



人類基因體

解碼後的商機

「基因」是由英文字gene翻譯而來，原指孟德爾遺傳定律中，用以產生性狀的因子，是學者約翰生（Wilhelm L. Johannsen）所創。在孟德爾的假設中，這因子是個顆粒，可在細胞內分散及組合。1910年代以後，基因就被想像成是染色體上一條條細的橫帶，

或是長鏈上的念珠。染色體是在細胞分裂期時才可觀察到的構造，每條染色體先複製再分裂為二，然後均勻地分散到新細胞中。1950年代以後，科學家們了解了染色體是由去氧核糖核酸（DNA）和蛋白質所組成的雙螺旋結構，而基因就是DNA分子上的一小段。



西元二〇〇三年，美國柯林頓總統宣布人類基因體內密碼均已解開，這項工作有何意義？伴隨而來的商機何在？以及由此衍生出的種種爭議，是本文探討的主題。

程樹德

因體研究所 (National Human Genome Research Institute) 主管此事。

雖然政治人物大多不懂尖端科學，但卻也想攬些「促進科學」的美名於其身，1989年，當時的布希總統 (George Bush) 在白宮頒發當年的國家科學及技術獎

1975年發明了分析及定序DNA分子上核苷酸序列的方法，到了1980年代，以基博特 (Walter Gilbert) 為首的一些學者，倡議以眾多科學家的力量，把人類23對染色體，總共大約30億對的核苷酸序列，都解讀出來，因為他們認為，能解讀製造人類特徵的基因藍圖，就能了解疾病及發育之類的過程。細胞內基因的總合稱為「基因體」或「基因組」(genome)。

此項建議談了幾年後，逐漸有人贊成，先是美國能源部想撥款資助，隨即被健康及福利部大肆炒作，甚至成立了國家人類基

時，不免要吹噓一下他及雷根總統主政期間的科學大事，諸如太空站、超導超級對撞機及基因體定序計畫，但他還不認得「genome」這個新創的字，老眼一花就將之唸成gnome (即地底小精靈之意)。在場的科學及技術界高級主管們，沒有一絲騷動，或許他們那時對這「人類地底精靈計畫」也不太清楚吧！但現在已由純科學研究，演變成了投資數百億美金的工業，讓我們來看看其來龍去脈。

解碼競賽

由美國官方所支持的人類基因體計畫 (Human Genome Project) 花費了約三十億美金，完成解讀三十億鹼基的工作，此計畫的領導人柯林斯 (Francis S. Collins) 對這約可塞滿兩百本電話簿A、T、G、C (腺嘌呤、胸腺嘧啶、鳥糞嘌呤、胞嘧啶) 四字母的序列，十分地著迷，他說：「這是一本製造你我的說明書，還有甚麼比這更吸引人呢？」

此計畫在美國設立了四個定序中心，另外英國劍橋的商格中心 (Sanger Center)、日本、法國、德國、大陸及台灣等定序中心也都協力合作，約有1,100位科學家投入工作，歷時近十年即將完成之際，不料卻在2000年4月，一家私人新創的小公司賽雷拉 (Celera Genomics) 卻搶得頭籌，宣告他們已完成人類基因組解碼的初稿。為何私人對人體基因有興趣呢？又何以賽雷拉公司能晚發先至？

原來官方的人類基因體計畫，是有條不紊地先從人類血球及精子細胞中，仔細分離23對染色體，將之切成片斷，先確定這些片斷的相對位置，然後再定各片斷內核苷酸的鹼基順序，正像是一頁一頁有序地撕下百科全書，讀完後，再一頁一頁地組裝回去。

而賽雷拉公司的策略大不同，它把百科全書拆散開來，同時找很多人來讀，再憑語句相重複之處予以連貫回來。在它的總部，同時有上百台基因定序儀工作，且運用超級電腦將數以千萬計的鹼基序列編纂起來。該公司在1998年才開張，但在2000年4月就公布已知基因染色體的位置草圖。

其實賽雷拉公司頗有取巧之嫌，它原先預定要取好幾個人的基因體，來回定序十次，才宣布「共識序列」，但它只拿一個人的基因體，來回念了三次。而且官方所得序列，都會立刻經網路輸入「基因銀行」(GenBank)予以儲存，並公之於世，賽雷拉公司就隨時取用官方成果，用以填補自己序列的空白處，且檢驗已知序列是否正確，其輸贏已定矣！

藥物基因體學的商機

官方既已投入三十億美元，何以私人又願意投下數十億美元做相同的事呢？因為不同的人，其細胞中的基因序列，可能有千分之一的不同。例如，某一相同的基因中如胰島素基因，某一位置你可能具有A這種核苷酸，而旁邊的人卻具有T這種核苷酸，此些微的差異稱為「單一核苷酸的多樣性」(single-nucleotide polymorphism, SNP)，也可能是麻煩所在。

