的前提下，分析特定組織區域的基因表達模式。對於發現細胞在特定組織環境下是如何相互作用和生物反應訊號非常重要，有助揭示各種疾病的發病機轉和治療目標(圖五)。

單細胞和空間定序技術的結合，正顛覆我們對生命科學的理解，尤其是在精準醫學和個體化治療方面的應用潛力巨大。隨著這些技術的不斷完善，我們有望更深入地探索和理解生命的多樣性以及細胞間的複雜交互作用。

基因體定序與應用：以珍貴的牛樟芝為例

牛樟芝是一種珍貴的藥用真菌。這種源自台灣的特有多孔菌，長久以來被視為對抗癌症、高血壓和宿醉的天然療方，其市場價值在台灣每年超過一億美元。因生長侷限在樟樹（尤其是台灣原生的牛樟樹）上，其珍稀程度常

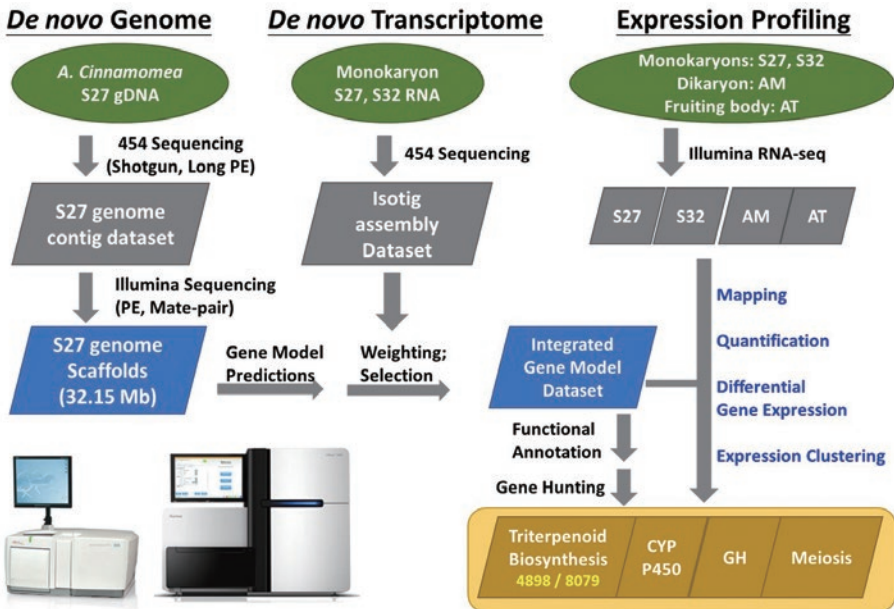


(Reference: <https://www.10xgenomics.com/platforms/visium>)

圖五、10× Genomic空間轉錄體分析。10× Visium玻片上印有空間條碼（類似郵遞區號）和分子標記，供組織貼片和cDNA合成和空間標定建庫定序（上圖）。藉由將染色的小鼠大腦組織切片（左）與RNA定序聚類分析（右）對齊疊合，可以將基因表達聚類映射到組織上（中），在特定區域的組織圖上展示基因表達的空間分布（下圖）。

引起樟樹的盜伐。

利用各式的次世代定序和分析，研究牛樟芝的基因體和轉錄體，我們完成了基因體草圖的繪製和基因註釋工作，獲得全世界第一個牛樟芝基因體草圖，總長32.15 Mb 包含了9,254個基因。透過基因體本體豐富度和途徑分析，



(Reference: Lu et al., 2014)

圖六、牛樟芝基因體定序研究基生物資訊流程圖。運用Illumina及Roche 454定序平台獲得牛樟芝DNA的高通量長短序列資料，並用生物資訊工具進行基因體組裝（左）；將牛樟芝RNA進行轉錄體定序（中），輔助基因模型預測和分析轉錄體的基因表現定量，以利尋找特定基因，研究次級代謝物的生合成和有性生殖的調控機制（右）。



(Reference: 呂美擘、張瑞仁、林玉儒, 2017)

圖七、藥用台灣牛樟芝基因體組序與功能特性分析。此圖展示了一個典型基因體研究計畫的步驟示例(圖示1~9, 左上至右下): 高品質的DNA和RNA製備及QC(1-3)、選擇適合基因體組裝的定序平台(4-5)、基因模型預測和轉錄體分析(6-7)、尋找重要性狀的顯著基因和家族(8), 以及通過合成生物學解決相關問題和/或需求(9)。

