

# 植物的身分證-DNA條碼

張惠如<sup>1</sup>、鍾文全<sup>2</sup>

## 壹、前言

生物的遺傳物質存在於細胞中，主要以去氧核糖核酸(DNA)的方式存在，由於近30年以來，分子生物學發展迅速，以DNA的分析資料作為植物分類與演化研究，也普遍被許多學者採納。自2002年Tautz等學者提出DNA序列作為生物分類系統的方法(DNA-taxonomy)後，加拿大Guelph大學的動物學家Paul Hebert亦在2003年提出了DNA條碼(DNA Barcode)的概念。簡單來說，DNA條碼技術就是利用一些特定的短片段DNA序列(約400-800bp)作為物種快速且準確識別的方法，如同便利商店掃描儀讀取條型碼那樣。這項技術是利用生物體中某些遺傳保留性高的DNA片段進行物種親緣分析，藉以了解各物種間的分類情形並推論其演化的方向等多方面研究，使得這項技術在近年內成為生物分類學及生態學上的重要研究工具，本文針對DNA條碼技術使用與選擇標準、技術特色及植物應用範圍進行概要介紹，並討論DNA條碼存在問題及展望。

## 貳、DNA條碼技術的使用與選擇標準

DNA條碼可說是DNA分子標誌的一種，且其技術也是應用PCR分析方法，因此其工作內容與一般利用PCR進行分析的

分子標誌技術一樣，需進行樣品收集、樣品DNA萃取、設計與合成特定引子對、PCR、膠體電泳分析與序列數據分析(MEGA、PAUP與neighbor joining tree等)。DNA條碼的選擇標準具有3種特性，即：1.盡量短的片段；2.可使用通用的引子進行PCR增幅；3.具有足夠的變異性。在這樣的標準下，動物方面的研究經常使用粒線體細胞色素C氧化酶I(*COI*或*cox1*)作為物種內的鑑定標記。在植物方面葉綠體與核糖體為細胞質遺傳，包含許多變異區域且進化速度快於粒線體，因此，植物學者多使用核糖體內轉錄間隔區域(Internal Transcribed Spacers, ITS)與葉綠體的DNA序列，進行DNA條碼的開發。生物條碼聯盟中的植物研究組(Consortium for the Barcode of Life, CBOL)一開始建議使用葉綠體基因片段如：*matK*、*rpoC1*、*rpoB*、*accD*、*ndhj*和*YCF5*，後來研究發現*accD*在禾本科植物中缺失，而*ndhj*基因除在松屬植物中缺失外，在部分蘭花中亦會變短或功能喪失，至於*YCF5*基因則在蘚苔類植物中缺失。另外，研究結果也顯示核糖體ITS序列、質體*trnH-psbA*序列、質體*trnL-trnF*序列、23S rDNA序列，可被應用為通用的植物DNA條碼。在多個植物分類研究結果中發現，使用單一片段序列會有變異性

