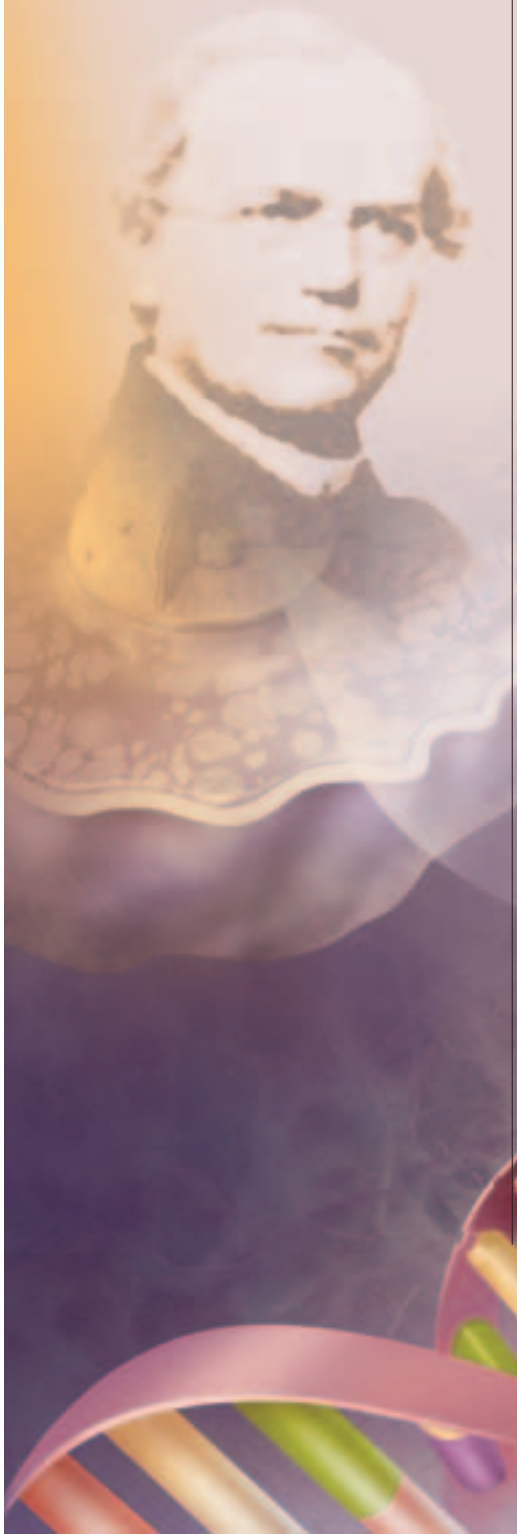


# 新興的 生物資訊學

這些年來，分子生物學的快速進展和基因組技術的研發，產生了許多生物相關資料，這些大量累積的資料，需要電腦協助儲存、組織、分析及搜尋。另一方面，資訊科技近年來突飛猛進，已臻成熟階段，新興的生物資訊學乃因應而生。

