



探索未知基因序列 —cDNA 末端快速擴增技術

郭裔培、劉清碩、楊順德

水產試驗所淡水養殖研究中心

前言

基因序列是分子生物技術重要的參考依據，可用於基因功能預測、基因表現分析和分子演化等應用，海洋生態系具有高度多樣性，根據世界海洋物種目錄 (world register of marine species) 統計資料，全球 2023 年已被確認的海洋生物高達 24 萬餘種，研究人員常會遇到目標研究物種的基因序列尚未發表的問題，本文介紹利用 cDNA 末端快速擴增 (rapid amplification of cDNA ends, RACE) 進行基因序列定序，以獲得未知的目標基因序列。

真核生物的 mRNA 轉錄

一、5' 端加帽 (capping)

以 DNA 為模板在細胞核內開始進行轉錄時，剛合成出來的 RNA 5' 端會很快地接上 7-甲基鳥苷 (7-methylguanosine, m⁷G)，保護 mRNA 前體不被核酸外切酶分解，並作為後續轉譯反應時的核糖體結合位點 (圖 1A)。

二、3' 端加尾 (tailing)

當 mRNA 前體轉錄出多腺苷酸化信號序列 (AAUAAA)，會被多聚腺苷酸特異性因子 (cleavage and polyadenylation specificity

factor, CPSF) 辨識接合，誘導切割刺激因子 (cleavage stimulatory factor, CstF) 與多腺苷酸化信號序列下游的富含尿嘧啶或是尿嘧啶/鳥嘌呤序列 (U/GU rich region) 結合，並從上述兩個接合位點間切割，將 mRNA 前體從 RNA 聚合酶分離 (圖 1B)。後續 RNA 聚合酶持續轉錄的序列，因為缺少 5' 帽端會很快被分解；釋出的 mRNA 前體則透過多聚腺苷酸聚合酶，於 3' 末端連接多聚腺苷酸尾 (poly-A tail, 圖 1C)，多聚腺苷酸尾具有保護 mRNA 和協助 mRNA 離開細胞核的功能。

三、mRNA 剪接 (mRNA splicing)

依據 DNA 模板轉錄出的 mRNA 前體含有內顯子 (intron) 和外顯子 (exon) 交互穿插，當 mRNA 前體離開 RNA 聚合酶後，內顯子會被剪接體 (spliceosome) 切除，保留的外顯子會拼接為一條成熟的 mRNA (圖 1D)，進行後續的蛋白質轉譯。此外，同一個 DNA 模板轉錄出的 mRNA 前體，剪切位點並非完全固定，透過不同的選擇性剪接 (alternative splicing)，可產生多種序列相似卻又不同的 mRNA 異構物 (isoform)，後續轉譯出不同功能的蛋白質。以圖 2 為例，mRNA 前體包含外顯子 1、2、3、4 和 5，經由 mRNA 選擇性剪接，可表現 2 種 mRNA 異構物，分別由外顯子 1-3-4-5 和外顯子 2-3-4-5 構成。

