



03

探求自身秘密  
的好時機

塞吉·布林（Sergey Brin）原本以為DNA檢測不過是供人消遣的娛樂。身為當紅網路搜尋引擎Google公司的共同創辦人，他讓整個世界擷取資訊的方式完全改觀，對不計其數的人來說，電腦網路已經成為情報蒐集站和聚會地點，這些都是以前無法想像的局面。

而他的妻子，安·沃吉斯基（Anne Wojcicki）與人合夥，共同創辦了基因檢測公司「23與我」。安邀請布林成為該公司頭幾位接受基因組檢測的客戶時，他答應了。布林甚至鼓勵其他親人參與此事，也興致勃勃想知道，他與親戚的哪些DNA片段是一樣的。

疾病風險預測的結果顯示，布林罹患某些疾病的風險比平均還低一些，罹患其他疾病的風險也許高一些，但對他來說，並沒有什麼結果是太讓人驚訝的。

然而，當「23與我」提供了另一項新的服務時，整件事就變得不太一樣了。布林的妻子鼓勵他接受一種新發展出來的*LRRK2*基因突變檢測。原因是布林的母親罹患了帕金森氏症，而近來的研究指出，在某些例子裡，這種遲發性的神經系統疾病是由於*LRRK2*基因的一種罕見突變所造成的。

布林仔細檢視了「23與我」以晶片所做的基因檢測的結果，發現他和母親都帶有這種*LRRK2*突變基因。到頭來，基因檢測不再只是一種娛樂。

這家公司提供給消費者的基因體檢，原先為布林評估出的風險因子大多都很平常，也沒有預測出其他特別高的風險。但後來發現了這個突變，布林頓時了解到自己在八十歲罹患帕金森氏症的風險，大約在74%左右。這結果不但十分戲劇化，也完全出乎布林的意料之外。布林現在還年輕，因此等到他年老、受到影響時，專家也許已經找出了治療的方式——但這當然誰也無法保證。

布林在自己的部落格上寫道：「這讓我處在很特殊的情境裡：我在還年輕

的時候，就事先知道自己非常可能會罹患什麼疾病。如今我有機會可以調整自己的生活，努力降低患病的風險。同時，在這個疾病開始影響我之前，我仍有充分的時間和機會，能夠多了解這個疾病，並支持相關研究。而且，先不論自己的健康，這麼做也可以幫助我的親戚和其他更多有類似處境的人。我很幸運能遇到這樣的事情。除非有人真的能發現傳說中的青春之泉，否則每個人到了年老的時候，或多或少會生一些病，我們只是無法事先知情罷了。我則比一般人更容易猜到可能會是什麼疾病在等著我——而我有足足數十年的時間可以作準備。」

歡迎來到基因組時代。那麼，你現在又想知道什麼呢？

## 尋找基因組裡的不定時炸彈

布林的故事讓人十分震驚，但他的情況絕非特例。我們每個人都有得到數十種疾病的風險，有可能因此生病，但也有可能不會，端看我們基因中風險因子的組合，以及是否暴露在會誘發疾病發作的環境中。**實際上，並沒有哪一種疾病完全與遺傳無關。**

也許你會抗議這種一概而論的說法。畢竟你如果被屋頂上掉下來的磚頭砸到，那總該與遺傳無關吧——不過，這種意外倒是有可能反映了丟磚頭的那個人的遺傳狀況。而你的基因，當然也一定會影響到你傷口復原的能力。

你自己的疾病風險因子，有一部分可以藉由詳細的家族病史推論出來，而你應該要好好利用這種「免費的基因檢測」來評估自身的風險。但並非所有人都知道自家的家族病史（特別是那些被領養的人），而且在小家庭為主流的時代，就算是最完整的家族病史，也很有可能無法概括所有的風險因子，畢竟常見疾病的遺傳模式往往很複雜，且難以預測。