這些微的不同，使有些人因而易患某病，而賽雷拉公司的創辦人溫特(Craig Venter)預測，很多藥物只對群體中百分之三十至五十的人有效，甚至有可能發生的是，某人的良藥是另一人的毒藥。溫特指出，瑞如林(Rezulin)是針對第二型糖尿病的藥物，但它在全球使用後，卻因對某些人的肝有毒性，導致60多人死亡。所以將來對症下藥時，應該設計簡單的基因檢驗方法，先篩選病人是否能接受這藥，以免反被毒害，這是一種商機，像賽雷拉這類的私人公司，



利用各人基因內鹼基序列的不同，來預測藥品的功效，只是新興藥物基因學的一部分目標。

不但想知道一般人的基因序列，更想為不同的個體定出其特異的順序，以期預測出此人將來可能易患何種病，警示他不要使用某一種對他有害的藥物，這種私人個別的服務，是要花大筆的金錢才能得到的！

利用各人基因內鹼基序列的不同，來預測藥品的功效，只是新興藥物基因學的一部分目標。用某段基因鹼基序列來預測其所轉譯出的蛋白質序列，並預測其蛋白質的空間形狀，設計出藥物來嵌入此蛋白質的立體構型中，使蛋白質功能被抑制或激發，工作極其複雜。因此一家公司無法獨立完成，需不同公司間相互合作，目前策略聯盟的公司已有不少。

例如，紐約市的輝瑞藥廠(Pfizer)與加州帕羅阿圖(Palo Alto)的因賽特基因體公司(Incyte Genomics)結盟，賓州費城的史凱必藥廠(SmithKline Beecham)與洛克維爾(Rockville)市的人類基因體科學公司(Human Genome Sciences)合作，而位於印地安那波里斯(Indianapolis)的禮來藥廠(Eli Lilly)則與位在麻州劍橋的千禧年藥品公司(Millennium Pharmaceutical)聯姻。有些商業分析家認為，個人化醫藥到了2005年，會有8億美元的市場。

專利爭議

如果台積電申請一項製程專利，只要它不侵犯別人的發明，沒有人反對。但在生物技術行業，若有人要以人類基因當專利來申請，一定會惹很多人反感，每個人都擁有的基因，怎麼能變成某一公司的獨有財產呢？

但這些生技公司說，如果我沒法擁有某些基因的專利，就不可能花大筆金錢，去研發有關某個基因之種種突變的偵檢試劑，因為若有別家公司也可以在稍後加入研發的競爭，但他們卻已省下了早期的研發投資，這樣，有誰願當第一個投入的傻子呢？

以上的論點雖不能說服很多人，但卻能說服美國專利及商標局，他們已發給因賽特公司(Incyte)五百多個基因專利。該公司另外也提出七千多個基因的專利申請，像人類基因體科學公司也得到一百多個

專利權，另外七千多個基因之專利也在申請中。

但很多科學家仍極力反對人類基因可變成專利的做法，其中一項理由是官方科學家所得到的序列，都報告給基因銀行了，而私人公司從銀行取走序列，稍加整理便申請專利，豈不太像強盜行徑了？例如2000年4月，光是一天之內，就有三萬五千人利用銀行的資料庫，其中當然有很多是屬私人公司的科學家。另一項理由是，私人公司還不明白某基因有何功能，便拿走專利，反而限制了此基因的研究，所以華盛頓大學的威爾生 (Richard K. Wilson) 教授警告大家：「如果資訊都鎖在私人資料庫裡，只有少數人花錢才能看到，會妨礙很多疾病的研究。」

反對的聲音雖大，既然專利及商標局已經頒發了那麼多基因專利，營利公司就可以用它來賺錢，2000年3月因賽特公司已推出基因體的網路商品，付錢上網的人可以訂購基因的序列資料，甚至實際已選殖到質體（可以獨立複製的DNA分子）的基因，目前付錢訂購的製藥大廠包括輝瑞、拜爾 (Bayer) 及禮來，而賽雷拉公司也在網路上販售各種基因的資訊，例如哪些基因與高血壓有關、哪些基因又跟大腸癌有關，大公司每年的訂購費，約在五百萬至一千五百萬美元間，而學術實驗室可少付些，每年在兩千元至一萬五千元之間。

基因測試

公司拿到基因專利後，便可開發基因檢驗試劑，目前已上市及開發中的，約有七百四十多種，可以查出受試者某特定基因是否有突變。但問題是，到目前為止，我們還不知道某突變基因與疾病發展過程的精確關係，例如，現在已有方法測定杭丁頓氏舞蹈症 (Huntington's Chorea) 基因的確切序列，但還無法