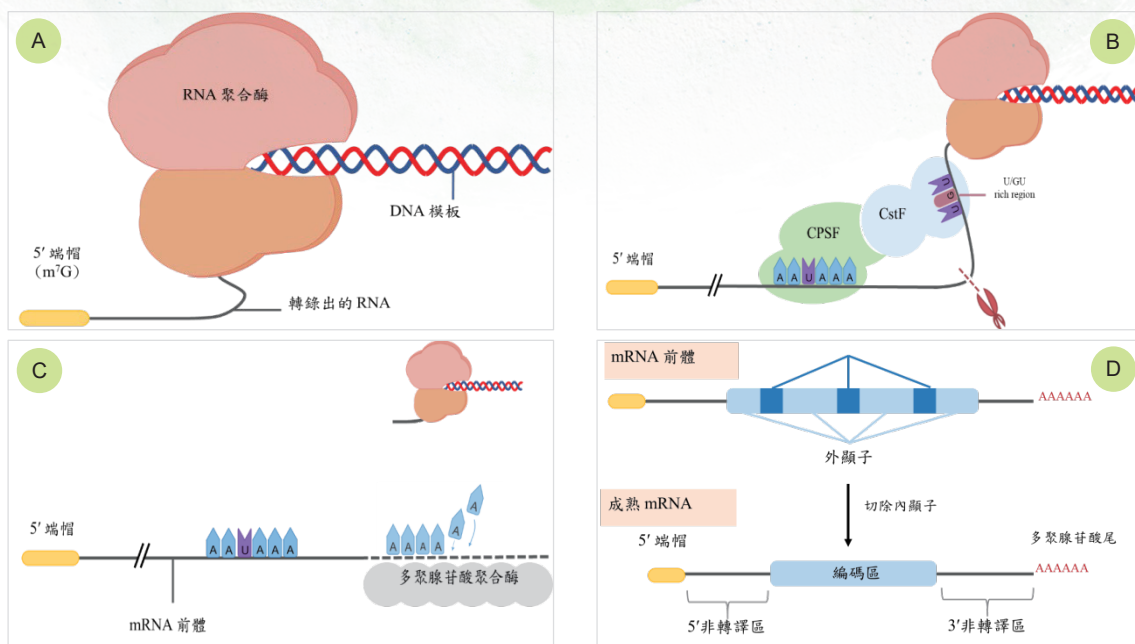


圖 1 真核生物的 mRNA 轉錄，包含 5' 加帽 (A)、mRNA 前體與 RNA 聚合酶分離 (B)、3' 端加尾 (C) 和 mRNA 剪接 (D)

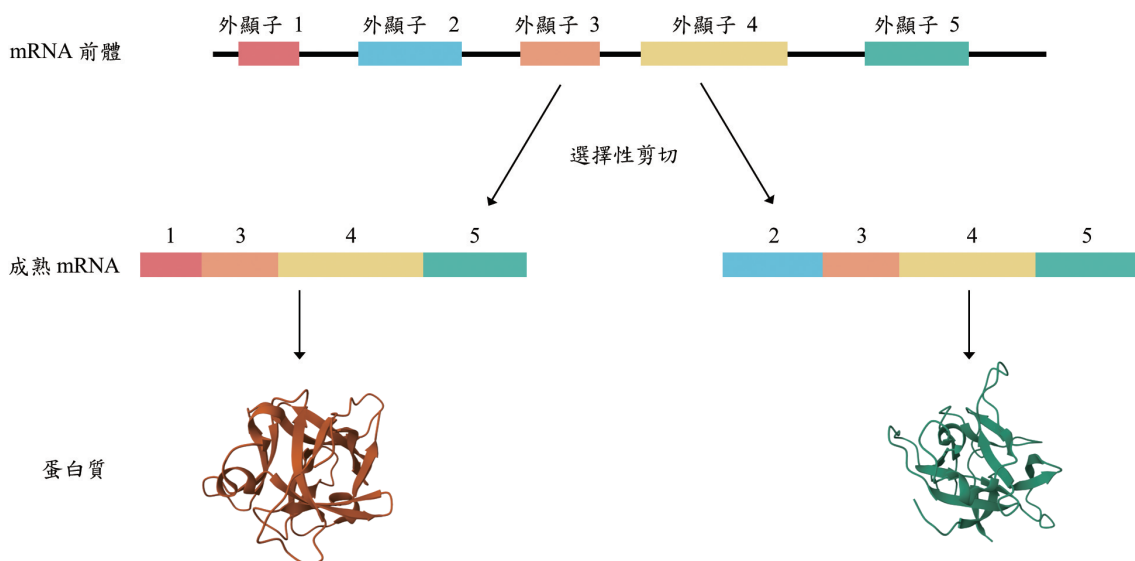


圖 2 mRNA 選擇性剪接

同一條 mRNA 前體透過選擇性剪切，可產生序列不同的 mRNA 異構物，增加後續轉譯蛋白質的多樣性，使得同樣的基因能在不同的生理狀況下，表達合適的功能性蛋白，增加生物的適應性