我們已經討論過一些遺傳性很高的罕見疾病，由於這些疾病是由單一基因的突變所引起，所以較容易預測。而現今基因革命的觸角，已延伸到更常見的

疾病上，諸如糖尿病、常見癌症、心臟病、中風以及精神疾病，這些疾病的遺傳模式都不單純，但仍深受基因所影響。

若想要了解這些疾病，我們需要更複雜的遺傳模式。近來科學上的發現，讓我們能夠提出下列幾項強烈的論點：(1) 每一種疾病都與特別的基因和環境風險因子相關，目前科學家發掘出這些因子的進展極快。(2) 這些發現讓我們針對疾病的處理和預防，開啟了嶄新而利多的視野。(3) 你愈早得到這些訊息，愈能及早調整自身的生活方式並進行醫療檢查，藉此預防疾病的發生，或在剛發病、還能治療的階段，便及早發現。

為了幫助大家了解，我們就來談談成人型（第二型）糖尿病吧。第二型糖尿病顯然與基因有關，因為當一個人有手足罹患此病，那麼他（或她）得病的機率，便增加為常人的3倍。

然而，若我們觀察那些有超過一人深受第二型糖尿病所擾的家族，就會發現光靠顯性、隱性或性聯遺傳模式，都無法解釋這個疾病發生的情況——可見第二型糖尿病並不是由單一基因所決定的。相反的，這個疾病必定牽涉到數十種遺傳變因，而每個變因都會些微增加個體罹病的機率。遺傳學家將這種遺傳疾病稱為**多基因疾病**。

我們每個人都有一長串這類的遺傳變因，這些變因可能使我們罹患第二型糖尿病的風險高於常人、與常人無異、或低於常人。基因上具高風險的個人，就算是所處的環境誘因相對的少，還是有可能會得到第二型糖尿病。而基因上具中等風險的個人，只有在其他變因的配合下，例如體重過重、缺乏運動、以及飲食習慣不良，才會跨過門檻而生病。至於基因上屬於低風險群的個人，就算生活習慣並不健康，還是有可能逃過一劫。

也就是說，決定你是否罹病的因素，除了你所遺傳到的基因外，還有你生活的外在條件。因此才会有這麼一句時常聽到的話：「**基因將子彈上膛，而扣下扳機的是外在環境。**」

諸如糖尿病、癌症或心臟病，這類疾病的基因風險因子都具有多重性，於是研究這些疾病的人，就像偵探在試著解開困難的犯罪事件一樣，努力想抓出藏在DNA裡的兇手。用來處理單一基因疾病（例如囊腫纖維症）的好策略，卻無法解決多基因疾病的難題，這真是相當遺憾的事。

研究者嘗試進行以家族為基礎的分析策略，卻接二連三失敗，於是開始嘗試其他捷徑。其中包括了所謂的候選基因策略——科學家試著臆測人類的2萬個基因裡面，有哪些基因與特定疾病有關，然後努力在病患的這些基因中找尋變異。

你也許聽過這則笑話，深夜暗巷中，有個傢伙的鑰匙從口袋掉了出來。那人過了好一陣子，才發現鑰匙不見，於是開始尋找。那老兄一直定點在路燈正下方尋找，怎麼也找不到。他的同伴覺得很奇怪，於是問他為什麼要在同一個地方打轉，他解釋說：「誰都曉得，如果沒有開燈，是找不到鑰匙的呀！」說來遺憾，候選基因策略往往面臨一樣的宿命——而從中我們並沒有找到答案。

科學家遭受到的挫折愈積愈多，直至2003年，他們找到的常見疾病（例如第二型糖尿病）基因風險因子，幾乎是微乎其微。但我們手上不是已經有基因組的序列了嗎？為什麼我們沒辦法更有系統的研究呢？為什麼我們沒辦法讓整個暗巷都亮起來？

就讓我們用想像力來做個實驗好了。假設你手上有1,000名糖尿病患的DNA，以及1,000名正常人的DNA。假設現在你已知這2,000人的完整基因組序列，然後並排作比較（圖3.1），你比較的不再限於候選基因，而是完整的基因序列，但你仍必須將資料中的雜訊從有用的信號裡剔除。