明確地預言某人何年將發病，及發病後的嚴重程度與進展速度。

另一項大問題是社會的歧視，如果測出一個人具有杭丁頓氏舞蹈症的基因，未來會發病，此人有可能會被雇主辭退，被保險公司退保。為了防止這類歧視，美國前總統柯林頓在2000年2月發布了一項行政

命令，禁止聯邦機構利用基因資料為依據，來聘用、擢升、或解僱人，私人公司是否會跟進，尚待觀察。

生物資訊學的商機

隨著人類基因體計畫的完成，其他許多生物的基因體也正在解讀中，業已完成的有果蠅、線蟲、酵母菌及近百種的古菌類 (Archaea)，正在進行中的則有小老鼠、黑猩猩、玉米、水稻、牛、羊等。由於各公私立機構的研究群不但有極強力的定序機器，也有好的軟體，因此他們定序能力驚人，例如因賽特公司每天能解讀出兩千萬個鹼基序列，

是台灣榮陽團隊幾年來成績的兩倍。

因此，全世界每天能解出的序列資訊何止億萬，如何處理是新興的「生物資訊學」範疇，有很多公司正爭取所衍生的商機，其中最有名的是總部在德國海德堡的獅子生科公司 (Lion Bioscience)、總部在馬里蘭州比賽大的最大資訊公司 (Informax)、總部在英國牛津的牛津分子群 (公司) (Oxford Molecular Group)、總部在俄亥俄州克里夫蘭的網路基因公司 (NetGenics)、總部在加州奧克蘭的雙扭公司 (Double Twist)、總部在以色列特拉維夫的電腦基因公司 (Compugen)。

最早處理DNA序列的機構，是前述的基因銀行，在1980年代初由美國能源部贊助成立，那裡有整屋子的技術員，將在各學術刊物所公布的基因序列鍵入電



隨著人類基因體計畫的完成，其他許多生物的基因體也正在解讀中。

腦，隨後基因銀行將資料轉到美國國家衛生研究院 (National Institutes of Health) 的生技資訊國家中心 (National Center for Biotechnology Information, NCBI) 讓研究人員可直接用網路將序列存入其電腦中。目前擁有的資訊已超過七十億個鹼基順序，私人公司也經常上網去擷取資訊。

生物資訊學是綜合DNA序列、基因在各組織內的表現時機、蛋白質的形狀及互動等資訊，以發展適合各種疾病的不同藥品。例如，骨質疏鬆症是因骨質不斷降低而鈣化，對中老年人造成嚴重的威脅。1993年，在費城史凱必公司 (SKB) 的研究人員從骨癌病人的破骨細胞 (osteoclast, 是骨組織中破壞骨質，而讓骨頭再生長、修補的細胞) 中分離出特殊的傳訊核酸 (mRNA)，並拿給人類基因體科學公司的科學家分析，他們先將互補DNA (cDNA也即是由基因表現後的mRNA反向轉錄成的DNA) 選殖及定序後，再將其序列與基因銀行已有的DNA序列比較，居然發現一個鈣蛋白 基因的序列與此相似，再回頭檢驗破骨細胞內這個蛋白質，果然其數量極高，似乎是破骨細胞消蝕骨質的利器。

幾週的功夫，就讓史凱必公司知道，若想治療骨質疏鬆症，設計一個藥物來抑制此鈣蛋白，應該是一個極佳的方向，這可省下了不少年的盲目實驗研究，若藥物能早一年上市，則所爭取到的額外專利權保護時效，即可能值五億美元呢！

所以生物資訊學的商機，在於提供一系列不同的服務，或為各大小公司提供資料庫，或設計程式以整合整個藥廠的資訊交流，或替公司預測蛋白質的形狀，或設計藥物，這種運用電腦來模擬細胞內基因互動及蛋白質結構的學問，也被叫做「矽晶內之生物學」(silicobiology)，以對照活體 (in vitro) 及試管內 (in vitro) 生物學。

蛋白體學的興起

雖然決定人類及各種動植物基因序列的工作很重要，但一旦完成，就成了過去的成就，並不足以再激發熱情。雖然還有不少剩餘工作有待處理，例如群體

內的各種變異、遺傳疾病的突變基因發生在何處、以及地球上尚有數千萬至數十億種生物基因體待解讀，但目前最熱門最炫的研究方向，已轉到「蛋白質體學」(proteomics) 了。

基因轉錄轉譯的過程，係基因轉錄出mRNA，需經過修剪和編輯以後，再運輸到細胞質中，由核糖體將之轉譯成蛋白質，蛋白質是執行各種生理生化功能的主要物質，也是細胞出現問題時的焦點，像是房子起火時的起火點。