■ 趙坤茂



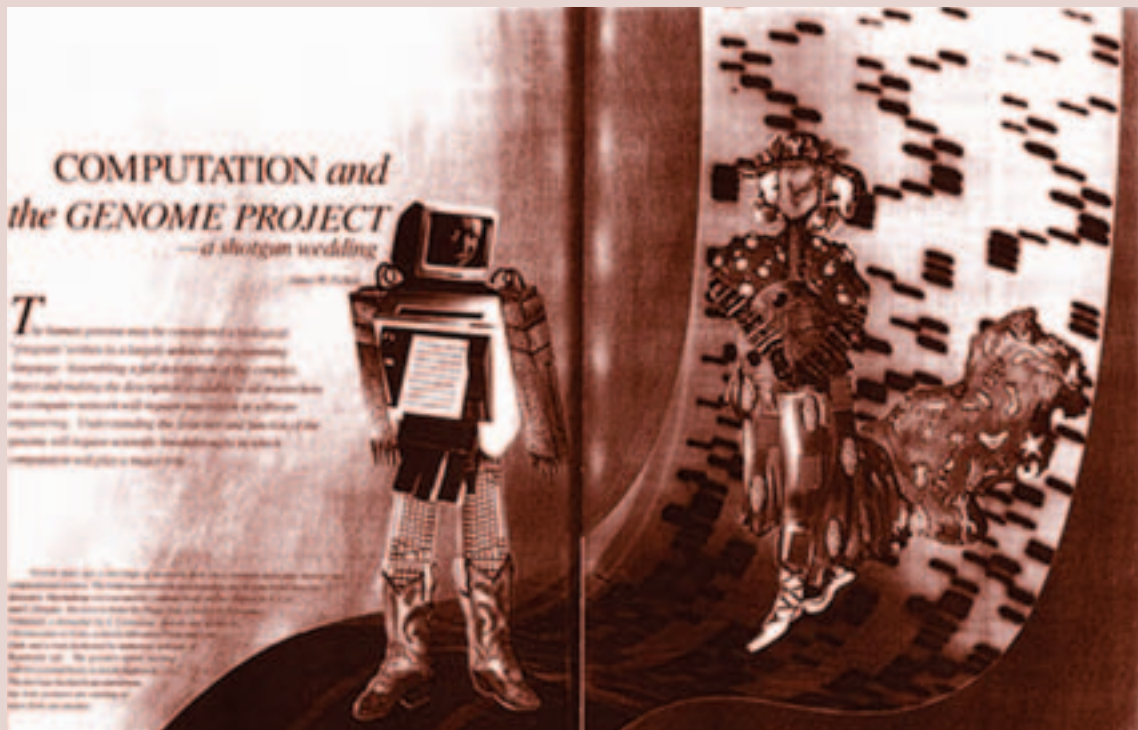
生物資訊學是近年來蓬勃發展的領域，它的研究是跨領域的，除了參與的研究者背景多元外，研究主題也隨著科技的演進時時在變，因此，每個人對它的詮釋也不盡相同。生物資訊學的英文名稱是 bioinformatics，在很多拼字檢查系統裡，還沒有被列為正確單字呢！

根據美國國家衛生研究院生物科技資訊中心的說法，生物資訊學是一個結合生物學、計算機科學及資訊科技所形成的新研究領域，最終的目標是發現新的生物認知，進而建立生物系統的大概念，以辨識生物學的各项準則。在基因組分析剛開始的時候，生物資訊學所關心的是如何以資料庫儲存各式各樣的生物資料，包括 DNA、RNA 及蛋白質序列。而發展這類型的資料庫，除了牽扯到設計議題外，也要考慮複雜的人機介面，以方便研究者查詢既存的資料，同時也可繳交或修改資料。

這些資料庫中的資訊，終究要結合成全面性的系統，釐清正常細胞各階段的表現，以利我們對疾病的發展能有更精確的了解。因此，目前最迫切的任務就是分析及解釋核酸序列、蛋白質序列、蛋白質結構、蛋白質家族及生化反應路徑等資訊。

總之，生物資訊學的研究結合了生物學、分子生物學、醫學、藥

**生物資訊學是一個結合生物學、計算機科學及資訊科技所形成的新研究領域，最終的目標是發現新的生物認知，進而建立生物系統的大概念，以辨識生物學的各项準則。**



分子生物學與計算機科學結合，產生了計算生物學。

[http://www.buid.org.tw/museum/s\\_star18.htm](http://www.buid.org.tw/museum/s_star18.htm)



遺傳學首創者孟德爾

學、資訊科學、數學、物理及化學等領域，終極目標是了解生物特性及生命本質。它的主要子領域有：大量生物資料的分析演算法及統計方法，各種生物序列、結構、功能及演化的分析與解釋，管理及使用各種型態資訊的軟體工具，生物醫學及藥物的資訊工具開發等。

