

# 知識天地

## 水稻基因體研究新里程碑

邢禹依特聘研究員、吳正杰博士候選人（本院植物暨微生物研究所）

### 摘要

稻米是全球最重要作物之一；當人口持續增加且氣候變遷帶來影響時，如何提高稻米品質與產能、抵抗生物性及非生物性逆境成為全球植物學家及育種學家極為關心的課題。而稻米的近親—野生稻—富含多樣的遺傳背景以適應環境的各類特性，成為改良現有栽培稻的重要資源。因此，解讀野生稻的全基因體序列成為最基礎也最關鍵的因素之一。藉由比較栽培稻與其祖先野生稻基因體的異同，了解栽培稻馴化過程，將更有助於挖掘野生稻中重要的遺傳資源。

### 一、現代栽培水稻的近親與遠親

稻屬（*Oryza* genus）中共有24個物種，除了亞洲栽培稻（*O. sativa*）、非洲栽培稻（*O. glaberrima*）及其祖先（*O. rufipogon*、*O. nivara*和*O. barthii*）外，同為AA基因體型的「近親」野生稻：*O. glumaepatula*、*O. meridionalis*及*O. longistaminata*，他們在世界上的分布如圖1所示<sup>[1]</sup>。其他16個「遠親」

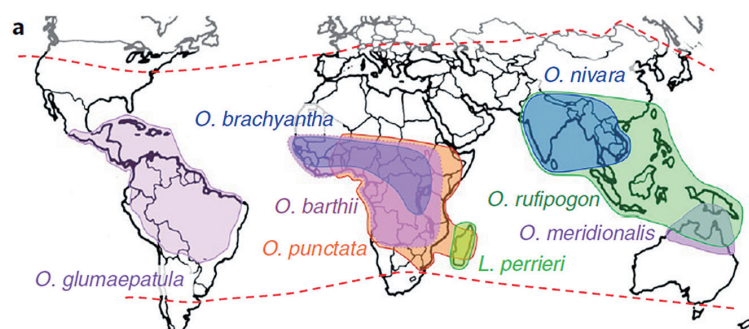


圖1：近親野生稻分布區域圖

野生稻，例如*O. punctata* (BB)、*O. officinalis* (CC)、*O. australiensis* (EE)、*O. brachyantha* (FF) 與*O. granulata* (GG) 分屬於差異較大的基因體型，尚有其他較複雜基因體型BBCC、CCDD、JJHH和JJKK等，外表型如圖2所示。野生稻外觀與栽培稻差異頗大，在野外甚至常被誤認為雜草，如亞洲野生稻（*O. rufipogon*）株高到達2公尺以上且穀粒有長芒，而*O. brachyantha*則矮於1公尺。然而，這些野生稻品種卻提供了重要的農業資源，這類仍待開發的重要農藝性狀將有潛力解決我們現今所面臨水稻生產中的許多問題，如：縮短生長期、提高產量及品質、抗病蟲害、耐旱及高鹽、耐高低溫等性狀。

稻米對於人類文明發展以及糧食安全的重要性無庸置疑。現代栽培稻可區分為亞洲栽培稻（*O. sativa*）以及非洲栽培稻（*O. glaberrima*），分別由其野生稻祖先獨立馴化而來：亞洲栽培稻於大約一萬年前在中國長江流域由其野生稻（*O. rufipogon*）馴化而來；非洲栽培稻則於約三千年前馴化自非洲野生稻（*O. barthii*）。野生稻與栽培稻在外表型態以及生理特徵上均有極大的不同。種植幾乎遍布全世界的亞洲栽培稻被視為人類最重要的主要糧食來源之一；然而，儘管栽培稻為主要的糧食作物、同時也被視作植物研究領域的主要模式系統，長期以來關於栽培稻的演化起源及其馴化過程仍然充滿爭議。

以現今的飲食習慣，亞洲栽培稻成為全世界近一半人口的主要糧食，提供了近五分之一的人類膳食熱量。由於全球人口在2050年將增加到30億，稻米育種家迫切地需要育成具高產量、高營養、高抵抗環境變化力的栽培稻品系。而栽培稻的近親「野生稻」遍布於不同的生物地理環境



圖2：栽培稻及其近、遠親野生稻外觀<sup>[7]</sup>。最左邊為亞洲栽培稻，他旁邊是非洲栽培稻。

(圖1)，從貧脊的乾旱地至深水區域甚至惡劣的生長條件都可以看到它們的蹤影；野生稻能抵抗多種生物性逆境及適應非生物性環境逆境，因此野生稻深深吸引了育種學者以及研究人員的目光，視為可用來改良現有栽培種的寶庫。