有些病人也許因為擁有二或三個罕見突變而患病；有些病人則可能擁有更多的常見突變基因。假設你手上的基因序列正確無誤，那麼你應該可以發現最主要的風險因子。靠著評估基因變異出現在實驗組（糖尿病患）和對照組（非糖尿病患）的頻率，你甚至應該可以決定出每個變異的重要性。

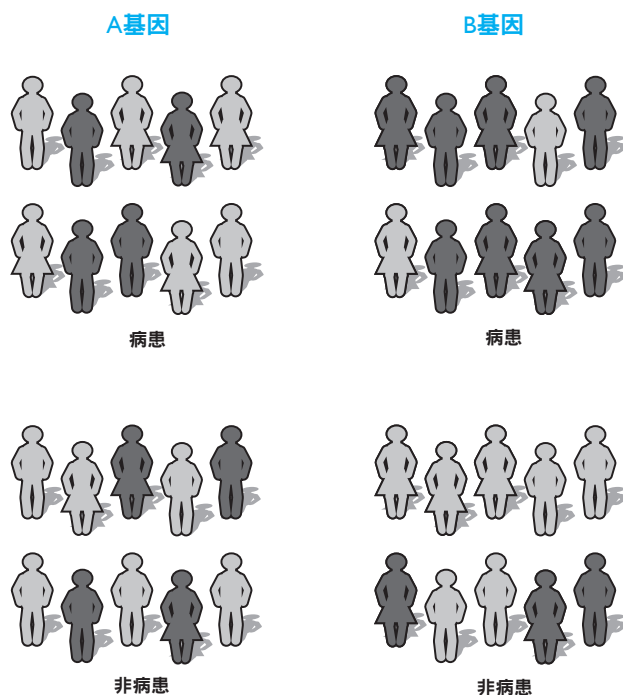


圖3.1  
想找出增加疾病風險的基因缺陷，我們必須能分辨出，有哪些DNA變異在病患身上出現的次數比非病患更頻繁（深色的人代表身上擁有該基因變異）。在這個例子裡，B基因似乎可能帶有重要的風險因子，而A基因則看似與此疾病無關。

2003年人類基因組計畫剛完成時，人們都覺得要實現上述的實驗，還需要很長的一段時間。當時我們甚至還不曉得基因組中那1千萬個常見變異的位置，更別提罕見變異了。所以，一開始專注在常見變異的研究，看似是很合理的策略。

首先，常見變異大多是每個人之間、單一鹼基對的差異，我們稱作**單核苷酸多型性**（single nucleotide polymorphism，以下簡稱SNP）。其次，2003年的時候，要在DNA樣本中解讀出其中一個基因變異的DNA序列（這個程序叫做「決定基因型」），實驗成本大約是0.5美元。因此倘若要針對實驗組的1,000人

和對照組的1,000人，逐步尋找他們基因組中所有與疾病相關的常見SNP（這個步驟稱作**全基因組關聯研究**），所需的費用是「1千萬個SNP」乘以「2,000人的DNA樣本」乘以「（每個SNP的成本0.5美元，也就是1百億美元！顯然是一件不可能的任務（這甚至不是基因組的全部定序，只是分析常見變異的DNA罷了）。

六年後我寫這本書的時候，上述在當時幾乎無法辦到的分析方法，如今已應用於數十種疾病的研究上，且分析了超過10萬人的DNA樣本，進展著實讓人震驚。我們幾乎從來沒有在任何科學領域中，看到過技術可以發展得如此迅速的情況。這樣的進展激起了大眾對於個人化醫療的熱烈期待，這也是為什麼我會著手寫書的主要原因。

能有這些進展，有一部分要感謝一個叫做**人類基因組單體型圖**（以下簡稱HapMap）的研究計畫。圖3.2列出大約2,000個字母的DNA密碼。幾乎所有人都擁有這段DNA，序列跟圖中大部分字母一樣，但常見變異會出現在其中的三個位置。而這只是人類基因組中，1千萬個常見SNP中的3個。