1 種苗改良繁殖場生物技術課 助理研究員

2 種苗改良繁殖場生物技術課 副研究員兼課長

不足的現象，因此學者進行了多種組合試驗，如：*rbcL+trnH-psbA*、*rpoC1+rpoB+matK*或*matK+trnH-psbA*等組合進行分析

後，所得到的鑑別結果較佳，建議採用多個片段的組合來做為植物 DNA 條碼。目前植物常用的DNA條碼引子序列如表一。

表一、植物 DNA 條碼片段之引子對序列

| 序列片段     | 引子名稱                           | F端/R端 | 引子序列(5'-3')                 |
|----------|--------------------------------|-------|-----------------------------|
|          | 2.1 <sup>[1]</sup>             | F     | CCTATCCATCTGGAAATCTTAG      |
|          | 2.1a <sup>[1]</sup>            | F     | ATCCATCTGGAAATCTTAGTTC      |
|          | 5 <sup>[1]</sup>               | R     | GTTCTAGCACAAAGAAAGTCG       |
|          | 3.2 <sup>[1]</sup>             | R     | CTTCCTCTGTAAAGAATTC         |
|          | matK-KEW <sup>[1]</sup>        | F     | AATATCCAAATACCAAATCC        |
| matK     | matK-KEW <sup>[1]</sup>        | R     | ACCCAGTCCATCTGGAAATCTTGGTTC |
|          | 390F <sup>[2]</sup>            | F     | CGATCTATTCAATCAATTTTC       |
|          | 1326R <sup>[2]</sup>           | R     | TCTAGCACACGAAAGTCGAAGT      |
|          | X <sup>[3]</sup>               | F     | TAATTTACGATCAATTCATTC       |
|          | F ( Equisetum ) <sup>[1]</sup> | F     | ATACCCCATTTTATTCATCC        |
|          | R ( Equisetum ) <sup>[1]</sup> | R     | GTACTTTTATGTTTACGAGC        |
|          | F ( Adiantum ) <sup>[1]</sup>  | F     | GATGTTGCAGTCTATTCATTC       |
|          | pkF1 <sup>[4]</sup>            | F     | TTTCTGATGAACAARTGGAA        |
|          | pkF3 <sup>[4]</sup>            | F     | CTACGATACTGGGTNAARGA        |
|          | pkF4 <sup>[4]</sup>            | F     | CCCTATTCTATTCAAYCCNGA       |
|          | pkF5 <sup>[4]</sup>            | F     | CCTCATTTTATTCAAYCCNGA       |
|          | pkF7 <sup>[4]</sup>            | F     | CTAATACCCTACCCNATHCA        |
|          | pkR1 <sup>[4]</sup>            | R     | CGTATCGTGCTTTTTRGYTT        |
| rpoC1    | 1 <sup>[1]</sup>               | F     | GTGGATACACTTCTTGATAATGG     |
|          | 2 <sup>[1]</sup>               | F     | GGCAAAGAGGGAAGATTTTCG       |
|          | 3 <sup>[1]</sup>               | R     | TGAGAAAACATAAGTAAACGGGC     |
|          | 4 <sup>[1]</sup>               | R     | CCATAAGCATATCTTGAGTTGG      |
|          | LP1 <sup>[1]</sup>             | F     | TATGAAACCAGAATGGATGG        |
|          | LP5 <sup>[1]</sup>             | R     | CAAGAAGCATATCTTGASTYGG      |
|          | ajfMossF <sup>[5]</sup>        | F     | GGCAAAGAAGGACGTTTTTCG       |
|          | ajfMossR <sup>[5]</sup>        | R     | CCAGAAGCATATCTTGACTTGG      |
| rpoB     | 1 <sup>[1]</sup>               | F     | AAGTGCATTGTTGGAAGTGG        |
|          | 2 <sup>[1]</sup>               | F     | ATGCAACGTCAAGCAGTTCC        |
|          | 3 <sup>[1]</sup>               | R     | CCGTATGTGAAAAGAAGTATA       |
|          | 4 <sup>[1]</sup>               | R     | GATCCCAGCATCACAAATTCC       |
|          | LP1.1 <sup>[1]</sup>           | F     | TCTAATATGCARCGTCAAGG        |
|          | LP3 <sup>[1]</sup>             | R     | TTTACCCAAAYRAACATCHCC       |
|          | LP4.3 <sup>[1]</sup>           | R     | ATAATACCTTTATTWCCATG        |
|          | LP5.2 <sup>[1]</sup>           | R     | AAATAAGGCATATCTTGTCT        |
|          | ajfF1 <sup>[6]</sup>           | F     | TCTAATATGCAICGTCAAGC        |
|          | ajfR1 <sup>[6]</sup>           | R     | GAGGIGTTAITTTACCTAC         |
| rbcL-a   | rbcL-a F <sup>[7]</sup>        | F     | ATGTCACCACAAACAGAGACTAAAGC  |
|          | RbcLajf634R <sup>[6]</sup>     | R     | GAAACGGTCTCTCCAACGCAT       |
|          | rbcL-a R <sup>[7]</sup>        | R     | CTTCTGCTACAAATAAGAATCGATCTC |
| cox1     | 42F <sup>[4]</sup>             | F     | GGATCTTCTCCACTAACCACAA      |
|          | ajf699R <sup>[8]</sup>         | R     | CCGAAAGAGATGCTGGTATA        |
| 23S rDNA | p23SrV F1 <sup>[9]</sup>       | F     | GGACAGAAAGACCCTATGAAGCTT    |
|          | p23SrV R <sup>[9]</sup>        | R     | TCAGCCTGTTATCCCTAGAGTAAC    |
| trnL     | c <sup>[10]</sup>              | F     | CGAAATCGGTAGACGCTACG        |

| 序列片段 | 引子名稱                 | F端/R端 | 引子序列(5'-3')            |
|------|----------------------|-------|------------------------|
|      | d <sup>[10]</sup>    | R     | GGGGATAGAGGGACTTGAAC   |
|      | g <sup>[11]</sup>    | F     | GGGCAATCCTGAGCCAA      |
|      | h <sup>[11]</sup>    | R     | CCATTGAGTCTCTGCACCTATC |
| rbcL | lf <sup>[12]</sup>   | F     | ATGTCACCACAAACAGAAAC   |
|      | 724r <sup>[12]</sup> | R     | TCCGATGTACCTGCAGTAGC   |

### 參、DNA 條碼技術之特色

DNA 條碼技術之特色，依與傳統方法相較後的優點主要有：不受發育階段與植株完整性的影響、得到的資料易於保存與分析、判別誤差較低、可鑑別的物種範圍大，分別細述如下：