我們對牛樟芝的有性生殖遺傳調控機制及其獨特的藥用成分一次級代謝物，如倍半萜、三萜、麥角甾、牛樟芝醇和牛樟芝素的生物合成途徑有了更深入的理解。

此研究中我們還鑑定出在菌絲體和子實體階段表達差異的基因，以及參與麥芽糖途徑、萜類途徑、細胞色素P450和多酮合成酶的242種蛋白質。這些蛋白質可能對牛樟芝中那些具有藥用價值的次生代謝物的生成有關鍵作用。有趣的是，我們還發現了一些特定的基因在不同組織中的表達豐富性，例如在子實體中轉化蘭石脂為麥角甾三 的14- α 去甲基酶（CYP51F1），在菌絲體中參與牛樟芝醇合成的輔酶Q（COQ），以及在子實體中生產牛樟芝素的多酮合成酶（圖六）。

這些發現不僅為我們提供了對牛樟芝生物合成途徑的深入洞見，而且還為未來開發提高這些有益代謝物產量的策略奠定了基礎。利用合成生物學，科學家能將牛樟芝生物合成途徑的基因導入細胞工廠（如酵母菌），製造並量產有藥用價值的次生代謝物（圖七）。

為推廣此研究成果的相關科普知識，我們製作了微電影及一本專刊，以視聽的平台和戲劇的方式，讓人一窺基因體研究人員的歷程（呂美曄、張瑞仁、林玉儒，2016）。

參考文獻

1. Giani, Alice Maria, et al. "Long walk to genomics: History and current approaches to genome sequencing and assembly." *Computational and Structural Biotechnology Journal* 18 (2020): 9-19.
2. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. *Genomics*. 2016 Jan; 107(1):1-8.
3. Lu, M. Y. J., Fan, W. L., Wang, W. F., Chen, T., Tang, Y. C., Chu, F. H., ... & Li, W. H. (2014). Genomic and transcriptomic analyses of the medicinal fungus *Antrodia cinnamomea* for its metabolite biosynthesis and sexual development. *PNAS*, 111(44), E4743-E4752.
4. Warburton, P. E., & Sebra, R. P. (2023). Long-Read DNA Sequencing: Recent Advances and Remaining Challenges. *Annual Review of Genomics and Human Genetics*, 24.
5. 呂美曄、張瑞仁、林玉儒（2017年07月）。生技創業時代與女男平權的優勢 *Advantages of gender equality in the era of biotech entrepreneurship* (ISBN : 978-986-05-3058-2)。台北。計畫內容：性別友善環境之研究、性別與科學學習之研究。經費補助：行政院科技部。自然司編號：105WIA0150685。出版者：中央研究院生物多樣性研究中心。

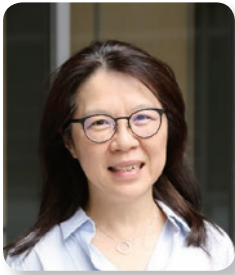
6. 呂美曄、張瑞仁、林玉儒（2016年10月）。【牛樟芝身世篇 # 我的研究日記】微電影（上）| 女性科學家與她的發源地。<https://www.youtube.com/@user-pm9nt6wr6x>



後記

基因體學的驚人進步，對當代科學造成了巨大且深遠的影響。從高通量定序技術的起源到其在單細胞和空間定序領域的突破，每一步都體現了對生命更深層次的理解。定序技術的演進和多樣化的應用使我們能夠更精確地瞭解生物組織的複雜性。隨著科技的不斷發展，我們相信未來將揭示更多關於生命的秘密。這不僅將擴展科學的疆界，還將為醫療、農業和環境保護等領域帶來革命性的變化。展望未來，這些發現將繼續推動科學的前沿，為探索生命奧秘提供新的工具和思路。

作者簡介



呂美曄

中央研究院生物多樣性研究中心研究技師、全院性新世代基因體定序核心實驗室經理及核心設施計畫主持人。專長為高通量定序技術服務及新穎技術研發和應用。研究範疇為功能基因體學、光合作用轉錄體、細菌比較基因體學、單細胞及空間轉錄體學，並推廣新穎定序相關課程與工作坊。

https://www.biodiv.tw/zh_pi-Meiyeh_Jade_Lu

<http://ngs.biodiv.tw/NGSCore/about-ngs-core/>