不過，倘若你想尋找某個疾病，例如糖尿病的致病因子，你不需要針對這1千萬個SNP，一個一個進行分析。科學家後來發現，這些SNP原來是群居動物，它們在演化過程中都是成群移動的。

很明顯的，相較於其他生物，人類的歷史較短，最初的祖先數目也並不龐大，這使得人類擁有的染色體種類十分有限。我們身上那1千萬個常見變異，並不是互不相干的。相反的，這些變異往往與鄰近的其他變異串聯在一起，我們稱此情況為**單體型**（haplotype）。

因此，一旦我們了解一兩個變異，便能預測同一區塊的其他變異會如何表現。有些區塊的範圍很小，有些則橫跨長長的一段DNA。平均來說，一個區塊通常包含了30或40個SNP，這些SNP往往是一同行動的。

倘若你知道區塊的邊界，而且也謹慎的選擇具有代表性的一組SNP，那麼

GAAATAATTAATGTTTTTCCTTCTCCTATTTTTGTCTTTACTTCAATTTATTTATTT  
ATTATTAATATTATTTATTTTTTGAGACGGAGTTTCACTCTTGGTCCAACTGGAGTGCA 1  
GTGGCGTGATCTCAGTCACTGCACACTCCGCTTTC  $\begin{matrix} C \\ T \end{matrix}$  GGTTC AAGCGATTCTCCTGC  
CTCAGCCTCCTGAGTAGCTGGACTACAGTACACACCCACACGCCCGGCTAATTTTTGT  
ATTTTTAGTAGAGTTGGGGTTTACCACATGTTGGCCAGACTGGTCTCGAACTCCTGACCTT  
GTGATCCGCCAGCCTCTGCCTCCCAAAGAGCTGGGATTACAGGCGTGAGCCACCGCGCTC  
GGCCCTTTGCATCAATTTCTACAGCTTGTTTTCTTTCCTGGACTTACAAGTCTTACCT  
TGTCTGCCTT CAGATATTTGTGTGGTCTCATTCTGGTGTGCCAGTAGCTAAAAATCCAT 2  
GATTTGCTCTCATCCCACCTCTGTGTTCATCTCCTCTTATCTGGGGTCA  $\begin{matrix} A \\ G \end{matrix}$  TATCTC  
TTCGTGATTGCATTTGATCCCAGTACTTAGCATGTGCGTAACAACCTCTGCCCTCTGCTT  
TCCAGGCTGTTGATGGGGTGCTGTTTCATGCCTCAGAAAAATGCATTGTAAGTTAAATTA  
TTAAAGATTTAAATATAGGAAAAAGTAAGCAAACATAAGGAACAAAAAGGAAAGAAC  
TGTATTCTAATCCATTATTTATTATACAATTAAGAAATTTGGAACTTTAGATTACACTG  
CTTTTAGAGATGGAGATGTAGTAAGTCTTTTACTCTTTACAAAATACATGTGTTAGCAAT  
TTTGGGAAGAAATAGTAACTCACCCGAACAGTGAATGTGAATATGTCACTTACTAGAGGA  
AAGAAAGCCTTGAAAAACATCTCTAAACCGTATAAAAACAATTACATCATAATGATGAA  
AACCAAGGAATTTTTTTAGAAAAATTACCAGGGCTAATAACAAAGTAGAGCCACATGT  
CATTATCTTCCCTTGTGTCTGTGTGAGAAATCTAGAGTTATATTTGTACATAGCATGG  
AAAAATGAGAGGCTAGTTTATCAACTAGTTCATTTTTAAAAGTCTAACACATCCTAGGTA  
TAGTGAACTGTCCCTCGCCAATGTATTGCACATTTGTGCCAGATCCAGCATAGGGTA  
TGTTTGCCATTTACAAACGTTTATGTCTTAAGAGAGGAAATATGAAGAGCAAAACAGTGC  
ATGCTGGAGAGAGAAAGCTGATACAAATATAAATGAAACAATAATTGGAAAAATTGAGAA  
ACTACTCATTTTCTAAATTAATCATGTATTTTCTAGAAATTAAGTCTTTTAAATTTTGA  
TAAATCCCAATGTGAGACAAGATAAGTATTAGTGATGGTATGAGTAATTAATATCTGTTA  
TATAATATTCATTTTCATAGTGGAAGAAATAAAATAAAGGTTGTGATGATTGTTGATTAT  
TTTTTCTAGAGGGTGTGTCAGGGAAGAAATGCTTTTTTTTCTCTCTTTCCACTAA  
GAAAGTTCAACTATTAATTTAGGCACATACAATAATTACTCCATTCTAAAAATGCCAAAA  
GGTAATTTAAGAGACTTAAACTGAAAAGTTTAAGATAGTCACACTGAACATATATTA  
AATCCACAGGGTGGTTGGAAC TAGGCCTTATATTAAGAGGCTAAAAATGCAATAAGAC  
CACAGGCTTTAAATATGGCTTTAAACTGTGAAAAGGTGAAACTAGAATGAATAAAAATCCTA 3  
TAAATTTAAATCAAAAGAAAGAAACAACT  $\begin{matrix} A \\ G \end{matrix}$  AAATTAAGTTAATATACAAGAAATG  
GTGGCTGGATCTAGTGAACATATAGTAAAGATAAAACAGAAATTTCTGAAAAATCCTG  
GAAATCTTTTGGGCTAACCTGAAAAACAGTATATTTGAAACTATTTTAAACCGAGTTAT  
GGCACACTTGGGCAATTT CAGAGATT