退化性引子 (degenerate primer)

在進行 cDNA 末端快速擴增前，必須先

取得目標基因序列上的其中一段片段，退化性引子是具有兼併鹼基 (在一個鹼基位置上有多種可能鹼基) 的引子，利用演化上同一



個基因在不同物種間會有相似度較高的保守序列特性，設計符合目標物種的退化性引子。以圖 3 為例，星號代表序列比對後完全相同的鹼基位置，紅色框為序列保守區，針對保守區內有差異的位點設計符合各物種序列的兼併鹼基 (W、M、S 和 H)，再以聚合酶連鎖反應進行目標物種的 cDNA 片段擴增與定序，即可得到目標基因的部分序列。

cDNA 末端快速擴增原理

cDNA 末端快速擴增能從部分已知的 mRNA 片段，分別往 5' 和 3' 端進行未知序列的擴增，完成整段 mRNA 的定序。(1)萃取目標物種的總 RNA 或是 mRNA 後，以連接酶將 RNA oligo 的 3' OH 基和 mRNA 的 5' 磷酸接合 (圖 4A)。(2)利用 mRNA 具有多聚腺苷酸尾的特性，以對應的 oligo dT 引子，將 mRNA 連同 5' 端接合的 RNA oligo 反轉錄為 RACE-ready cDNA，作為後續未知序列擴增模板 (圖 4B)。(3)未知序列擴增：從部分已知的 mRNA 片段設計正向 (forward) 和反向 (reverse) 引子，引子設計要盡可能地分別靠

近 3' 和 5' 末端，以降低聚合酶連鎖反應產物的長度；3' 端擴增：已知序列設計的正向引子和 3' oligo dT 反向引子進行聚合酶連鎖反應 (圖 4C)；5' 端擴增：5' oligo 正向引子和從已知序列設計的反向引子進行聚合酶連鎖反應 (圖 4D)。(4)選殖：聚合酶連鎖反應產物以電泳確認是否為單一產物，如果引子專一性不足，可在擴增的序列內再設計專一性引子，將第一次的聚合酶連鎖反應產物進行巢式聚合酶反應再次擴增，巢式聚合酶引子設計時，比第一次聚合酶連鎖反應的長度小 100 bp 以上為佳，以便電泳時區分兩次聚合酶連鎖反應的產物。最後將電泳後的單一條帶回收純化，透過質體送入勝任細胞進行選殖，培養後挑選數個菌落分別進行擴培、質體純化與定序。

長片段基因定序策略

由於反轉錄酶無法完成 4 kbp 以上的 mRNA 反轉錄，當未知的基因片段過長，反轉錄時可將 oligo dT 引子改為隨機引子 (random primer)，反轉錄出多條隨機 cDNA，



圖 3 退化性引子設計 (紅色框為序列保守區；* 代表序列比對後完全相同的鹼基位置)