生物資訊學可提供實驗設計者更宏觀的看法，

從以往個別基因的研究，邁向整個基因組的研究。有人形容早期的單一基因研究就像用釣竿

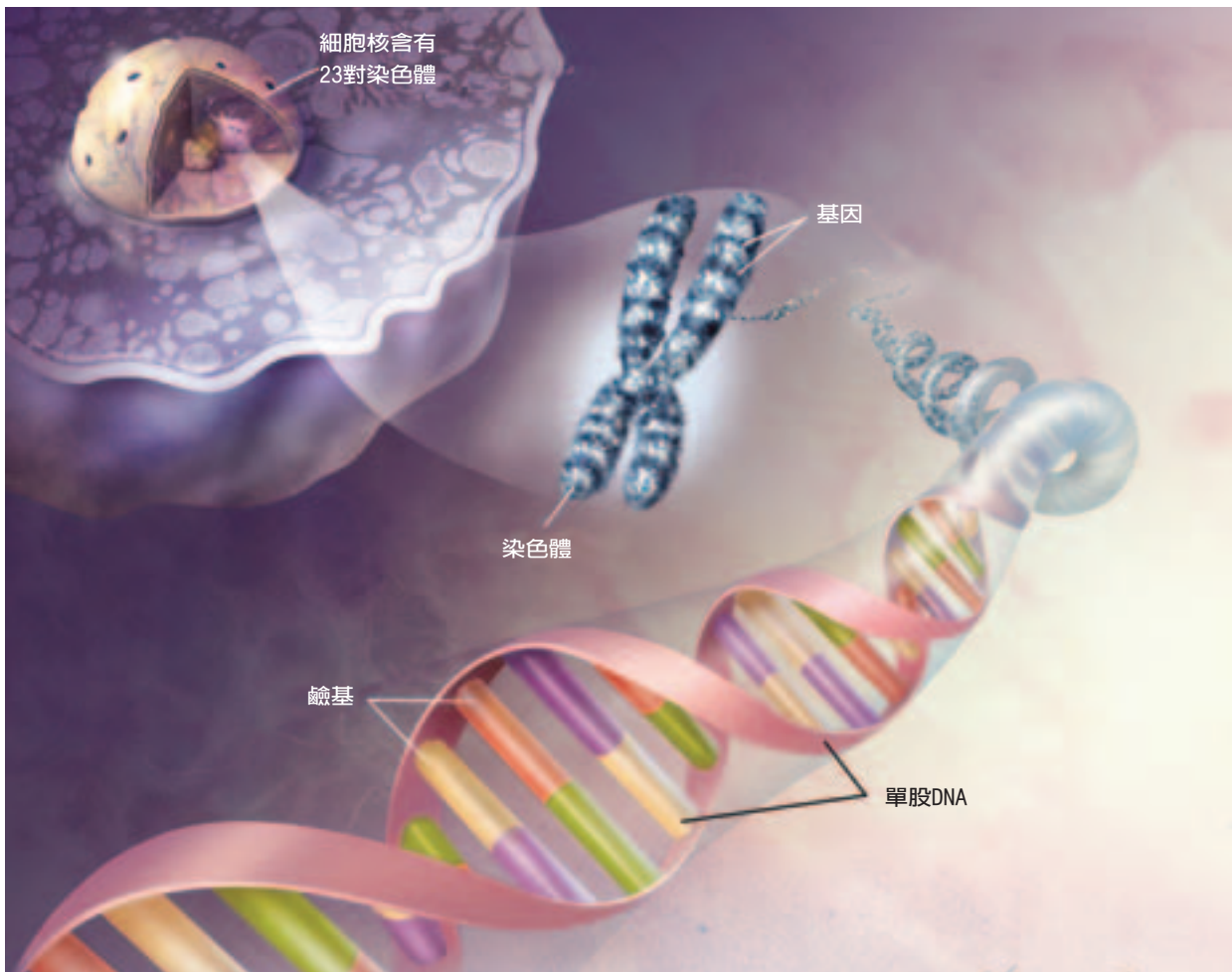
釣魚，一次一條，現在透過生物資訊學的研究，則有如用漁網撈魚般，效果自然強多了。

生物資訊學的核心領域主要是「基因組學」(genomics)及「蛋白體學」(proteomics)，有時也被戲稱為「基蛋學」(雞蛋學)，相關的研究領域還包括「計算生物學」及「系統生物學」等。