方框內的是三個常見變異，也就是SNP。

但是後來發現，這些變異之間有很強的相關性：SNP 1的C鹼基，總是與SNP 2的A鹼基以及SNP 3的G鹼基同時出現。

圖3.2 DNA密碼中的2,000個字母（只列出單股序列）。



就算不在實驗室灑大錢、檢測每一個 SNP，你還是可以得到相當充足的資訊。

HapMap 計畫的目標，就是為了界定各個區塊的邊界，如此便能為基因組分析提供一條捷徑，讓工作量減少到原本的 1/40。我有幸成為 HapMap 的計畫主持人，這個國際型計畫很有組織、進展飛速，研究團隊包含了來自六個國家的 2,000 名科學家，大家一同努力羅列出人類基因中的變異，並將這些變異分類成不同區塊，迅速把這些資訊免費供給大眾。

2003 年之後的另一項巨大進展，則是基因型分析成本的驟降，從 0.5 美元，降至不到 0.01 美元。這受惠於許多很有創意的新技術，其中大多數的技術都結合了電腦晶片與 DNA 化學而製造出的「DNA 晶片」，多達 1 百萬個 SNP 可以塞進郵票般大小的陣列晶片中，然後用在檢驗上。

在 HapMap 計畫的幫助下，檢測所需的 SNP 數目減少了，再加上基因型分析成本的劇降，到了 2006 年，只要不到 1 百萬美元的成本，便可以針對 1,000 個實驗組加 1,000 個對照組個體，進行全基因組關聯研究。你們看看，短短幾年間，就有這麼大的進展！

## 第一個成功的案例：黃斑病變

我的姑姑瑪莎是個很特別的人。她聰明、頑固、又學富五車，後來當上一所私立學校的女校長，學生對她既敬又畏。而我身為她的姪子和教子，對她也是抱持一樣的態度。她是很棒的老師，全心奉獻給學生，但她那強烈的性格，和如同名廚茱莉亞·柴爾德（Julia Child）聲如洪鐘的嗓音，足夠讓一個小男孩嚇破了膽。

讓人難過的是，她生命中最重要娛樂之一——閱讀，就在她退休後遭受嚴重阻礙，因為她的視力逐漸退化，後來診斷為老年性黃斑病變。她在將近八十歲時開始發病，晚年視力幾乎完全喪失，殘忍的疾病剝奪了她閱讀的樂趣。

你的視力退化了嗎？