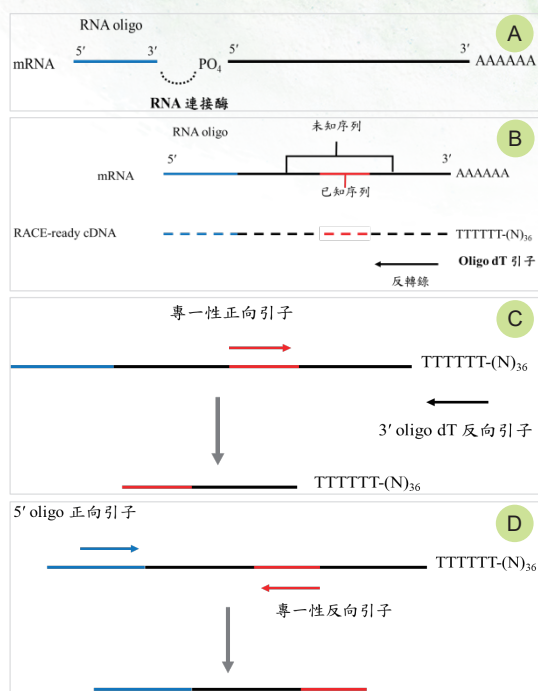


圖 4 cDNA 末端快速擴增技術原理

其中包含具有已知序列片段的 cDNA，再利用已知序列引子搭配隨機引子分別進行 5' 和 3' 定序，即可拼接出更長的已知序列，第二次 cDNA 末端快速擴增再利用上一輪新定

序出來的區域，重新設計新的專一性引子，反覆進行 cDNA 末端快速擴增，依序向兩端延伸直到完成整段 mRNA 定序 (圖 5)。

結語

cDNA 末端快速擴增技術是建立在 RNA 反轉錄的基礎上，因此在抽取 RNA 時，須考量到目標基因在不同組織的表現量差異，以及是否需要特定條件誘導基因表現，如免疫相關基因可採集脾臟、頭腎或白血球等免疫器官或細胞，並透過抗原誘導免疫基因表現，提高 cDNA 末端快速擴增成功率。相較於 DNA 序列同時包含外顯子和內顯子，mRNA 反轉錄的 cDNA 僅保留外顯子序列，可設計即時定量聚合酶連鎖反應引子，檢測目標基因表現量，此外，mRNA 作為後續蛋白質轉譯的模板，亦可應用於合成重組蛋白、製作抗體以及蛋白質結構和功能預測等多種應用。

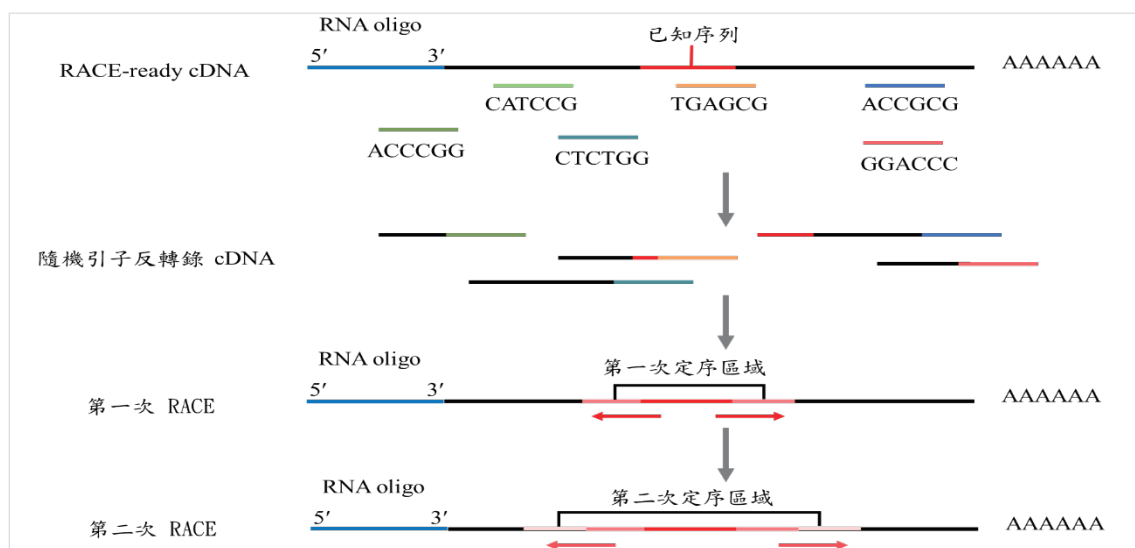


圖 5 利用隨機引子進行 cDNA 末端快速擴增