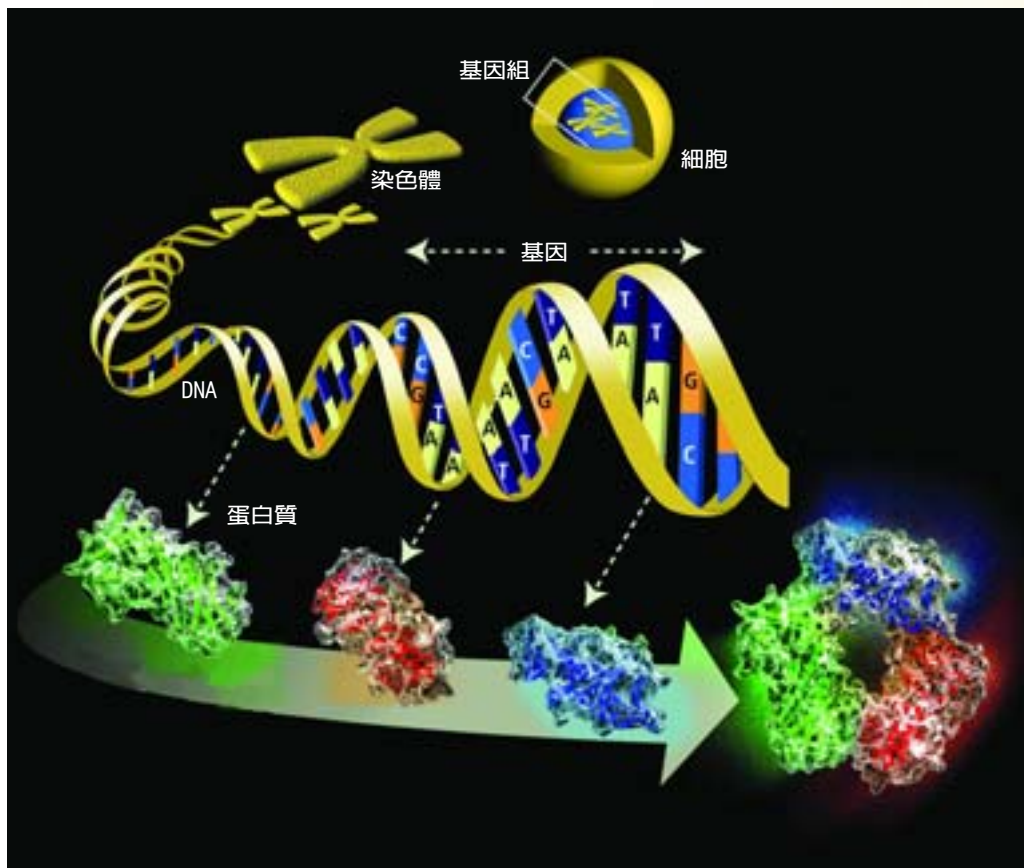
隨著生物科技及資訊科技的演進，生物資訊學的研究課題日新月異，常見的包括序列組合、序列分析、比較基因組學、生物資訊資料庫、基因認定、種族樹建構、蛋白質三維結構推測、微陣列晶片分析、反應路徑分析、分子演化、藥物設計、計算遺傳學等。

生物資訊學之所以能在短期內竄起，成為當代的顯學，最大的推動力應是來自於前陣子

[http://www.alzheimers.org/mediu/IMAGES/high/DNA\\_HIGH.JPG](http://www.alzheimers.org/mediu/IMAGES/high/DNA_HIGH.JPG)



人類細胞含有23對染色體，而染色體又由極長的DNA所構成。



[http://www.ornl.gov/sci/techresources/Human\\_Genome/graphics/slides/images/molecularmachine.jpg](http://www.ornl.gov/sci/techresources/Human_Genome/graphics/slides/images/molecularmachine.jpg)

基因中含有製造蛋白質的指令，而蛋白質或複合體使細胞發揮各種功能。

剛完成的「人類基因組解讀計畫」(Human Genome Project)，現在就讓我們簡短地回顧一下這個計畫。

一百多年前，遺傳學首創者孟德爾神父以栽培的豌豆，進行異種交配實驗，得到一個結論：植物繼承了來自父母雙方各一的遺傳性成分。這遺傳性的成分，就是我們現在所熟悉的基因。因為孟德爾，我們了解到遺傳基因是數位的，也就是說，各種基因不會混雜在一起，它們有就是有，沒有就是沒有。

很神奇的是，單一的基因也是數位的！怎麼說呢？DNA（去氧核糖核酸）已被證明是遺傳的基本物質，它是由四種鹼基（縮寫成A、G、C及T）組合而成的長鏈分子，而基因的訊息，就是由這些沿著DNA不同種鹼基的排列順序來傳遞。換句話說，基因的訊息，是由一種使用A、G、C及T四個字母的語言所寫成的。

所謂基因，就是指那些儲存蛋白質製造模具的DNA片段。

以人類為例，每個細胞有23對染色體，它們其實是由捲得很緊密的DNA所構成。有趣的是，若把細胞內的DNA拉直，大約有兩公尺長。若再把一個人體內所有的DNA拉直串接起來，它的長度可以由地球延伸到太陽，再返回地球，也可以自地球來回月亮8千次！不可思議吧！更令人驚嘆的是，光是一個小小細胞單套染色體的DNA，就已包含了32億個鹼基對，我們真能解讀這份建構生命奧秘的藍圖嗎？