整個基因體雖像是大樓藍圖，但建設某一房間時，只用藍圖中的極小部分，所以各組織內活躍的基因可能只有幾百個，這一小集合，可以稱為這特殊組織的轉錄體 (transcriptome)。例如，胰臟組織內的蘭氏小島細胞充滿了胰島素的mRNA，而在腦內的神經細胞內，則沒有這種mRNA。

了解了各特化細胞內的轉錄體，對純科學或應用研究都有價值。例如，1998年美國的國家癌症研究所 (National Cancer Institute) 推動「人類腫瘤基因索引」計畫，想把各種癌症內所出現的mRNA都登錄下來，由政府、學術界與一些藥廠共同合作。目前已找到各種癌症中共有五萬個活躍的基因，單在乳癌細胞中就有5,692個，而有277個基因在他種組織中是關閉的，這些基因所轉譯出的蛋白質，就是人類要以化學藥品來破壞乳癌細胞的目標 (蛋白質)。為了這個目的，美國國家癌症研究所也與食品藥物管理局 (Food and Drug Administration, FDA) 推動「組織蛋白質體創始計畫」(Tissue Proteomics Initiative)，準備花費幾百萬美元，找出癌細胞內特別出現的蛋白質。

其商機很明顯，如果藥廠營收每年要成長百分之十到二十，以滿足投資人的期望，那麼每年至少要有三到五個有希望的新藥進入臨床試驗的流程中，而目前每家大藥廠每年只有半個到一個半的新藥進入試驗。故要增加數倍的生產力，就只有仰賴轉錄體學找到可以用藥品來攻擊的目標蛋白質，而精通這領域的千禧年藥品公司答應幾年內交給拜爾藥廠二百二十五個已經初步試驗過可受藥品攻擊的目標蛋白質。

晶片的興起

目前台灣許多生技公司以發展基因晶片為主要產品，希望利用積體電路產業所精通的微蝕刻技術來製作晶片。

在非常光滑乾淨的一小片玻璃表面上，以微小探針針頭點上成千上萬個小點，每個小點內是一個生物體內特殊的互補DNA (cDNA)，這些單股的DNA固著在玻璃表面上後，有能力與互補的RNA序列形成很多氫鍵，於是RNA留在玻璃表面上，不被沖洗掉。

若我們分離同一生物某一種組織細胞內的mRNA，並裝上一個會發螢光的小分子於RNA的鏈上，然後均勻地散布在晶片表面，讓RNA與cDNA配對，洗掉沒配對的mRNA，再用雷射光來激發螢光分子，用機器閱讀光點並測量強度，就可知道這生物的某基因在這特殊組織細胞裡是否有表現了。

這種晶片可以探知某一基因在何種組織何種生理狀態會關閉或活化，對純化研究極有幫助。相似原理的晶片，也可以用做癌症及遺傳病的篩檢，目前在美國已有很多晶片公司及成品上市，其中最有名的是加州聖塔克拉拉的親和網路 (Affymetrix) 公司，它有一種產品可同時檢查六萬種人類mRNA的表現，另一產品可檢驗一千七百個人類與癌症有關的mRNA，價格極昂貴，要做個正確的檢驗，須花費新台幣六十萬元左右。



了解了各特化細胞內的轉錄體，對純科學或應用研究都有價值。

新一波商機所在，例如加州的大規模生物學公司 (Large Scale Biology in Vacaville) 及英國牛津的牛津糖科學公司 (Oxford Glycosciences) 均開發機器人來做二維膠蛋白質電泳的工作。所謂電泳，即帶電分子在電場裡會向異極游動，而分子跑的速度與它的電荷量成正比，與分子大小成反比，因此不同種的分子片斷就會分開來。如果電場在含水的膠體內，不同

分子片斷分開後，不容易再混合，就達到分離的目的了。蛋白質是質量很大的分子，用二維電泳，可以把成千的蛋白質分散在平面上。

機器人的長處，在於精確地不停地把膠上的蛋白質切割下來，分析甚至定其氨基酸順序，就知道其為何種基因所表現的。

除了鑑定蛋白質的表現及功能外，蛋白質間如何互動及蛋白質的空間結構均是科學家及生技界所注目的焦點。蛋白質結構的決定，以前是用X光結晶學的方法，分析其繞射圖，此

法需先將蛋白質製成結晶，要耗極大功夫及碰運氣，常需數年努力才能完成一個蛋白質。

目前，美國國家衛生院以兩千萬美元支助學術中心發展「結構基因體學」，許多生技公司也相繼投入，大家都想研發出半自動且可處理大量樣品的X光繞射儀。惟有掌握了蛋白質的精細三維空間，才能進行「理性的藥物設計」。

程樹德

陽明大學微生物及免疫學研究所

蛋白質體學的挑戰

既然蛋白質才是重點，處理極多蛋白質的技術是