由野生稻導入重要功能的基因並非天方夜譚；例如在70年代早期，有一群水稻育種家開始注意野生稻*O. longistaminata*，此品種對於白葉枯病具廣效的對抗性，而由細菌病原菌*Xanthomonas campestris pv. Oryzae*所引起的水稻白葉枯病，對亞洲及非洲栽培稻皆造成嚴重的病害，進而也影響稻作產量<sup>[2]</sup>。學者從野生稻*O. longistaminata*中鑑定一群具有抵抗白葉枯病的基因，其中最著名的*Xa-21*就是在1990時由Khush所發現，它是第一個在*O. longistaminata*中被鑑定出的抗病基因<sup>[3]</sup>，進而廣泛的運用於水稻對抗白葉枯病的育種研究。

因應人口快速成長、耕種面積減少、環境天候變化急遽可能造成的糧食不足問題，為能更加有效率利用這些現代栽培稻遠親—野生稻，進一步了解野生稻全基因體DNA序列解讀以及基因體之間的差異及相似性是極為迫切的課題。

## 二、新世代定序技術橫空出世

野生稻基因體定序及註解功能基因的計畫得以完成，主要歸因於定序技術的進步。自1998年由來自10個國家的科學家合作國際水稻基因體定序計畫（International Rice Genome Sequencing Project, IRGSP），針對日本現代粳稻品種「日本晴」（Nipponbare）進行全基因體DNA序列定序及註解工作，日本晴水稻基因體全長約3億8千萬鹼基對（bp），大小約為人類基因體的八分之一。首先，利用細菌人工染色體（Bacterial artificial chromosome, BAC）建立之大片段選殖基因庫，依其製作詳實的基因圖譜（physical maps），再逐步定序轉殖載體（cloning vectors），定序深度約10倍，最後利用測序片段重疊區域將片段接起。其中，臺灣負責水稻第5條染色體的定序工作，建構全長近3千萬個鹼基對，定序及基因註解工作主要由植微所負責，達成99%以上的覆蓋率（筆者有幸參成為小小螺絲釘，亦於有榮焉）。國際水稻基因體定序計畫於2004年底達成目標。前後歷時7年，成功解開95%的序列，連最困難的高度重複序列區域中異染色質與染色體中節也被部分定序。整個團隊於2005年將成果發表於Nature期刊<sup>[4]</sup>。當時的定序技術主要原理仍為傳統Sanger定序法，例如Applied Biosystems™ 3700，其總讀序數量少，但優勢為讀序片段長（讀序片段約900bp）且讀序品質、正確率高，定序後進行組裝可藉由前後片段間重疊區域進而組成完成大片段基因體。

而後，次世代定序技術Next Generation Sequencing（NGS）橫空出世，以大量平行定序（massively parallel sequencing）的概念建構在第一代定序基礎上開發出之新興技術<sup>[5]</sup>，藉由大量序列片段達到同時高速且高通量的核酸定序；同時，樣品不須經過質體複製，因此減少製備時間與人力，並降低複製過程中可能出現的鹼基錯誤。次世代定序平台主要有三：Illumina、ABI SOLiD與Roche 454。次世代定序產出的序列數據量幾乎為前一代的十萬倍，但讀序片段較短（100~500bp）；因此，利用次世代定序平台得到大量短片段序列欲組裝成大片段基因體時，面臨到電腦運算效能不足以應付如此大量資料的窘境。同時，基因體內序列重複性區域常較讀序片段長，也面臨到組裝時短片段讀序無法判定位置的困境。雖然自次世代定序技術成熟後，各種動植物、微生物基因體文獻如雨後春筍般發表，但都面臨相似的問題。所幸，第三代定序技術平台持續軍備競賽，例如Oxford Nanopore、Pacific Biosciences，在保留高通量的優勢下延長讀序長度，因此，可預期在不久的將來，基因體測序將不再是科學研究的瓶頸而是起點。

## 三、水稻基因體研究里程碑

在栽培稻日本晴之全基因體DNA序列解讀完成之後，Rod Wing博士隨即號召許多原本國際水稻基因體定序計畫研究成員及相關研究人員，合作進行比較稻屬物種基因體計畫（The Oryza Map Alignment Project, OMAP）<sup>[6]</sup>，臺灣也參與此計畫進行*O. nivara*的部分。雖囿於初代定序技術的耗時耗力且所費不貲，但依循著相仿的模式，利用細菌人工染色體（BAC）建立選殖基因庫，而後製作基因圖譜（physical maps），準備好所有的前置作業，靜待時機。所幸，次世代定序技術的成熟，隨後搭著新技術的順風車趁勢崛起，解讀的基因體序列包含6個野生稻（*O. rufipogon*（亞洲），*O. nivara*（亞洲），*O. barthii*（非洲），*O. glumaepatula*