在1953年，華生（James Watson）和克里克（Francis Crick）解開了DNA結構之謎，證實它是一個雙螺旋結構，也開啓了光輝燦爛的分子生物學時代。1988年美國國家衛生研究院生物科技資訊中心（NCBI）成立，以開放式的

# 專題報導 生物資訊

## 新興的生物資訊學

「人類基因組解讀計畫」在2003年完成，令人訝異的是，老天爺雖然賦予我們每個人獨特的生命之書，但你我之間的差異也只不過千分之二而已，真是四海之內皆兄弟啊！

資料庫「基因銀行」(GenBank)儲存來自全球各地的生物序列，頓時成爲研究生物資訊的重鎮。

號稱生命科學領域登月計畫的「人類基因組解讀計畫」，目標就是要解讀出人類基因組 32億個字符所寫成的生命之書，並找出其中的基因位置。這計畫從 1990年開始執行，由美國能源部及國家衛生研究院主導，有十幾個國家的實驗室組成全球定序團隊共同參與。

另一方面，私人公司 Celera 成立不久就造

**目前已知的遺傳疾病已達數千種，我們希望生物資訊學的進步，能協助生物科技及基因工程得以治療這方面的疾病，或至少能在診斷上有所助益。**

文分別刊登在《自然》及《科學》期刊，而整個精確的解讀計畫則在 2003年完成。

地球上的每一個生命體，都有屬於自己的有字天書。在生物科技與資訊科技共同合作

成一陣旋風，它以不到 1年的時間，就把果蠅的生命之書，約 1億 6,500萬個字符，解讀出來，其結果發表在 2000年 3月的《科學》期刊，作者共有兩百多位，分別來自 8個國家。

2001年 2月，全球定序團隊及 Celera 更把人類基因組初稿論

文下，第一部人類的「生命之書」草圖已經出爐了。令人訝異的是，老天爺雖然賦予我們每個人各自獨特的生命之書，但你我之間的差異也只不過千分之二而已，真是四海之內皆兄弟啊！

現在的生物科技，正處於如同哥倫布找尋新大陸的探險時代，會帶領我們在這世紀裡逐步探索每一部數位化的生命樂章。這樣的旅程，勢必充滿了刺激、突破、爭議及風險，沒有人確切知道我們究竟會航向何方！

目前已知的遺傳疾病已達數千種，我們希望生物資訊學的進步，能協助生物科技及基因工程得以治療這方面的疾病，或至少能在診斷上有所助益。這方面的研究會使我們更認識生命的本質，但不願見到它被應用在違反倫理的事，如改變 IQ 或發展致命武器等。專家指出，生物科技產業會繼資訊電子產業，在 20年內成爲最當紅的產業。隨著人類基因組圖譜的完成，生物科技產業的接班態勢已趨明顯。



<http://www.photo-power.net/photorevolution/pages/PP-13002.html>



翻自：張志強

早期的單一基因研究就像用釣竿釣魚，一次一條。現在透過生物資訊學的研究，有如用漁網撈魚般，效果強多了。

我們在20世紀裡發現了原子、解構了DNA以及發明了電腦，21世紀會是原子、DNA及電



<http://www.infogenetic.com/images/nature682.jpg>

西元2001年2月15日，由美國能源部及國家衛生研究院主導的人類基因組解讀計畫初稿論文發表於《自然》期刊。

腦等科技整合發揮的時代。也許不久的將來，我們就能設計最完美的後代，或是植入功能最強大的矽腦。到那時候，我們將處於通通是最完美，但卻也是最不完美的時代，因為月亮如果沒有了陰晴圓缺，那還值得我們歌詠嗎？從事科學研究者都必須審慎評估其研究在倫理道德上的衝擊。

下面列了三個生物資訊學的入門網站，值得有心的初學者去瞧一瞧：

人類基因組解讀計畫

[http://www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)

美國國家衛生研究院生物科技資訊中心

<http://www.ncbi.nlm.nih.gov>

國際計算生物學學會

<http://www.iscb.org/>

趙坤茂

台灣大學資訊工程學系

# 專題報導 生物資訊

## 新興的生物資訊學