1. 不受發育階段的影響：同種生物的 DNA 序列在不同的生長階段，如種子、小苗、成熟植株、開花階段基本上是完全相同的，因此不需等到某個生長階段才能進行差異性判別。
2. 不受植株完整性的影響：傳統分類學上需要收集完整型態的資料，依據型態資料進行物種的判別，但 DNA 層次上進行檢測時，則僅需取植株某個部份的樣品，即可進行辨別。
3. 得到的資料易於保存與分析：透過電腦軟硬體協助下，可將各物種、個體的 DNA 分析資料進行儲存與比對，進而建立 DNA 條碼資料庫。此外，更可一次分析大量的樣品，較傳統辨別方法減少檢測時間。
4. 判別誤差較低：在生物演化過程中，由於外在環境因子的改變，單憑外部型態來界定親緣關係或分類可能有趨同演化造成的錯誤，而 DNA 條碼分析方法的運用不易受這樣的影響，將能解決許多物種型態、生物功能性上

的爭議，故減少了判別誤差的產生。

5. 可鑑定的物種種類範圍大：由於 DNA 條碼引子序列，在篩選時就以在物種間具有高保留性的序列進行設計(其所增幅的片段卻需具高變異性)，因此這些引子對可在多種物種間被使用。

### 肆、DNA 條碼在植物學中的應用

DNA 條碼技術在植物學方面的應用研究，主要有：分類學、植物多樣性、古生物學及譜系群聚生態學上，分別敘述如下：

1. 應用於分類學上：植物分類需要仰賴專業人士與完整的標本，進行樣品型態的區分而能將物種分類，當缺乏專業人士或是植株型態標本收集不完全(例如花、果實、種子不易取得)；或是新物種的出現，此時要以傳統分類方法快速準確地進行鑑別，在執行上就較為困難。如果利用 DNA 條碼技術，則可立即進行鑑定並提供初步分類的方向。另外，DNA 條碼技術也成為發現那些型態相似但實際上存在遺傳分化的隱存種的有效途徑，譬如巨藻、膜蕨(*Polyphlebium borbonicum*)等新種，為植物分類學貢獻許多。進一步，也可以將本技術應用於檢測植物產品的純度，如檢測茶包或草本茶產品之組成，看似否有未標示的成份出現；

以及偵測管制物種，如植物外來入侵物種和瀕臨危險的植物物種。

2. 在多樣性研究上：在植物多樣性評估中，DNA 條碼技術可以克服傳統型態分類的缺陷，可以在多種植物群落(如高原、熱帶雨林等)進行分析，透過分析結果去探討群落的變化，更具有快速、操作簡便的優點。
3. 在古生物學上：透過植物化石進行古世代植物群落是非常不容易的事情，因為植物化石發現不易，其完整性也不足。在乾燥寒冷的環境下，分子保存情況較好的情況下，可以利用 DNA 條碼技術協助分析古植物群落。
4. 在譜系群聚生態學( phylogenetic community ecology )上：應用 DNA 條碼技術(使用多組序列片段組合分析)可以快速獲得大量植物序列分析資料，建構群聚譜系樹及大樣區的生態資料，並結合群落生態學進行植物群落的譜系分析研究。而這方面的概念及相關的分析軟體，近來年逐步發展成熟，已成為快速發展的新興整合領域。例如：史密森自然史博物館採用的策略為由熱帶雨林豐富樹種進行評估；以巴拿馬 Barro Colorado Island (BCI) 50 公頃大樣區 296 樹種建立實驗流程，並且以植物 DNA 條碼建構群聚譜系樹，延伸 DNA 條碼在生態學研究上的運用。

## 伍、植物 DNA 條碼存在問題與展望

DNA 條碼技術操作具有簡便性與高效性，對於物種鑑定和物種演化研究上也

提供了很多有用的資料，加速了相關研究的發展，故有許多研究學者的支持。但同時也有不少學者抱持著存疑的態度，因為這項技術仍有缺點與無法克服的問題。許多存疑者認為這項技術對於分布地區廣泛的物種鑑定可能效果較差，原因是：1. 許多較晚分化出來的類群會被排除在局部區域分析之外；2. 由於不考慮地理變異，因此種內變異在地區性研究上可能被嚴重地低估。特別是在植物的研究上，因為植物基因組進化較慢，難以確定由哪一部份的 DNA 片段適合作為識別的 DNA 條碼。此外，在植物中譜系距離較遠且雜交的現象更加普遍，故需要從不同基因組中篩選標識物來確保鑑定結果的準確度，增加了篩選 DNA 條碼的難度。另外需要進一步探討的問題還有：如何確定種內跟種間的變異範圍、DNA 條碼的鑑定能力、該選擇多少數量的 DNA 條碼進行物種鑑別以及林奈二分法的分類方法所攜帶資訊可能消失等問題。

DNA 條碼技術的確為分類學、生物多樣性、物種鑑別等研究領域提供許多資訊、且能夠自動化與標準化，為相關研究領域帶來一個新思維外，也加快了研究的發展。希望隨著更多相關研究成果的發表與相關分析軟體的發展，讓 DNA 條碼技術更為完善，進而應用於藥用植物鑑定、食品成分鑑定、作物系統發育等研究與檢測方法。