(美洲), *O. meridionalis* (澳洲), *O. punctate* (非洲, BB基因體型))、外群 (outgroup) *Leersia perrieri* (來自馬達加斯加島) (圖1), 以及2個亞洲栽培稻 (N22及IR8), 品質達到堪與日本晴全基因體序列相比的參考基因體水準。

團隊並利用新一代定序平台Pacific Biosciences定序完成亞洲栽培稻N22及IR8。栽培稻Nagina 22 (N22) 具有耐旱、耐熱特性, 同時因其深根的性狀備受矚目, 成為淹水稻研究領域良好的素材。IR8 (又稱為「奇蹟稻」) 是第一個高產水稻品系, 由國際稻米研究中心 (International Rice Research Institute, IRRI) 育出。由於50年代的亞洲處於饑荒邊緣, 來自臺灣的水稻專家張德慈博士利用臺灣本土的半矮品系低腳烏尖 (Dee-geo-woo-gen), 育出以IR8為首的高產量IR系列品系, 進而引發了亞洲綠色革命 (Green Revolution), 並於60年代在菲律賓、越南、緬甸、柬埔寨、印度、馬來西亞等地減緩了糧食危機, 2016年正是IR8育成50周年慶。

完成基因體定序及組裝後, OMAP計畫將這9個基因體與先前已發表的4個基因體 (亞洲栽培稻 *O. sativa* *vg.* *Japonica* (粳稻) 及 *Indica* (秈稻), 非洲栽培稻 *O. glaberrima*、非洲遠緣野生稻 *O. brachyantha* (FF基因體型)) 利用相同的方式註解以減少誤差, 同時進行比較基因體分析, 解決了關於稻屬系統演化的爭議, 並支持秈稻 *Indica* 與粳稻 *Japonica* 的獨立起源一說。藉由研究成果, 推測出AA基因體型物種演化分支應少於兩百五十萬年; 此基因體型物種之間, 不同染色體複雜的基因滲入 (introgression) 歷史及基因移徙 (gene flow) 足跡仍可被追蹤; 共有 21,448個基因家族被辨識出來, 可區分為4個不同的年代, 其中最古老的基因家族們可被追溯至被子植物的共同祖先。

此外, 我們也發現數千個基因間的長鏈非編碼RNA (long intergenic noncoding RNA, lincRNA)。在外群 *L. perrieri* 中具有比其他稻屬更多的lincRNAs; 而兩栽培種 *O. sativa* 和 *O. glaberrima* 中, lincRNAs顯著少於其野生稻祖先。此lincRNAs有著種別特異性 (species specific), 可被視為新型態的基因。

大多數植物抗病基因編碼NLR (nucleotide-binding, leucine-rich repeat) 能直接或間接辨識病原。NLR數量在稻屬間是高度可變的, 研究成果鑑定了5,408個候選NLR基因, 範圍從 *O. brachyantha* 的237個到秈稻 *O. sativa* *vg.* *indica* 的535個。此外, 與野生稻祖先相比, 秈稻與粳稻的NLR增幅明顯得高出許多; 此結果與使用人為選拔方式所增加NLR多樣性的結果相一致。

野生稻基因體定序及註解功能基因的計畫完成, 並於今 (2018年) 年初發表於Nature Genetics期刊<sup>[1]</sup>, 成為當期封面故事, 研究成果亦公開與其他研究學者共享。

## 參考文獻

1. Stein JC, Yu Y, Copetti D, Zwickl DJ, Zhang L, Zhang C, Chougule K, Gao D, Iwata A, Goicoechea JL et al: Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nat Genet* 2018, 50(2):285-296.
2. Song W-Y, Wang G-L, Chen L-L, Kim H-S, Pi L-Y, Holsten T, Gardner J, Wang B, Zhai W-X, Zhu L-H: A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *science* 1995, 270(5243):1804-1806.
3. KHUSH GS, BACALANGCO E, Ogawa T: 18. A New Gene for Resistance to Bacterial Blight from *O. longistaminata*. 1990.
4. Sasaki T: The map-based sequence of the rice genome. *Nature* 2005, 436(7052):793.
5. Goodwin S, McPherson JD, McCombie WR: Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics* 2016, 17(6):333.
6. Wing RA, Ammiraju JS, Luo M, Kim H, Yu Y, Kudrna D, Goicoechea JL, Wang W, Nelson W, Rao K et al: The *oryza* map alignment project: the golden path to unlocking the genetic potential of wild rice species. *Plant molecular biology* 2005, 59(1):53-62.
7. Sanchez PL, Wing RA, Brar DS: The Wild Relative of Rice: Genomes and Genomics. 2013:9-25.